Least Squares Phase Estimation of Mixed Signals

Carlos Eduardo Cancino Chacón[†‡] *and P. Mowlaee*[†]

[†]Signal Processing and Speech Communication Lab., Graz University of Technology, Graz, Austria
[‡]Austrian Research Institute for Artificial Intelligence, Vienna, Austria
cancinochacon@student.tugraz.at   pejman.mowlaee@tugraz.at

Abstract

Estimating the phase of sinusoids in noise in contrast to amplitude and frequency components has been less addressed in previous studies. In this paper, we derive the least squares phase estimator (LSPE) solution to recover the phase of an underlying signal observed in noise. Through Monte-Carlo simulations, we demonstrate the robustness of the proposed phase estimator against the modeling error. The proposed phase estimator is further evaluated in the speech enhancement setup to asses how much improvement is obtained by replacing the noisy phase with the LSPE when reconstructing the enhanced speech signal. Significant improvement in speech quality and speech intelligibility is obtained by replacing the noisy phase with the estimated phase provided by the proposed LSPE once the ambiguity in the phase candidates is removed.

Index Terms: Least squares, phase estimation, speech enhancement, speech quality.

1. Introduction

In estimation theory and signal detection, the ultimate goal is to find optimal estimators to recover a sinusoid observed in noise. In particular, we are interested to find estimates for the triple parameters: amplitude, frequency and phase for sinusoidal components of a signal of interest [1]. Such a general problem definition holds for a wide range of signal and audio processing applications, to name a few: signal detection [2], speech enhancement [3] and radar signal processing [4].

While there are many previous studies focused on deriving estimators of the parameters of sinusoid observed in noise [1,5], the issue of estimating the phase of the signal in noise has been often neglected. For example, in speech enhancement, the phase spectrum has been considered to be unimportant [6, 7] as the noisy phase was shown to provide the optimal estimate for clean phase for the signal components of high signal to noise ratio (SNR) [8, 9]. This optimality is only valid under the independence assumption of the STFT signal coefficients which is not always true in practice. Therefore, the choice of noisy phase limits the achievable performance by the existing signal enhancement methods. In fact, the importance of phase spectrum has been reported with positive results in the speech quality [10–18], speech intelligibility [19], speech recognition [20], and other signal processing applications [7] (see [21] for an overview). These findings motivate us to find estimators for the phase of the sinusoidal components observed in noise.

In this paper, we propose a least squares (LS) solution for estimating the phase of the underlying signals from a mixed signal. The derivations presented here require no prior assumption on the distribution on the underlying signals, in contrast to the conventional estimators where spectral amplitude and phase are assumed to be independent. By means of Monte-Carlo simulations on the synthetic sinusoids in noise, the proposed estimator is evaluated in terms of the mean square error (MSE) and an unwrapped phase distortion metric proposed in [22]. This estimator is eventually tested on real speech signals in the context of speech enhancement and is shown to improve the perceived quality as well as speech intelligibility of the signals using noisy phase for different noise scenarios and signal-to-noise ratios.

2. Problem formulation

Consider a single channel mixed signal can be modeled, using a physical interpretation, as the sum of two signals coming from different sources (positioned at $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively). The received mixed signal is given by

$$y(t) = s_1(\mathbf{r}_1, t) + s_2(\mathbf{r}_2, t). \tag{1}$$

Taking the Fourier transform with respect to the time t from Eq. (1), we obtain

$$Y^c(\omega) = S_1^c(\mathbf{r}_1, \omega) + S_2^c(\mathbf{r}_2, \omega), \tag{2}$$

where $Y^c, S_1^c, S_2^c \in \mathbb{C}$, can be represented in a polar form as follows:

$$Y(\omega)e^{j\phi_y(\omega)} = S_1(\omega)e^{j\phi_1(\omega)} + S_2(\omega)e^{j\phi_2(\omega)}, \tag{3}$$

where ϕ_y and ϕ_i denote the phase of the mixed signal and the underlying ith signal in the mixture with $i \in \{1, 2\}$ while $S_i(\omega)$ represents the ith signal magnitude with $Y(\omega)$ as the mixture magnitude spectrum at frequency ω. Throughout this paper, for simplicity we will use $Y(\omega) = Y$, $S_i(\omega) = S_i$ and $\phi_i(\omega) = \phi_i$. The signal + noise interaction in Eq. (3) can be geometrically interpreted as shown in Figure 1, where Y^c is the sum of vectors S_1^c and S_2^c in the complex plane.

