AN ALTERNATING OPTIMIZATION APPROACH FOR PHASE RETRIEVAL

Huaiping Ming[1], *Dongyan Huang*[2], *Lei Xie*[1], *Haizhou Li*[2] *and Minghui Dong*[2]

[1]School of Computer Science, Northwestern Polytechnical University, Xi'an, China
[2]Institute for Infocomm Research, A*STAR, Singapore

minghuaiping@gmail.com, huang@i2r.a-star.edu.sg, xielei21st@gmail.com,
hli@i2r.a-star.edu.sg, mhdong@i2r.a-star.edu.sg

ABSTRACT

In this paper, we address the problem of phase retrieval to recover a signal from the magnitude of its Fourier transform. In many applications of phase retrieval, the signals encountered are naturally sparse. In this work, we consider the case where the signal is sparse under the assumption that few components are nonzero. We exploit further the sparse nature of the signals and propose a two stage sparse phase retrieval algorithm. A simple iterative minimization algorithm recovers a sparse signal from measurements of its Fourier transform (or other linear transform) magnitude based on the minimization of a block l_1 norm. We show in the experiments that the proposed algorithm achieves a competitive performance. It is robust to noise and scalable in practical implementation. The proposed method converges to a more accurate and stable solution than other existing techniques for synthetic signals. For speech signals, experiments show that the voice quality of reconstructed speech signals is almost as good as the original signals.

Index Terms: Phase Retrieval, Damped Gauss-Newton Method, Sparse Coding

1. INTRODUCTION

Recovery of a signal from the magnitude of its Fourier transform, or equivalently, from its auto correlation is known as phase retrieval [1]. This problem arises in many applications such as X-ray crystallography [1, 2], astronomical imaging [3], microscopy [4] and speech processing applications [5, 6]. There was a recent interest in studying the phase importance of signals in speech processing [6]. Many studies elaborated the potential of using phase information in audio and speech signal processing applications such as speech enhancement [7], speech recognition [8] and speech analysis/synthesis [9].

Conventionally, most speech features only contain Fourier transform amplitude information (Mel-frequendy Cepstral coefficients is an example) without the phase information. The reasons why phase information was neglected vary in different applications. A detailed discussion can be found in [6]. The recent studies have revealed the importance of phase information and how it impacts on system performance. For example, a minimum phase technique improves the performance of text-to-speech synthesis systems [10], and an estimated phase from the spectral amplitude improves the performance of source separation [11] and speech enhancement [12]. It is generally believed that an effective phase retrieval algorithm would be helpful for many speech signal processing applications.

Given a signal, there could be many other different signals whose Fourier transforms share the same magnitude, so the problem is generally ill-posed [13]. For any given Fourier transform magnitude, every possible retrieved phase may lead to a different signal. A common approach to overcome this ill-posed problem is to exploit the prior information of the signal such as the support of signal (region in which the signal is nonzero), non-negativity and the magnitude of signal is used by most methods.

Two main categories of approaches have been proposed to solve this problem, namely, semi-definite programming-based approaches (SDP-based) [14, 15, 16, 17] and iterative projection approaches (Fienup-type) [18, 19, 20, 21]. Despite tremendous progress, phase retrieval remains a challenging problem. The SDP-based approaches are not suitable for large scale problems, and the iterative projection approaches suffer from convergence issues especially for 1D signals [22, 23]. Recently, many researchers exploit the sparsity of a signal to recover signal, and many sparse signal processing approaches are proposed [23, 24, 25]. Among these methods, an algorithm named GESPAR [23] outperforms conventional SDP-based and Fienup-type method in terms of complexity, success probability and robustness to noise. The method derived in this paper is motivated in many aspects by the work in [23, 24, 26].

