



Novel Shifted Real Spectrum for Exact Signal Reconstruction

Meet H. Soni, Rishabh Tak, and Hemant A. Patil

Speech Research Lab,

Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, Gujrat

{meet_soni, rishabh_tak, hemant_patil}@daiict.ac.in

Abstract

Retrieval of the phase of a signal is one of the major problems in signal processing. For an exact signal reconstruction, both magnitude, and phase spectrum of the signal is required. In many speech-based applications, only the magnitude spectrum is processed and the phase is ignored, which leads to degradation in the performance. Here, we propose a novel technique that enables the reconstruction of the speech signal from magnitude spectrum only. We consider the even-odd part decomposition of a causal sequence and process only on the real part of the DTFT of the signal. We propose the shifting of the real part of DTFT of the sequence to make it non-negative. By adding a constant of sufficient value to the real part of the DTFT, the exact signal reconstruction is possible from the magnitude or power spectrum alone. Moreover, we have compared our proposed approach with recently proposed phase retrieval method from magnitude spectrum of the Causal Delta Dominant (CDD) signal. We found that the method of phase retrieval from CDD signal and proposed method are identical under certain approximation. However, proposed method involves the less computational cost for the exact processing of the signal.

1. Introduction

For a long time, researchers thought that the “human ear is phase deaf”, i.e., phase of the sound signal does not affect the human perception system. In 1843 Ohm discussed that humans perceive sound in the form of harmonics. He proposed that the frequencies that are present in the sound signal are the only factor that affect the perception quality of signal [1]. Helmholtz, through his studies on periodic signals, arrived at an observation that the cochlea of human ear works as a spectrum analyzer. He observed that magnitude of frequency components was the *only* factor that helps human ears to perceive the musical tone [1]. However, this theory does not apply for quasi-periodic signals. After that, for the first time Seebeck argued that perception of tones not only depends on the strength of harmonics but also on the phase of the sound signal, which then became famous as Ohm-Seebeck dispute [1]. During the past three decades, there are many successful attempts for reconstructing the signal from its magnitude spectrum, which is referred to as *Phase Retrieval* problem. This idea was basically originated from the area of optical imaging. Initially, Gerchberg and Saxton [2] and Fineup [3–5] had started working on this research problem. They proposed an iterative algorithm for phase retrieval. Moravec et. al. [6] proposed the idea of compressive phase retrieval. Ohlsson and Eldar et. al. [7] who found some sufficient conditions on the measurement ensembles to guarantee unique reconstruction for sparse signal. Netrapalli et. al. [8] gave theoretical guarantee regarding the convergence of an alternating minimization for phase retrieval problem. Eldar et. al. [9] addressed the problem of reconstruction from the magnitude of

Short-Time Fourier transform (STFT). Sparse spectral factorization problem was solved by Lu and Vetterli [10]. Candes et. al. [11] proposed the idea of *Phase Lift* for phase retrieval problem. These studies were primarily focused on applications related to image processing.

Oppenheim et. al. [12–15] brought this idea into the signal processing community. They demonstrated synthesis of signal using only partial information in the Fourier-domain and they proposed signal reconstruction techniques. They arrived at two most significant results that minimum-phase sequences can be characterized by their magnitude spectrum and the associated Hilbert transform relations between log-magnitude and phase spectrum [16]. In [13], Hayes et. al. proposed the method of phase retrieval for real sequences. They actually solved problem for those sequences whose \mathcal{Z} -transform does not contain reciprocal pole-zero pair. In [17], Quatieri et. al. proposed an iterative technique for signal reconstruction using only its magnitude spectrum. They studied that how to reconstruct the time-domain signal from its magnitude spectrum as well as studied how to use the phase information to achieve a unique reconstruction of a time-domain signal. Then, in [18], Yegnanarayana et. al. analyzed the importance of group delay for minimum-phase sequences for the recovering of signal from its magnitude spectrum.