We are interested in computing the phase of S_i given the noisy phase ϕ_y and good estimations of the magnitude spectra[1] S_i and mixed signal observations Y and ϕ_y. In this paper, we derive the Least Squares Phase Estimator (LSPE) only for ϕ_1. Without loss of generality, an extension for the proposed LSPE is straightforward for the second source phase (ϕ_2). The problem can be stated as finding the phase that minimizes the difference between Y^c and the sum of signals S_1^c and S_2^c in the least squares sense, given by

$$d_{SE}(\hat{\phi}_1) = |Y^c - (S_1^c + S_2^c)|^2. \tag{4}$$

---

[1]estimate of signals magnitude spectrum S_i are provided using one of the plethora of methods extensively studied for amplitude estimation from a mixed signal in e.g. speech enhancement literature [3].
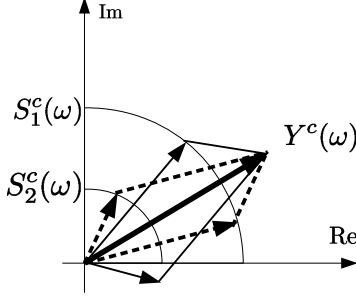
Figure 1: Representation of Y^c as the sum of S_1^c and S_2^c in the complex plane. Due to the parity in the sign of $\sin(\Delta\phi)$ in Eq. (21), there is an ambigiuity in the set of phase candidates.

Nevertheless, as noted in [22], an additional optimality criterion on the underlying phase can be incorporated via combining the cost function Eq. (4) with an unwrapped cosine distortion as

$$d_{cos}(\phi_1, \hat{\phi}_1) = 1 - \cos(\phi_1 - \hat{\phi}_1). \tag{5}$$

The distortion metric, in contrast to the mean square error (MSE) criterion, is invariant under modulo 2π transformation of the estimation error. For small estimation errors, it resembles the squared error distortion criterion[2]. By doing this, the resulting estimate does not affect the estimation of the spectral magnitude [22]. Using the method of Lagrangian multipliers, this constrained problem can be formulated as minimizing the Lagrangian function

$$\Lambda(\hat{\phi}_1, \lambda) = |Y e^{j\phi_y} - S_1 e^{j\hat{\phi}_1} - S_2 e^{j\phi_2}|^2 - \lambda(1 - \cos(\phi_1 - \hat{\phi}_1)), \tag{6}$$

and the resulting LS phase estimate is given by

$$\phi_1^{ls} = \arg\min_{\hat{\phi}_1} \Lambda(\hat{\phi}_1, \lambda). \tag{7}$$

3. Least Squares Phase Estimation

3.1. Closed-form Solution

From Eq. (3), the magnitude of Y^c is given by

$$Y = \sqrt{S_1^2 + S_2^2 + 2S_1 S_2 \cos(\Delta\phi)}, \tag{8}$$

where $\Delta\phi = \phi_2 - \phi_1$, is the phase difference. From the above equation, $\Delta\phi$ is given by

$$\Delta\phi = \pm\arccos\left(\frac{Y^2 - S_1^2 - S_2^2}{2S_1 S_2}\right), \tag{9}$$

which has ambiguity in sign due to the lack of knowledge regarding the phase-lead or phase-lag relationship between the underlying phase values (see Figure 2). Therefore, there is always a set of two phase candidates satisfying all the observations in the least squares sense (see also [23] for a detailed derivation of the two sets). Using the phase difference, the mixed signal in the Fourier domain can be written as

$$Y e^{j\phi_y} = \left(S_1 + S_2 e^{j\Delta\phi}\right) e^{j\phi_1}. \tag{10}$$

---

[2]as we have: $1 - \cos(\phi_1(\omega) - \hat{\phi}_1(\omega)) \approx \frac{(\hat{\phi}_1(\omega) - \phi_1(\omega))^2}{2}$ for $\hat{\phi}_1(\omega) - \phi_1(\omega) \ll 1$.