In this paper, we propose an effective phase retrieval method, which leads to accurate recovery of sparse signals with very high probability. The proposed approach is based upon the damped Gauss-Newton method (DGN) [27, 28] and sparse coding method [29]. The DGN method is used to minimize the objective function and obtain a suboptimal estimation of the original signal. The local linearization of the constraints induces a group-sparse structure on the variables. The sparse coding method is used to get new measurements of the estimated signal under different complete basis to enforce this structure. We alternate the above two steps. Such an iterative minimization achieves a good performance by finding the optimal solution under different complete basis, which is called an alternating optimization method. Due to the noise reduction ability of sparse coding method, we expect that the proposed method is robust to noise.

Numerical simulations for sparse synthetic signals show that our method is more accurate than current techniques and robust to noise. Experiments for speech signal show that the voice quality of the recovered signal is almost as good as the original speech signal.

2. PROBLEM FORMULATION

Let $\overline{\mathbf{x}} = (x_1, x_2, ..., x_n)$ be a real-valued signal of length n. In order to further exploit the sparsity of the signal, a $(N-n)$ zero padding version of $\overline{\mathbf{x}}$ is given by $\mathbf{x} = (x_1, x_2, ..., x_N)$. The square of an N point discrete Fourier transform magnitude of vector $\mathbf{x}$ given by Eq. (1) is a known measurement vector. Zero padding does not add any extra information to the original signal and the Fourier transform spectrum, but it helps to enforce the sparsity of the signal. We denote the measurement vector as $\mathbf{y} = (y_1, y_2, ..., y_N)$.

$$y_i = \left| \sum_{m=1}^{N} x_m e^{-\dfrac{2\pi j(m-1)(i-1)}{N}} \right|^2, i = 1, ..., N. \quad (1)$$

We can express $\mathbf{y}$ as $\mathbf{y} = |\mathbf{F}\mathbf{x}|^2$, where $\mathbf{F} \in \mathbb{C}^{N\times N}$ is the DFT matrix with elements $\exp\{-\frac{2\pi j(m-1)(i-1)}{N}\}$, and $|\bullet|^2$ is the element-wise absolute square operation. The vector $\overline{x}$ is known to be l-sparse, that is, it has l non-zero elements constraint by the support set. In case of the signal is none sparse, the support is $\mathbf{s} = (1, 2, ..., n)$. Our goal is to recover $\overline{\mathbf{x}}$ or $\mathbf{x}$ given the measurement vector $\mathbf{y}$ under the support $\mathbf{s}$.

Define $\mathbf{A}_i = \Re(\mathbf{F})_i^H \Re(\mathbf{F}_i) + \Im(\mathbf{F}_i)^H \Im(\mathbf{F}_i) \in \mathbb{R}^{N\times N}$, where $\mathbf{F}_i$ is the ith row of the DFT matrix $\mathbf{F}$. We need to minimize the sum of squared errors subject to the support constraint:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) = \sum_{i=1}^{N} (\mathbf{x}^T \mathbf{A}_i \mathbf{x} - y_i)^2 \\ s.t. \quad & \|x\|_1 \le \alpha l, \\ & \text{supp}(\mathbf{x}) = \mathbf{s} = (s_1, s_2, ..., s_l), \\ & \mathbf{x} \in \mathbb{R}^N. \end{aligned} \quad (2)$$

where $\|\bullet\|_1$ is the l_1 norm, and α is a constant. Eq. (2) is the mathematical formulation of phase retrieval problem that we consider. Due to the loss of Fourier phase information, the solution of this problem is trivial degenerated [23, 13]. Hence there is no guarantee for a unique recovery of $\mathbf{x}$. This ambiguity cannot be resolved using any method that use sparsity constraint and Fourier magnitude measurements alone.

2.1. DGN and sparse coding method

The DGN method and sparse coding method are two iterative steps in the proposed algorithm. We briefly describe next.

