Reviewer Comments:

Reviewer: 1

Recommendation: RQ - Review Again After Major Changes

Comments:
The manuscript compares different (partial) acoustic channel equalization techniques for speech dereverberation. To improve the robustness of these techniques to channel estimation errors, which are unavoidable in practical implementations, regularization is proposed and investigated for the different techniques. For implementing the regularization in practical scenarios, a non-intrusive way to estimate the regularization parameter is proposed.

Experiments in a simulated test environment based on measured room impulse responses compare the different approaches with and without regularization both with respect to the dereverberation performance measured by energy decay curves and with respect to perceptual quality measured by PESQ scores. The experimental results show that the P-MINT approach, recently proposed by the authors, achieves the best perceptual performance when regularization is applied, while the dereverberation performance of the CS and RMCLS approaches, previously proposed by other authors, is slightly better than that of P-MINT in most cases.

The manuscript is very well written, the equations and symbols are very well explained, and the notation is very well structured so that the paper is a pleasure to read. The overview of channel equalization tasks in II.B - II.D is concise and very insightful paving the way for the following discussion. The reviwer considers the regularized version of the P-MINT algorithm with the non-intrusive approach for estimating the regularization parameter as a novel technique that represents an important step for making channel equalization techniques ready for practical applications. Therefore the paper is almost ready for publication in T-ASL. Since there are a few important open points (mainly regarding the performance in real-world scenarios), the reviewer proposes to review the manuscript againg after the following revisions:

1. So far, the manuscript neglects background noise, which is usually recorded by the microphones in addition to reverberation. Since the robustness of channel equalization to background noise is extremely important in practical scenarios, background noise should be included in the paper at the following points:
a) In equation (1)
b) In the discussion of regularization, equations (40) - (42)
c) There should be at least one experiment including realistic background noise (recorded noise, not white noise)

2. The experimental results indicate that the performance gain obtained by the channel equalization technique strongly depend on the channel estimation error (as could be expected). For the experiments, scaled white noise is added to the true RIRs as suggested in [32] at a normalized channel mismatch Em of -33dB and -15dB. However it is not clear how realistic these conditions are in practical scenarios. To give the reader an idea of how the channel equalization techniques work in real-world scenarios either experiments with RIRs obtained by state-of-the-art channel identification

techniques should be added or at least it should be discussed whether the conditions described above(scaled white noise at Em of -33dB or -15dB) can be achieved by current channel identification algorithms.

Further details:
3. Page 2, paragraph 3 of the abstract:
"... outperforms all other intrusively regularized ... techniques"
should be replaced by
"... outperforms all other considered intrusively regularized ... techniques"

Similar in the second paragraph of the conclusion (page 20).

4. Lh=2000 appears to be very short for a reverberation time of 600ms. To cover at least 50% of the reverberation time would require Lh=4800 at fs=16kHz for a reverberation time of 600ms. Therefore, the simulations should be performed with RIRs of length Lh >= 4800 and corresponding Lg. If such long RIRs are a problem to any of the algorithms, then you should at least use Lh >= 4800 for generating (some of) the microphone signals and investigate the difference to Lh=2000.

5. Figure 4 and Figure 6: The best performance with regularized P-MINT is achieved with the longest considered desired window length of Ld=50ms. Therefore please add tests with longer windows to see were the maximum performance is obtained. Fruthermore, please provide color diagrams.

6. Figure 8: Please add the PESQ score for the unprocessed micophone signal x_1(n) to the plots.

7. Figure 5: From the reviewers point of view, there is no need for the vertical axis to start at -60dB. The plots would be better readable if the vertical axis started at -30dB.

Additional Questions:
1. Is the topic appropriate for publication in these transactions?: Yes

2. Is the topic important to colleagues working in the field?: Yes

Explain:

1. Is the paper technically sound?: Yes

why not?:

2. Is the coverage of the topic sufficiently comprehensive and balanced?: Yes

3. How would you describe technical depth of paper?: Appropriate for the Generally Knowledgeable Individual Working in the Field or a Related Field

4. How would you rate the technical novelty of the paper?: Novel

1. How would you rate the overall organization of the paper?: Satisfactory

2. Are the title and abstract satisfactory?: Yes

Explain:

3. Is the length of the paper appropriate? If not, recommend how the length of the paper should be amended, including a possible target length for the final manuscript.: Yes

4. Are symbols, terms, and concepts adequately defined?: Yes

5. How do you rate the English usage? : Satisfactory

6. Rate the Bibliography: Satisfactory

null:

1. How would you rate the technical contents of the paper?: Good

2. How would you rate the novelty of the paper?: Sufficiently Novel

3. How would you rate the "literary" presentation of the paper?: Totally Accessible

4. How would you rate the appropriateness of this paper for publication in this IEEE Transactions?: Excellent Match

Reviewer: 2

Recommendation: RQ - Review Again After Major Changes

Comments:

This paper discusses the multi-channel dereverberation techniques, focusing on the robustness issue of the inverse filter. To make the multi-channel inverse filter more robust to the RIR estimation errors, they proposed mainly the following two things in the paper:

1.
MINT-like dereverberation technique that suppresses only the late reverberation in the observed signal (referred to as P-MINT)

2.
Regularization for P-MINT and two other conventional methods proposed by others. The regularization is done in very similar way to the conventional method ([15] in the paper).

In the paper, they also shortly discuss the incorporation of automatic determination of the regularization parameters.

Although, according to the experimental results, we can see the effectiveness of the above approaches, the paper has to be improved in the following points to make it more technically sound.