Solving Eq. (10) for ϕ_1 we obtain

$$\phi_1 = \phi_y + j\log\left(\frac{S_1 + S_2 e^{j\Delta\phi}}{Y}\right). \tag{11}$$

Substituting $\Delta\phi$ in Eq. (7) and using that the absolute value of $z^c \in \mathbb{C}$ is given by $z = \sqrt{z^c z^*}$, with z^* representing the complex conjugate of z^c, and that $2\cos(x) = e^{jx} + e^{-jx}$, the Lagrangian function can be rewritten as

$$\begin{aligned}\Lambda(\hat{\phi}_1, \lambda) =\ & Y^2 + S_1^2 + S_2^2 - 2S_1 Y \cos(\phi_y - \hat{\phi}_1) \\ & - 2S_2 Y \cos(\phi_y + \Delta\phi - \hat{\phi}_1) \\ & + 2S_1 S_2 \cos(\Delta\phi) - \lambda(1 - \cos(\phi_1 - \hat{\phi}_1)).\end{aligned} \tag{12}$$

In order to find the minimum of the Lagrangian function, it is necessary to find $\hat{\phi}_1$ and λ at which the gradient of the Lagrangian vanishes, i.e. $\nabla_{\hat{\phi}_1, \lambda} \Lambda = \mathbf{0}$. The derivative with respect to $\hat{\phi}_1$ is given by

$$\begin{aligned}\frac{\partial\Lambda}{\partial\hat{\phi}_1} =\ & -2S_1 Y\left(\sin(\phi_y)\cos(\hat{\phi}_1) - \sin(\hat{\phi}_1)\cos(\phi_y)\right) \\ & 2S_2 Y\left(\sin(\phi_y + \Delta\phi)\cos(\hat{\phi}_1) - \sin(\hat{\phi}_1)\cos(\phi_y\Delta\phi)\right) \\ & + \lambda\left(\sin(\phi_1)\cos(\hat{\phi}_1) - \sin(\hat{\phi}_1)\cos(\phi_1)\right) = 0,\end{aligned} \tag{13}$$

and the derivative with respect to the multiplier λ is

$$\frac{\partial\Lambda}{\partial\lambda} = 1 - \cos(\phi_1 - \hat{\phi}_1) = 0. \tag{14}$$

By grouping the terms with $\sin(\hat{\phi}_1)$ and $\cos(\hat{\phi}_1)$ from Eq. (13), $\hat{\phi}_1$ can be expressed as

$$\tan(\hat{\phi}_1) = \frac{A - \lambda B}{C - \lambda D}, \tag{15}$$

with

$$\begin{aligned}A &= 2S_1 Y \sin(\phi_y) + 2S_2 Y \sin(\phi_y + \Delta\phi) \\ B &= \sin(\phi_1) \\ C &= 2S_1 Y \cos(\phi_y) + 2S_2 Y \cos(\phi_y + \Delta\phi) \\ D &= \cos(\phi_1).\end{aligned} \tag{16}$$

Substituting (15) in (14) gives

$$\phi_1 - \arctan\left(\frac{A - \lambda B}{C - \lambda D}\right) = 2\pi m, \tag{17}$$

with $m \in \mathbb{Z}$, i.e. an integer number. Solving for λ results in

$$\lambda = \frac{C\tan(\phi_1) - A}{D\tan(\phi_1) - B}. \tag{18}$$

By substituting Eq. (18) in Eq. (15) we obtain

$$\tan\hat{\phi}_1 = \tan\phi_1. \tag{19}$$

Using Eq. (11) and replacing ϕ_1 in (19), the LSPE is given by

$$\boxed{\phi_1^{ls} = \phi_y + j\log\left(\frac{S_1(\omega) + S_2(\omega)e^{j\Delta\phi}}{Y(\omega)}\right),} \tag{20}$$

2

where $e^{j\Delta\phi} = \cos(\Delta\phi) + j\sin(\Delta\phi)$ with

$$\sin(\Delta\phi) = \pm\sqrt{1 - \cos^2(\Delta\phi)}. \qquad (21)$$

3.2. Limit cases

In order for the LSPE to be real valued, the following constraint defined on Eq. (20) should be satisfied

$$\text{Re}\left(\log\left(\frac{S_1 + S_2(\cos(\Delta\phi) + j\sin(\Delta\phi))}{Y}\right)\right) = 0, \qquad (22)$$

Using the fact that $\log(z^c) = \log(z) + j\phi_z$, (22) can be rewritten as

$$\frac{S_2\sin(\Delta\phi)}{Y} = \pm\sqrt{1 - \left(\frac{S_1 + S_2\cos\Delta\phi}{Y}\right)^2}. \qquad (23)$$

By substituting Eq. (9) in the above equation, it can be seen, that this condition is always met, and therefore, $\phi_1^{ls} \in \mathbb{R}$ for all ϕ_y, Y, S_1 and S_2.