2.1.1. DGN method

The DGN method can be used to solve the problem of minimizing the objective function f under a given support $\mathbf{s} = (s_1, s_2, ..., s_l)$:

$$\min\{f(\mathbf{U_s x}) : \mathbf{x} \in \mathbb{R}^l\}, \quad (3)$$

where $\mathbf{U_s} \in \mathbb{R}^{N\times l}$ is the matrix consisting of the columns of the identity matrix $\mathbf{I}_N$ corresponding to the support set $\mathbf{s}$. Combining Eq. (2) and Eq. (3), we have the minimization formulation as

$$\min\{f(\mathbf{x}) = \sum_{i=1}^{N} (\mathbf{x}^T \mathbf{U_s}^T \mathbf{A}_i \mathbf{U_s} \mathbf{x} - y_i)^2 : \mathbf{x} \in \mathbb{R}^l\}. \quad (4)$$

Let $\mathbf{B}_i = \mathbf{U_s}^T \mathbf{A}_i \mathbf{U_s}$, and $h_i(\mathbf{x}) = \mathbf{x}^T \mathbf{B}_i \mathbf{x} - y_i$. Then the function $f(\mathbf{x})$ from Eq. (4) can be written as

$$f(\mathbf{x}) = \sum_{i=1}^{N} h_i^2(\mathbf{x}). \quad (5)$$

This is a nonlinear least square problem. The DGN method begins with an arbitrary vector $\mathbf{x}_0$. At each step, h_i is replaced by a linear approximation around $\mathbf{x}_{k-1}$:

$$\begin{aligned} h_i &\approx h_i(\mathbf{x}_{k-1}) + \nabla h_i(\mathbf{x}_{k-1})^T (\mathbf{x} - \mathbf{x}_{k-1}) \\ &= \mathbf{x}_{k-1}^T \mathbf{B}_i \mathbf{x}_{k-1} - y_i + 2(\mathbf{B}_i \mathbf{x}_{k-1})^T (\mathbf{x} - \mathbf{x}_{k-1}). \end{aligned} \quad (6)$$

We choose x_k to be the solution of the problem

$$\min_{x} \sum_{i=1}^{N} \left(\mathbf{x}_{k-1}^T \mathbf{B}_i \mathbf{x}_{k-1} - y_i + 2(\mathbf{B}_i \mathbf{x}_{k-1})^T (\mathbf{x} - \mathbf{x}_{k-1}) \right)^2. \quad (7)$$

Then this problem can be written as a linear least square problem

$$\widetilde{\mathbf{x}}_k = \arg\min \|\mathbf{J}(\mathbf{x}_{k-1})\mathbf{x} - \mathbf{e}_k\|_2^2, \quad (8)$$

with the ith row of $\mathbf{J}(\mathbf{x}_{k-1})$ being $\nabla h_i(\mathbf{x}_{k-1})^T = 2(\mathbf{B}_i \mathbf{x}_{k-1})^T$, and the ith component of $\mathbf{e}_k$ given by $y_i + \mathbf{x}_{k-1}^T \mathbf{B}_i \mathbf{x}_{k-1}$ for $i = 1, 2, ..., N$. The solution $\widetilde{\mathbf{x}}_\mathbf{k}$ is

$$\widetilde{\mathbf{x}}_\mathbf{k} = (\mathbf{J}(\mathbf{x}_{k-1})^T \mathbf{J}(\mathbf{x}_{k-1}))^{-1} \mathbf{J}(\mathbf{x}_{k-1})^T \mathbf{e}_k. \quad (9)$$

We then define the direction vector as $\mathbf{d}_k = \widetilde{\mathbf{x}}_\mathbf{k} - \mathbf{x}_{k-1}$, which is used to update the solution to a stationary point of $f(\mathbf{x})$. At last, $\mathbf{x}$ is updated by $\mathbf{x}_k = \mathbf{x}_{k-1} + t_k \mathbf{d}_k$, where t_k is the step-size. A more detailed description of the method is given in [23].