S. Mukherjee et. al. [19], [20] found the application of phase retrieval in the area of frequency-domain optical-coherence tomography. Recently, B. A. Shenoy et. al. have proposed the idea of exact phase retrieval in principal shift-invariant spaces [21], [22]. They have identified a class of continuous-time signals that are neither causal nor minimum-phase and yet guarantees the retrieval of exact phase. Very recently, Chandra proposed the concept of phase-encoded speech spectrogram using the property of Causal Delta-Dominance (CDD) [23]. He proposed the algorithm to encode the phase in the magnitude spectrogram as well as an algorithm for signal reconstruction.

In this paper, we propose a novel technique for exact signal reconstruction from its power spectrum alone. We consider the real part of the Discrete-Time Fourier Transform (DTFT) for our analysis. Since any signal can be represented in terms of its even and odd parts, we use symmetry property of the Fourier transform of even and odd signal to encode phase information in magnitude spectrum. For any real sequence, the real part of DTFT belongs to the even part of the signal [16]. Hence, if we consider causal signals, the signal can be retrieved completely using the real part of DTFT. We use this property of DTFT and propose a novel method to encode phase information in the magnitude spectrum, and also the exact signal reconstruction technique from the magnitude spectrum. We also establish a connection of the proposed method with the method proposed in [23] using certain approximation. Our finding suggests that under some approximations, the method described in [23] and

our proposed method are identical in terms of phase encoding in the magnitude spectrum. Moreover, the proposed method provides better signal reconstruction along with less computational complexity.

2. Proposed Shifted Real Spectrum

Any real causal sequence $x[n]$ can be written in terms of its even and odd part.

$$x[n] = x_{\text{even}}[n] + x_{\text{odd}}[n], \quad (1)$$

where x_{even} and x_{odd} are even and odd parts of $x[n]$, respectively. By taking DTFT of Eq. (1), we get,

$$X(e^{j\omega}) = X_{\text{even}}(e^{j\omega}) + X_{\text{odd}}(e^{j\omega}), \quad (2)$$

where $X(e^{j\omega})$, $X_{\text{even}}(e^{j\omega})$ and $X_{\text{odd}}(e^{j\omega})$ are DTFT of $x[n]$, $x_{\text{even}}[n]$ and $x_{\text{odd}}[n]$, respectively.

By using properties of DTFT, we can say that $X_{\text{even}}(e^{j\omega})$ will be purely real and $X_{\text{odd}}(e^{j\omega})$ will be purely imaginary. Hence, the real part of $X(e^{j\omega})$ will be $X_{\text{even}}(e^{j\omega})$ and the imaginary part of $X(e^{j\omega})$ will be $X_{\text{odd}}(e^{j\omega})$. For any discrete-time sequence $x[n]$,

$$x_{\text{even}}[n] = \frac{x[n] + x[-n]}{2}, \quad (3)$$

and

$$x_{\text{odd}}[n] = \frac{x[n] - x[-n]}{2}. \quad (4)$$

If $x[n]$ is real, causal, and if $x(0) = 0$, then it can be recovered from either from $x_{\text{even}}[n]$ or $x_{\text{odd}}[n]$, by simply taking its causal part. Hence, the exact signal can be reconstructed using either real part or imaginary part of the DTFT. We use real part of the DTFT for further processing. However, imaginary part of the DTFT can also be used in the similar manner.

Fig. 1 shows the block diagram of proposed method. We propose shifting of the real part ($X_R(e^{j\omega})$) in frequency-domain by adding a constant α to it. Here, the purpose of adding the α is to make the real part of the DTFT non-negative. If the DTFT spectrum of any signal is real, then phase of the spectrum can be written as,

$$\angle X_R(e^{j\omega}) = \begin{cases} 0 & \text{if } X_R(e^{j\omega}) \geq 0, \\ \pi & \text{if } X_R(e^{j\omega}) < 0. \end{cases} \quad (5)$$

By adding the proper value of α , we make the spectrum under consideration purely real and nonnegative, which will make phase of the real spectrum zero. The value of alpha must be chosen such that $\alpha + X_R(e^{j\omega})$ is nonnegative. If we choose $\alpha \geq$