When $S_2 \to 0$ i.e. contribution of the s_2 to the mixed signal y is negligible, the LSPE simplifies to

$$\lim_{S_2\to 0}\phi_1^{ls} = \phi_y, \qquad (24)$$

which is consistent with [9], and partially explains why phase estimation has been regarded as unimportant in scenarios were the amplitude has a large local SNR. Similar to the discussion in [9], we take into account the fact that a threshold SNR (denoted by SNR_{th}) exists above which the noisy phase suffices, and we are motivated to propose the following phase estimator:

$$\phi_1^{LS}(\omega) = \begin{cases} \phi_1^{ls}(\omega) & \frac{S_1^2(\omega)}{S_2^2(\omega)} < \text{SNR}_{th}, \\ \phi_y(\omega) & \text{otherwise.} \end{cases} \qquad (25)$$

The SNR_{th} is found from the Monte-Carlo simulation (see Figure 2) found equal to 6 dB which is consistent with previous findings by Vary [9].

4. Experiments

To demonstrate the effectiveness of the proposed LSPE method, in this section we consider two experiments: 1) Monte-Carlo simulation for synthetic signals and 2) real speech signals explained in the following.

4.1. Monte-Carlo Simulation

In the Monte-Carlo simulation, we consider one sinusoid in white Gaussian noise. Both frequency and phase of the sinusoid are taken from a uniform distribution $\phi \sim U[-\pi, \pi]$ while its amplitude is set to produce a signal-to-noise ratio within the range of -30 to 30 dB. The number of realizations in the Monte-Carlo simulation is set to 200. As the evaluation criterion, we first inspect the MSE phase distortion criterion $d_{MSE}(\phi_1, \hat{\phi}_1) = (\phi_1 - \hat{\phi}_1)^2$ for the LS-estimated phase versus the clean phase as the reference. The results are shown in Figure 2 where the performance obtained by the noisy phase and the two sets of LS-estimated phase values are shown and denoted by LS1 and LS2. The two LS-estimates are still ambiguous due to the parity in the sign $\sin(\Delta\phi)$ given in (21) (see Figure 2). On average, both LS1 and LS2 are similar, and therefore, they seem to overlap. To show the potential of the proposed least-square approach after incorporation of an ambi-
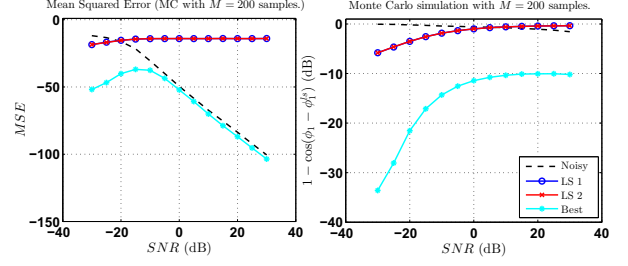

Figure 2: Monte-Carlo simulation for one sinusoid in white noise to evaluate the performance of the proposed least squares phase estimator under two distortion criteria: (left) mean square error: $d_{\text{MSE}}(\phi_1, \hat{\phi}_1) = (\hat{\phi}_x(k,l) - \phi_x(k,l))^2$, and (right) unwrapped phase distortion: $d_{\cos}(\phi_1, \hat{\phi}_1) = 1 - \cos(\hat{\phi}_x(k,l) - \phi_x(k,l))$.

guity removal method (e.g. the one explained in [23] relying on group-delay constraint[3]), the results obtained after removing the ambiguity in the phase candidates is also shown in Figure 2, denoted by "Best".

In the MSE sense, the ambiguous LS-estimated phase outperform the noisy phase for SNR ≤ -20 (dB). In contrast, the ambiguity removed LS-estimates (denoted as "Best") consistently leads to the best estimates compared to noisy and LS1 or LS2. The noisy phase asymptotates to this best result as the input SNR gets large enough (SNR ≥ -5 (dB)).