2.1.2. Sparse coding method

The sparse coding method is used to find a representation of input vectors approximately as a weighted linear combination of a small number of basis vectors. Let $\mathbf{X} \in \mathbb{R}^{N\times M}$ be the input matrix where each column is an input vector $\mathbf{x} \in \mathbb{R}^N$, let $\mathbf{D} \in \mathbb{R}^{N\times K}$ be the basis matrix where each column is a base vector $\mathbf{d} \in \mathbb{R}^N$, and let $\mathbf{Z} \in \mathbb{R}^{K\times M}$ be the coefficient matrix where each column is a coefficient vector $\mathbf{z} \in \mathbb{R}^K$. Then the input matrix $\mathbf{X}$ can be expressed as

$$\mathbf{X} = \mathbf{D}\mathbf{Z} + \mathbf{N}(\sigma^2), \quad (10)$$

where $\mathbf{N}(\sigma^2)$ is the reconstruction error matrix which is usually assumed as a zero-mean Gaussian distribution with covariance σ^2. In the sparse coding method, the distribution is defined as $\mathbf{N}(\sigma^2) \propto \exp(-\beta\|\mathbf{Z}\|_1)$, where β is a constant. Assuming a uniform prior on the basis, the maximum posteriori estimate of the base matrix and the coefficient matrix is the solution of following optimization problem:

$$\begin{aligned} \min_{\mathbf{D},\mathbf{Z}} \quad & \frac{1}{2\sigma^2} \|\mathbf{X} - \mathbf{D}\mathbf{Z}\|_F^2 + \beta\|\mathbf{Z}\|_1 \\ s.t. \quad & \sum_i \mathbf{D}_{i,j}^2 \le c, \forall j = 1, ..., N, \end{aligned} \quad (11)$$

where c is a constant number. This problem can be solved using convex optimization methods, and we adopt the algorithm described in [29].

2.2. The proposed alternating optimization

Suppose that we have M unknown signals given by $\mathbf{X} \in \mathbb{R}^{N\times M}$, with each column of the matrix is a vector $\mathbf{x}_i \in \mathbb{R}^N, i = 1, 2, ..., M$ corresponding to one signal. The square Fourier measurement of $\mathbf{X}$ is given by $\mathbf{Y} \in \mathbb{R}^{N\times M}$, and the corresponding support is given by $\mathbf{S} \in \mathbb{R}^{l\times M}$.

The basic idea of our algorithm are two folds: 1) to find the local optimal solution of Eq. (2) using DGN method; 2) to reconstruct the local optimal solution using sparse coding method, which enhance the group-sparse structure and the robustness to noise.

The proposed algorithm starts by applying the DGN method to input $\mathbf{Y}, \mathbf{S}$ and $\mathbf{x}_0$ to get an initialized suboptimal solution matrix $\widetilde{\mathbf{X}}_0$. The initial parameter $\mathbf{x}_0$ is a Gaussian random vector with zero mean and unit variance. Usually, the DGN method will stuck in some suboptimal solutions with random initial vector. In order to enhance the group-sparse structure and avoid the stagnation, we apply sparse coding method to the input matrix $\widetilde{\mathbf{X}}_0$, and the output are the basis vectors $\mathbf{D}_1$ and the coefficient vectors $\mathbf{Z}_1$. Then DGN method is used by taking $\mathbf{X}_1 = \mathbf{D}_1\mathbf{Z}_1$ as the initial matrix. It is important to note that the local optimal reconstructed signal by sparse coding is a near perfect reconstruction which increases the group-sparse structure of the signal, and enhances the robustness to noise. Comparing to random initialization, using the reconstructed signal by sparse coding method as the initial vector will increase the probability of convergence to global optimum. After the initialization step, the sparse coding method and DGN method alternate to optimize iteratively the signal as described in Algorithm 1. Finally, we can get a solution matrix $\mathbf{X}$ where each column $\mathbf{x}_i$ is the solution for Eq. (2).