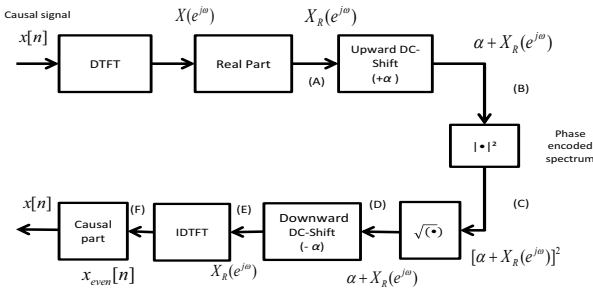


Figure 1: Block diagram of proposed approach.

$\sum |x[n]| \geq |X(e^{j\omega})|$ [23], it will ensure that $\alpha + X_R(e^{j\omega})$ is nonnegative. Hence, after adding the α , we can reconstruct the signal only from magnitude of $\alpha + X_R(e^{j\omega})$, which implies that phase of the signal is encoded in $|\alpha + X_R(e^{j\omega})|^2$. Hence, the equation for shifted real spectrum ($X_{\text{SRS}}(e^{j\omega})$) is,

$$X_{\text{SRS}}(e^{j\omega}) = \alpha + X_R(e^{j\omega}). \quad (6)$$

The speech signal can be reconstructed from $|\alpha + X_R(e^{j\omega})|^2$ by simply inverting all the operations. Fig. 1 shows the steps for encoding and decoding of the phase in magnitude spectrum. Square magnitude can be inverted by square root operation. Subtracting alpha from magnitude spectrum will give $X_R(e^{j\omega})$. Taking its inverse DTFT will give $x_{\text{even}}[n]$. By taking the causal part of $x_{\text{even}}[n]$, we can get the exact reconstructed signal $x[n]$.

3. Relation with Previous Approach

3.1. Phase encoded spectrogram

In order to generate the intelligible speech, we must require the phase information. Since speech signal is in general a mixed phase signal [17], if we want to reconstruct the speech signal using only from magnitude spectrum, we can only recover the minimum-phase component of speech signal. Thus, we cannot reconstruct the speech signal exactly from its magnitude spectrum alone. However, by doing the slight modification, in particular, adding a Kronecker impulse of right amplitude at the origin of the signal makes it a CDD sequence [22]. These sequences allow signal reconstruction from their magnitude spectrum and it is not necessary that these sequences (before adding impulse at origin) are minimum-phase. An important result is that finite-length CDD sequences are minimum-phase [22]. This modification allows the phase of the signal to be encoded in the magnitude spectrogram. This is called the phase-encoded spectrogram, and if the signal is speech then it is phase-encoded speech spectrogram. Fig. 2 shows the block diagram of the technique by which phase is encoded in magnitude spectrogram, as well as the signal reconstruction steps.

Consider a causal finite-length sequence, $\{x[n]\}_{n \in [1, N]}$. Adding a Kronecker impulse of amplitude α , we get,

$$\tilde{x}[n] = x[n] + \alpha\delta[n], \quad (7)$$

By taking DTFT of Eq. (7):

$$\tilde{X}(e^{j\omega}) = \sum_{n=0}^N \tilde{x}[n]e^{-j\omega n} = \alpha + X(e^{j\omega}), \quad (8)$$

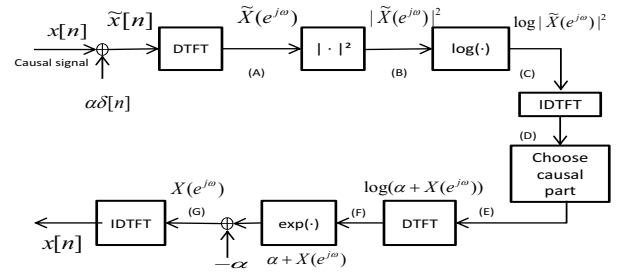


Figure 2: Block Diagram of Phase encoding and reconstruction [23].

where α produces the DC-shift in $X(e^{j\omega})$ across all the frequencies. It is claimed in [23] that, if $\alpha > |X(e^{j\omega})|$, $\forall \omega \in [-\pi, \pi]$, then $\tilde{x}[n]$, and therefore, $x[n]$ can be exactly recovered from $|\tilde{X}(e^{j\omega})|$. The proof of this claim can also be found in [22, 23]

The value of α is conditioned as,

$$\alpha = k \sum_{n=1}^N |x[n]|, \quad (9)$$

where $k > 1$. We can take any arbitrary value of $k > 1$. Moreover, it is also shown in [22, 23], that as the value of α increases, the reconstruction of signal becomes better. Hence, the value of α is critical for exact signal reconstruction using this method.