1.
It is not clear why the inverse filter becomes robust if it equalizes only the late reverberation part of RIR. Please show clearly based on which equations/operations/nature we can expect it becomes robust to RIR estimation error. It is generally required to have such theoretical validation in addition to the experimental results to support the claim of the paper.

2.
In the experimental section, they compared MINT, Relaxed multi-channel least squares (RMCLS), channel shortening (CS), P-MINT and their regularized versions, and showed the superior performance of RMCLS in terms of energy decay curve (EDC). It is not clear to me why RMCLS is superior to the others. Please add more technical discussions to the experimental section that can explain these differences in performance. It is also curious to see how the performance of the regularized RMCLS and CS may change if the regularization parameters are selected not from the view point of perceptual quality, but from the view of reverberation suppression rate.

3.
Please add the PESQ score of the ideally dereverberated signal for the reference purpose. It can be realized for example as the convolution of clean signal and the first part of RIR (\mathbf{h}_{p}^{d}). By doing so, we can see how close the performances of the discussed dereverberation methods are to ideal one.

4.
Please remove or modify the last paragraph of the section V-B (the one starting with "Summarizing the simulation results..."). Although it is mentioned as if the regularized P-MINT is the best among the compared methods, we can see that it is just not true, for example, by taking a look at Fig.3, where we can see that RMCLS suppresses the late reverberation more effectively.

5.
Please add some comments on the estimation accuracy for the regularization parameter. We can see that it is working reasonably well by observing PESQ score (Fig. 7). In addition to the PESQ score,

it may be helpful for readers to understand the effectiveness of the proposed scheme, if the paper includes how close the value itself is to the intrusively selected parameter.

Additional Questions:
1. Is the topic appropriate for publication in these transactions?: Yes

2. Is the topic important to colleagues working in the field?: Moderately So

Explain:

1. Is the paper technically sound?: No (why not?)

why not?: Please see my comments.

2. Is the coverage of the topic sufficiently comprehensive and balanced?: Treatment somewhat unbalanced, but not seriously so.

3. How would you describe technical depth of paper?: Appropriate for the Generally Knowledgeable Individual Working in the Field or a Related Field

4. How would you rate the technical novelty of the paper?: Somewhat Novel

1. How would you rate the overall organization of the paper?: Satisfactory

2. Are the title and abstract satisfactory?: Yes

Explain:

3. Is the length of the paper appropriate? If not, recommend how the length of the paper should be amended, including a possible target length for the final manuscript.: Yes

4. Are symbols, terms, and concepts adequately defined?: Yes

5. How do you rate the English usage? : Satisfactory

6. Rate the Bibliography: Satisfactory

null:

1. How would you rate the technical contents of the paper?: Fair

2. How would you rate the novelty of the paper?: Slightly Novel

3. How would you rate the "literary" presentation of the paper?: Totally Accessible

4. How would you rate the appropriateness of this paper for publication in this IEEE Transactions?: Excellent Match

Reviewer: 3

Recommendation: AQ - Publish In Minor, Required Changes

Comments:
The authors proposed a regularized partial multichannel equalization technique based on the multiple-input/output inverse theorem (MINT) for speech de-reverberation. The proposed algorithm shows improved robustness in the presence of RIR estimation error and gives better PESQ score for the dereverberated speech. Though the idea is not entirely novel (the partial MINT is already proposed by the authors in [19], the associated regularization term is not new [15], automatic computation of the regularization parameter already exists in [26]), the paper is well organized and the results are presented nicely. However, the reviewer has the following concerns to be considered before publication.

1. The authors have considered the estimation error in the RIRs but ignored the additive noise (with the clean speech signal) that causes the estimation error (see Eq. 42). Even if we are given the true RIR, the output signal obtained from the MINT equalization filters is severely distorted due to the additive noise. The regularization term may be robust against the additive noise but it should be demonstrated in the paper.
2. In order to simulate estimation errors, the measured RIRs have been perturbed by adding scaled white noise as proposed in [32]. However, noise robust RIR estimation technique from the noisy speech signal has been already proposed in the literature (see [18] and the ref. therein). These techniques can be used to simulate more realistic scenario.
3. The RMCLS algorithm produces better PESQ score than the P-MINT without the regularization term. The performance of RMCLS does not change even if the regularization term is incorporated (Fig. 4). However, the performance of regularized P-MINT is significantly better than the P-MINT. Will you please explain that?
4. Can you please demonstrate how the performance of regularized P-MINT varies with selection of different acoustic channels in Eq. 29?

Additional Questions:
1. Is the topic appropriate for publication in these transactions?: Yes

2. Is the topic important to colleagues working in the field?: Moderately So

Explain:

1. Is the paper technically sound?: Yes

why not?:

2. Is the coverage of the topic sufficiently comprehensive and balanced?: Yes

3. How would you describe technical depth of paper?: Appropriate for the Generally Knowledgeable Individual Working in the Field or a Related Field

4. How would you rate the technical novelty of the paper?: Somewhat Novel

1. How would you rate the overall organization of the paper?: Satisfactory

2. Are the title and abstract satisfactory?: Yes

Explain:

3. Is the length of the paper appropriate? If not, recommend how the length of the paper should be amended, including a possible target length for the final manuscript.: Yes

4. Are symbols, terms, and concepts adequately defined?: Yes

5. How do you rate the English usage? : Satisfactory

6. Rate the Bibliography: Satisfactory

null:

1. How would you rate the technical contents of the paper?: Good

2. How would you rate the novelty of the paper?: Slightly Novel

3. How would you rate the "literary" presentation of the paper?: Totally Accessible

4. How would you rate the appropriateness of this paper for publication in this IEEE Transactions?: Good Match