The MSE criterion on phase parameters suffers from wrapping due to phase jumps of $\pm\pi$ leading to a not necessarily a faithful metric for the evaluation of a phase estimator. This insufficiency of the MSE criterion for quantifying the distortion in phase has been studied in speech coding (see for example [24, 25]). To alleviate this, we also report the performance in terms of $d_{cos}(\phi_1, \hat{\phi}_1)$ defined in Eq. (5) which was proposed in [22] for speech enhancement and in [24] for speech coding applications. This makes the metric invariant under module 2π transformation of the estimation error term $(\phi - \hat{\phi})$.

From the unwrapped metric, we observe that the noisy phase starts outperforming LS1 and LS2 for input SNRs larger than 6 dB. This confirms the previous observation made in [9] stating that the choice of the noisy phase suffices for local signal-to-noise ratio of larger than 6 dB. However, that observation in [9] was obtained under the prior assumption that the noise probability density function is Gaussian with no restriction on noise pdf. Furthermore, the SNR $= 6$ (dB) reported in [9] was calculated from the maximum phase deviation before reaching the phase perception threshold[4].

4.2. Proof-Of-Concept Using Phasegrams

As the instantaneous phase spectrum shows a random pattern, in order to demonstrate the effectiveness of the proposed phase estimation method, here we consider group delay as the frequency derivative of phase which excludes the wrapping issue and exhibits a clear harmonic structure following that in the spectrogram. Here, we consider a male utterance corrupted with white noise at SNR $= 0$ dB. Figure 3 shows the spectrogram and group delay plots from left to right, for clean, noisy input, noisy phase reconstructed signal, and reconstructed signal using the proposed LSPE. The PESQ scores obtained by each method is also shown at the top of each panel. The proposed phase estima-

---

[3]A future study will be dedicated to fully address the impact of the errors introduced by the ambiguity removal step.

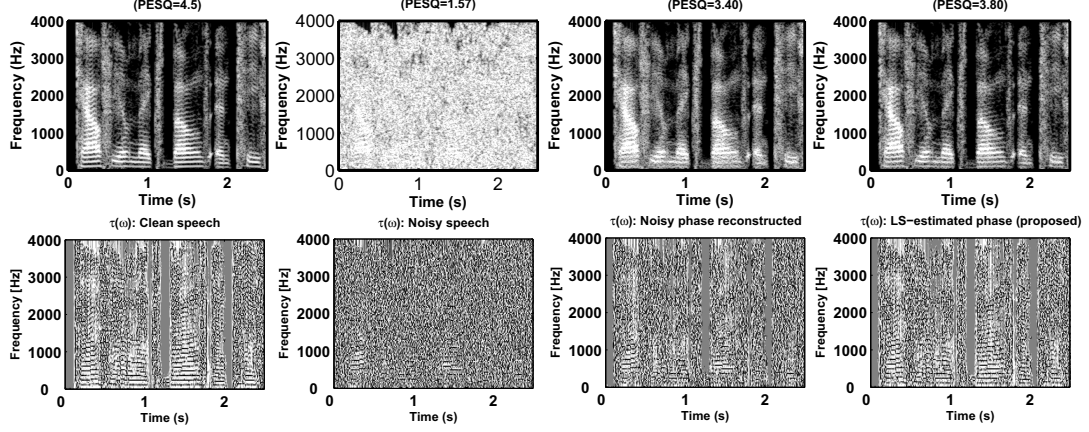[4]Phase deviation is defined as Eq. (9) in Appendix-1 in [9]

Figure 3: (Top) Spectrogram and (bottom) group delay plots for clean, noisy (unprocessed), noisy phase reconstructed signal, signal reconstructed using the LS-estimated phase. The PESQ score of each method is shown at the top of the panels.
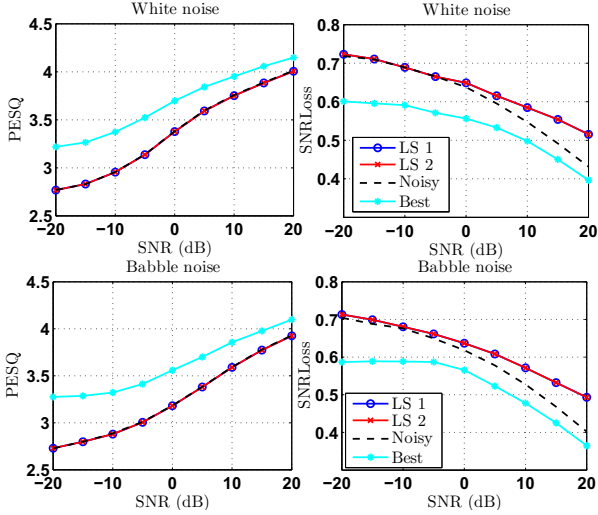

Figure 4: Speech quality and speech intelligibility results shown for noisy, LS-estimated phase with and without ambiguity removal.

tor recovers the harmonic structure existing in the clean phase spectrum, missing in the noisy signal due to added noise.