3.2. Relationship with existing phase encoded spectrograms

To establish a relationship between proposed shifted real spectrum and phase encoded spectrum, we take DTFT on both side of Eq. (7),

$$\tilde{X}(e^{j\omega}) = X(e^{j\omega}) + \alpha. \quad (10)$$

Now, taking magnitude square on both sides, we get,

$$\begin{aligned} |\tilde{X}(e^{j\omega})|^2 &= |(X_R(e^{j\omega}) + \alpha) + jX_{Im}(e^{j\omega})|^2, \\ &= (X_R(e^{j\omega}) + \alpha)^2 + (X_{Im}(e^{j\omega}))^2, \\ &= X_R^2(e^{j\omega}) + 2\alpha X_R(e^{j\omega}) + \alpha^2 + X_{Im}^2(e^{j\omega}), \\ &= \alpha^2 + 2\alpha X_R(e^{j\omega}) + X_R^2(e^{j\omega}) + X_{Im}^2(e^{j\omega}). \end{aligned} \quad (11)$$

If the value of α is very high, we can take the approximation $\alpha^2 + 2\alpha X_R(e^{j\omega}) \gg X_R^2(e^{j\omega}) + X_{Im}^2(e^{j\omega})$. For very high value of α , the contribution of the term $X_R^2(e^{j\omega}) + X_{Im}^2(e^{j\omega})$ in Eq. (11) can be ignored. However, we are keeping the term $X_R^2(e^{j\omega})$ in order to keep the equation in terms of real part of the signal. Hence, after neglecting the $X_{Im}^2(e^{j\omega})$ part in Eq. (11), we get,

$$|\tilde{X}(e^{j\omega})|^2 \approx \alpha^2 + 2\alpha X_R(e^{j\omega}) + X_R^2(e^{j\omega}), \quad (12)$$

$$|\tilde{X}(e^{j\omega})|^2 \approx (\alpha + X_R(e^{j\omega}))^2. \quad (13)$$

As it can be observed from Eq. (9), the value of alpha is similarly conditioned as in Section 2. Hence, $\alpha + X_R(e^{j\omega})$ will be nonnegative. Therefore, the magnitude spectrum of CDD signal as per our approximation will be,

$$|\tilde{X}(e^{j\omega})| \approx \alpha + X_R(e^{j\omega}). \quad (14)$$

As it can be observed, this result is identical to proposed shifted real spectrum shown in Eq. (6). Hence, findings of [23] and the proposed approach are similar for large value of α . Moreover, by comparing Fig. 1 and Fig. 2, it can be observed that the reconstruction is simpler using proposed approach than existing phase encoded spectrum [23]. In particular, proposed approach does not require the cepstral-domain processing to reconstruct the signal. This leads to faster reconstruction of the exact signal. Other key advantage of proposed method is that accuracy of reconstruction is independent of α . The reconstruction accuracy in case of phase encoded spectrum is proportional to the value of α [22, 23]. In proposed method, any value of α that is greater than $\sum_{n=1}^N |x[n]|$ will give exactly similar reconstruction accuracy.

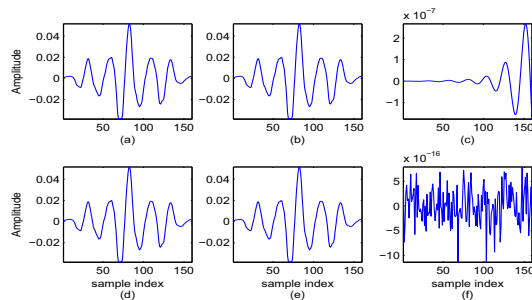


Figure 3: (a) Voiced frame from an utterance from NOIZEUS database, (b) frame reconstructed using phase encoded spectrum method [23], and (c) reconstruction error. Similarly, (d) same frame, (e) reconstructed frame using proposed shifted real spectrum approach, and (f) reconstruction error.

4. Experiments and Results

We present our studies on NOIZEUS database [24]. We have chosen a clean utterance from the database to demonstrate phase encoding and reconstruction using both approaches. For analysis, we have divided the speech signal into frames using 20 ms hanning window with 5 ms overlap. Speech signals are sampled at 8 kHz. To ensure enough oversampling in frequency domain, we use 512-point Fast Fourier Transform (FFT) instead of DTFT. First, we present frame-level reconstruction to analyze the frame wise reconstruction error using both the techniques. To ensure the condition $x(0) = 0$ is satisfied, we pad a zero to each frame before analysis. Fig. 3 shows a voiced frame of an utterance from the database. The first row corresponds to the phase encoded spectrum shown in Fig. 2. The second row represents the results of proposed approach shown in Fig. 1. It can be observed that reconstruction error using both approaches is very less (less than 6-7 orders of magnitude). However, the reconstruction error using proposed method is far less (9 orders of magnitude less) than using the method shown in Fig. 2. Hence, proposed method gives better reconstruction using simpler encoding-decoding procedure.

In the next part, we show the utterance-level reconstruction properties of both approaches. Fig. 4 shows the time-domain waveforms, and spectrograms of the signal under consideration and reconstructed signals using both methods. It was observed that both methods are able to reconstruct the speech signal with same reconstruction Signal to Noise Ratio (SNR). The reconstruction SNR was 39.67 dB for both the methods. Moreover, we also found the reconstruction SNR for ideal condition, i.e. for framing and overlap-add system. It was also 39.67 dB. Similarly, the Log Spectral Distortion (LSD) was also calculated using all the methods. The value of LSD was 0.0015 dB for both the methods as well as for ideal reconstruction using overlap-add system. Hence, it can be concluded that both the methods give ideal reconstruction SNR. Similar trend was observed in all the utterances of the database, including noisy utterances. It was observed during our experiments that speech signals reconstructed using both approaches had similar perceptual qualities.

To compare the time complexity of both the algorithms, we have noted the time required for analysis/synthesis of the utterance shown in Fig. 4. The time required for analysis/synthesis was 0.0585 sec and 0.1340 sec for our proposed method and method proposed in [23], respectively. To demonstrate the effect of different values of α on reconstruction accuracy, we have

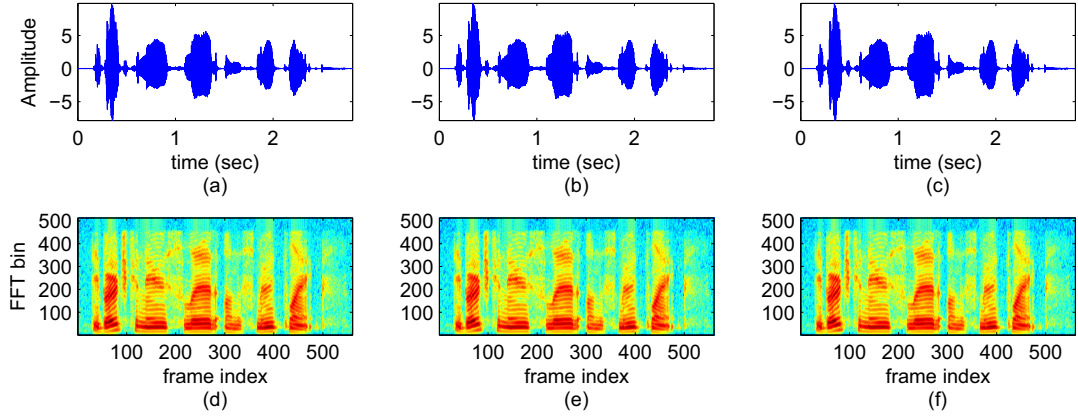


Figure 4: (a) Time-domain waveform of an utterance from NOIZEUS database, (b) signal reconstructed using existing phase encoded spectrum method [23], and (c) signal reconstructed using our proposed method. Similarly, (d), (e), and (f) are spectrograms of (a), (b), and (c), respectively. Reconstruction SNR for (e) and (f) is 39.67 dB, and Log Spectral Distortion (LSD) for (e) and (f) is 0.0015 dB.