4.3. Realistic Noisy Speech Signals

The proposed phase estimation algorithm can be applied to any single channel speech enhancement method where amplitude and frequency are estimated. To demonstrate the effectiveness of the proposed phase estimation method, we further extend the experiments to realistic speech signals considered as sum of sinusoids. The goal is to justify how much improvement is achievable via replacing the noisy phase with the estimated phase when reconstructing the enhanced speech signal. The amplitude and frequency of the underlying speech signal are assumed to be given[5]. Here we choose 100 sinusoids to represent the underlying speech signal at each frame. We select 20 utterances from the TIMIT database [26] corrupted with white and babble noise both selected from NOISEX-92 [27] at global SNR ranging within -20 to 20 dB with 5 dB step. The signals are

---

[5]Future work will report the phase estimation results for a blind scenario where amplitude and frequency are estimated using the conventional speech enhancement methods in the literature [8].

downsampled to 8 kHz. As our frame setup a Hanning window of length 256 (32 ms) was used with the frameshift of 10 ms. To quantify the obtained improvement in the speech enhancement performance, we consider two metrics: ITU standard in [28] for perceptual evaluation of speech quality (PESQ) and SNR-loss [29] for speech intelligibility. Both were shown to have the high correlation with subjective listening results [30]. We note that a lower SNRloss indicates better speech intelligibility.

The results are shown in Figure 4 categorized to white and babble noise scenarios. Significant improvement in both speech quality and speech intelligibility is obtained by the LS-estimated phase followed by ambiguity removal. In terms of PESQ, the LS-estimated phase with ambiguity (denoted by LS1 and LS2), performs on average the same as the noisy phase since they are equal to the noisy phase plus an additive term as a function of $\Delta\phi$ (as seen in Eq. (20)). In terms of speech intelligibility, the remaining ambiguity in the LS-estimated phase results in information loss, hence degrades the speech intelligibility compared to the input speech signal while after removing the ambiguity in the LS-phase candidates, significant improvement is obtained in intelligibility compared to noisy phase, justified by a lower score of SNRloss. Both observations are consistent with previous observation that the sign of the phase information largely contributes to speech intelligibility [31].

5. Conclusion

In this paper, a least squares (LS) solution is derived for estimating the phase spectrum of an underlying signal corrupted in noise. The proposed solution makes no prior assumption on the probability density of the desired source or interfering noise signals. Throughout Monte-Carlo simulations and realistic speech enhancement experiments, it was demonstrated that the proposed least-squares phase estimator provides lower phase distortion in Monte-Carlo simulations as well as significant improvement speech quality and speech intelligibility compared to when noisy phase spectrum is chosen for signal reconstruction stage in single-channel speech enhancement.

It is a known result that least squares (without any constraint or regularization) is equivalent to maximum likelihood estimation in presence of Gaussian Noise. Nevertheless, the proposed LS phase estimator (LSPE) is a more general approach, that do not necessarily relies on the assumption of the probability distribution on the involved signals

6. References

[1] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.

[2] H. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968, no. Part I.

[3] P. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton: CRC Press, 2007.

[4] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proceedings of the IEEE*, vol. 86, no. 2, pp. 421–441, Feb 1998.

[5] H. Li, P. Stoica, and J. Li, "Computationally efficient parameter estimation for harmonic sinusoidal signals," *Signal Processing*, vol. 80, no. 9, pp. 1937 – 1944, 2000.

[6] D. Wang and J. Lim, "The unimportance of phase in speech enhancement," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 30, no. 4, pp. 679–681, 1982.

[7] A. Oppenheim and J. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529 – 541, May 1981.

[8] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, no. 6, pp. 1109–1121, Dec 1984.