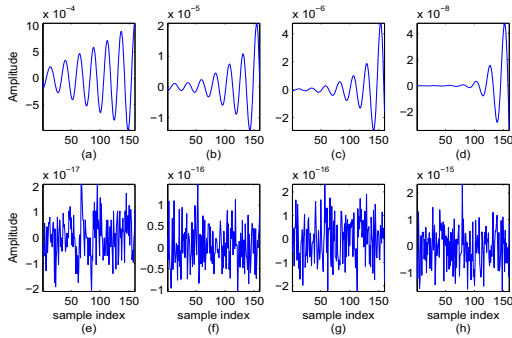


Figure 5: (a), (b), (c) and (d) are reconstruction error for the frame shown in Fig. 3 using phase encoded spectrum for $k=1, 5, 10$ and 100 , respectively. Similarly, (e), (f), (g) and (h) are reconstruction error using proposed method.

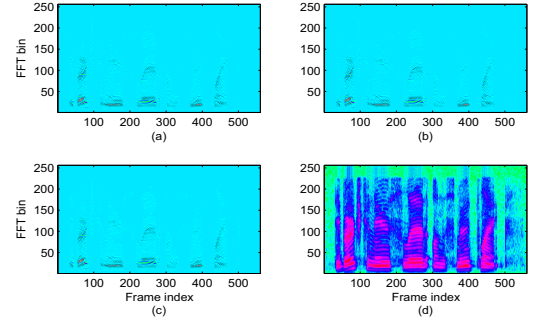


Figure 6: (a) Log Power spectrum extracted from block (C) of Fig. 2, (b) Log Power spectrum extracted from block (C) of Fig. 1, (c) real part of the FFT spectrum of the signal under consideration and (d) powers spectrum of the signal.

taken the similar frame shown in Fig. 3. We calculated value of α according to Eq. (9) for different values of k . As it can be seen from Fig. 3, as value of α increases, the reconstruction error decreases in case of method described in Fig. 2. However, the reconstruction error is almost similar in case of proposed method for different values of α . The differences visible in Fig. 5 using our proposed method are suspected due to finite precision of the computer while performing computations. Hence, we can say that reconstruction accuracy using our proposed method is independent of the value of α .

To observe the properties of the phase encoded magnitude spectra using both methods, we have shown the log-magnitude plots extracted using both methods. In Fig. 6, we have shown the output of block (C) of Fig. 1 and output of block (C) of Fig. 2. These blocks corresponds to the log-energy after adding α in the process. For reference, we have shown the real part of STFT of the signal. As it can be observed, the log-energy spectra after adding α have similar properties as the real part of STFT. It implies that ultimately, both the features manipulates the real part of the spectrum to ensure encoding of the phase, and consequently, exact reconstruction of the signal.

5. Summary and Conclusions

In this paper, we have shown a simple yet novel way to encode the phase of the signal into its magnitude spectrum for exact signal reconstruction. We use odd-even decomposition of a real causal signal and some basic properties of the Fourier transform to manipulate the signal encoding and decoding. We use shifted real spectrum of the signal to encode phase in magnitude spectrum itself. Moreover, we established the relationship of our proposed approach with the recently proposed method of phase encoded speech spectrograms. We have showed that using some approximation, both the approaches are identical in concept. However, proposed approach is less computational expensive than the earlier one. We found that both approaches ultimately manipulates the real part of the STFT spectrum. Hence, if we can come up with an efficient method to encode and decode the real part of STFT spectrum, the phase can be encoded and signal can be reconstructed exactly using that method. This kind of encoding/decoding may be useful in many speech technology applications such as speech enhancement, speech synthesis, voice conversion, etc.

6. References

- [1] R. S. Turner, "The Ohm-Seebeck dispute, Hermann von Helmholtz, and the origins of physiological acoustics," *The British Journal for the History of Science*, vol. 10, no. 01, pp. 1–24, 1977.
- [2] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, p. 237, 1972.
- [3] J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Optics Letters*, vol. 3, no. 1, pp. 27–29, 1978.
- [4] Fienup, James R, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, 1982.
- [5] J. R. Fienup, "Phase retrieval algorithms: a personal tour [invited]," *Applied Optics*, vol. 52, no. 1, pp. 45–56, 2013.
- [6] M. L. Moravec, J. K. Romberg, and R. Baraniuk, "Compressive phase retrieval," in *Proceedings of SPIE International Symposium on Optical Science and Technology*, vol. 6701, San Jose, California, United States, 2007, pp. 670 120–670 120–11.
- [7] H. Ohlsson and Y. C. Eldar, "On conditions for uniqueness in sparse phase retrieval," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, 2014, pp. 1841–1845.
- [8] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems*, South Lake Tahoe, California, 2013, pp. 2796–2804.
- [9] Y. C. Eldar, P. Sidorenko, D. G. Mixon, S. Barel, and O. Cohen, "Sparse phase retrieval from short-time Fourier measurements," *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 638–642, 2015.
- [10] Y. M. Lu and M. Vetterli, "Sparse spectral factorization: Unicity and reconstruction algorithms," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 2011, pp. 5976–5979.
- [11] E. J. Cands, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 199–225, 2013.
- [12] A. V. Oppenheim, J. S. Lim, and S. R. Curtis, "Signal synthesis and reconstruction from partial Fourier-domain information," *Journal of the Optical Society of America (JOSA)*, vol. 73, no. 11, pp. 1413–1420, 1983.
- [13] M. Hayes, J. Lim, and A. Oppenheim, "Signal reconstruction from phase or magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 6, pp. 672–680, 1980.
- [14] P. Van Hove, M. Hayes, J. Lim, and A. Oppenheim, "Signal reconstruction from signed Fourier transform magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 5, pp. 1286–1293, 1983.
- [15] S. Nawab, T. Quatieri, and J. Lim, "Signal reconstruction from short-time Fourier transform magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 986–998, 1983.
- [16] A. V. Oppenheim, *Discrete-Time Signal Processing*, 2nd ed. Pearson Education India, 1999.
- [17] T. Quatieri and A. Oppenheim, "Iterative techniques for minimum phase signal reconstruction from phase or magnitude," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 6, pp. 1187–1193, 1981.
- [18] B. Yegnanarayana, D. Saikia, and T. Krishnan, "Significance of group delay functions in signal reconstruction from spectral magnitude or phase," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 3, pp. 610–623, 1984.
- [19] S. Mukherjee and C. S. Seelamantula, "A non-iterative phase retrieval algorithm for minimum-phase signals using the annihilating filter," *Sampling Theory in Signal & Image Processing*, vol. 11, no. 2-3, pp. 165–193, 2012.
- [20] Mukherjee, Subhadip and Seelamantula, Chandra Sekhar, "An iterative algorithm for phase retrieval with sparsity constraints: application to frequency domain optical coherence tomography," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 2012, pp. 553–556.
- [21] B. A. Shenoy and C. S. Seelamantula, "Exact phase retrieval for a class of 2-D parametric signals," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 90–103, 2015.
- [22] B. A. Shenoy, S. Mulleti, and C. S. Seelamantula, "Exact phase retrieval in principal shift-invariant spaces," *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 406–416, 2016.
- [23] C. S. Seelamantula, "Phase-encoded speech spectrograms," in *INTERSPEECH*, San Francisco, USA, 2016, pp. 1775–1779.
- [24] Y. Hu and P. C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms," *Speech communication*, vol. 49, no. 7, pp. 588–601, 2007.