[9] P. Vary, "Noise suppression by spectral magnitude estimation mechanism and theoretical limits," *Signal Processing*, vol. 8, no. 4, pp. 387 – 400, 1985.

[10] P. Mowlaee and R. Saeidi, "Iterative closed-loop phase-aware single-channel speech enhancement," *IEEE Signal Processing Letters*, vol. 20, no. 12, pp. 1235–1239, Dec. 2013.

[11] P. Mowlaee, R. Saeidi, and R. Martin, "Phase Estimation for Signal Reconstruction in Single-channel Speech Separation," in *Proceedings of the International Conference on Spoken Language Processing (INTERSPEECH)*, 2012.

[12] P. Mowlaee and R. Saeidi, "On phase importance in parameter estimation in single-channel speech enhancement," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 7462–7466.

[13] P. Mowlaee, M. Watanabe, and R. Saeidi, "Show & tell: Phase-aware single-channel speech enhancement," in *14th Annual Conference of the International Speech Communication Association*, 2013.

[14] P. Mowlaee and M. Watanabe, "Partial phase reconstruction using sinusoidal model in single-channel speech separation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 832–836.

[15] K. K. Paliwal, K. K. Wojcicki, and B. J. Shannon, "The importance of phase in speech enhancement," *Speech Communication*, vol. 53, no. 4, pp. 465–494, 2011.

[16] T. Gerkmann, "MMSE-optimal enhancement of complex speech coeffcients with uncertain prior knowledge of the clean speech phase," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014, pp. 4511–4515.

[17] T. Gerkmann and M. Krawczyk, "MMSE-optimal spectral amplitude estimation given the STFT-phase," *Signal Processing Letters, IEEE*, vol. 20, no. 2, pp. 129 –132, Feb. 2013.

[18] P. Mowlaee and R. Saeidi, "Time-frequency constraint for phase estimation in single-channel speech enhancement," *The International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014.

[19] L. D. Alsteris and K. K. Paliwal, "Further intelligibility results from human listening tests using the short-time phase spectrum," *Speech Communication*, vol. 48, no. 6, pp. 727 – 736, 2006.

[20] T. Kleinschmidt, S. Sridharan, and M. Mason, "The use of phase in complex spectrum subtraction for robust speech recognition," *Computer Speech and Language*, vol. 25, no. 3, pp. 585 – 600, 2011.

[21] P. Mowlaee, R. Saeidi, and Y. Stylanou, "INTERSPEECH 2014 special session phase importance in speech processing applications," *in Proceedings of the 15th International Conference on Spoken Language Processing*, 2014.

[22] I. Cohen, "Relaxed statistical model for speech enhancement and a priori SNR estimation," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 870–881, 2005.

[23] P. Mowlaee, R. Saeidi, and R. Martin, "Phase estimation for signal reconstruction in single-channel speech separation," in *Proceedings of the International Conference on Spoken Language Processing*, 2012.

[24] H. Pobloth and W. B. Kleijn, "Squared error as a measure of perceived phase distortion," *J. Acoust. Soc. Am.*, vol. 114, no. 2, pp. 1081–1094, 2003.

[25] Y. Agiomyrgiannakis and Y. Stylianou, "Wrapped Gaussian Mixture Models for Modeling and High-Rate Quantization of Phase Data of Speech," *IEEE Trans. on Audio, Speech, and Lang. Proc.*, vol. 17, no. 4, pp. 775–786, 2009.

[26] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, "DARPA TIMIT acoustic phonetic continuous speech corpus cdrom," 1993.

[27] A. Varga, H. J. M. Steeneken, M. Tomlinson, and D. Jones, "The NOISEX–92 Study on the Effect of Additive Noise on Automatic Speech Recognition," *Technical Report, DRA Speech Research Unit*, 1992.

[28] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 749–752, Aug 2001.

[29] J. Ma and P. C. Loizou, "Snr loss: A new objective measure for predicting the intelligibility of noise-suppressed speech," *Speech Commun.*, vol. 53, no. 3, pp. 340–354, Mar. 2011.

[30] Y. Hu and P. Loizou, "Evaluation of objective quality measures for speech enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 1, pp. 229–238, 2008.

[31] L. D. Alsteris and K. K. Paliwal, "Iterative reconstruction of speech from short-time fourier transform phase and magnitude spectra," *Computer Speech & Language*, vol. 21, no. 1, pp. 174–186, 2007.