If we ignore the reconstruction error $\mathbf{N}(\sigma^2)$ in Eq. (10), the minimization function of DGN method Eq. (4) can be rewritten as

$$\min\{f(\mathbf{z}) = \sum_{i=1}^{N}(\mathbf{z}^T\mathbf{U}_\mathbf{s}^T\mathbf{D}^T\mathbf{A}_i\mathbf{D}\mathbf{U}_\mathbf{s}\mathbf{z} - y_i)^2 : \mathbf{z} \in \mathbb{R}^m\}, \quad (12)$$

where $\mathbf{z}$ is the coefficient vector under the new basis $\mathbf{D}$ learned by sparse coding method, and m is the dimension of the coefficient vector. Let $\mathbf{B}_i = \mathbf{U}_\mathbf{s}^T\mathbf{D}^T\mathbf{A}_i\mathbf{D}\mathbf{U}_\mathbf{s}$, and $h_i(\mathbf{z}) = \mathbf{z}^T\mathbf{B}_i\mathbf{z} - y_i$, then the minimization function under the new basis can be written as

$$f(\mathbf{z}) = \sum_{i=1}^{N} h_i^2(\mathbf{z}), \quad (13)$$

which is the same as applying the DGN method to the new basis learned by sparse coding method. Therefore, the iterative optimization by DGN method and sparse coding method seeks to find a suboptimal solution near to global optimal solution for Eq. (2) under different complete basis, which is supported by the simulation results in Section 3.

Algorithm 1

Input:

$\mathbf{Y} \in \mathbb{R}^{N\times M}$ - measurement matrix, each column is a Fourier transform magnitude measurement vector $\mathbf{y}_i, i = 1, 2, ..., M$.
$\mathbf{S} \in \mathbb{R}^{l\times M}$ - support set matrix, each column is a support vector $\mathbf{s}_i, i = 1, 2, ..., M$.
$ITER$ - iteration numbers.
$\mathbf{x_0} \in \mathbb{R}^l$ - initial vector for DGN method.
β - penalty factor in sparse coding method.
$baseNum$ - number of basis vectors for sparse coding method.

General step:

1. **Initialization**
   - Apply DGN method with parameter $\mathbf{y}_i, \mathbf{s}_i$ and $\mathbf{x_0}$ for $i = 1, 2, ..., M$. The parameter $\mathbf{y}_i$ and $\mathbf{s}_i$ are the ith column of matrix $\mathbf{Y}$ and $\mathbf{S}$. The output $\widetilde{\mathbf{x}}_i$ is a suboptimal solution given by Eq. (9).
   - Repeat for $i = 1, 2, ..., M$, we get a solution matrix $\widetilde{\mathbf{X}}_0 \in \mathbb{R}^{N\times M}$ where each each column is vector $\widetilde{\mathbf{x}}_i$.
2. **Alternating optimization**
   Repeat
   - Apply sparse coding method with input matrix $\widetilde{\mathbf{X}}_{k-1}$ and parameter: $\beta, baseNum$. The output is the basis vectors $\mathbf{D}_k$ and the coefficient vectors $\mathbf{Z}_k$ described Eq. (11).
   - Apply DGN method with parameter $\mathbf{y}_i, \mathbf{s}_i$ and $\mathbf{x}_{i,k}$ for $i = 1, 2, ..., M$. Denoting $\mathbf{x}_i$ as the ith column of matrix $\mathbf{X}_k = \mathbf{D}_k\mathbf{Z}_k$, then the initial vector $\mathbf{x}_{i,k} = \mathbf{x}_i(\mathbf{s}_i)$. The output is a updated suboptimal solution matrix $\widetilde{\mathbf{X}}_k$.

   Until iteration number $k \geq ITER$.

Output:

$\mathbf{X} \in \mathbb{R}^{N\times M}$ - a matrix where each column $\mathbf{x}_i \in \mathbb{R}^N, i = 1, 2, ..., M$ is the solution for Eq. (2).

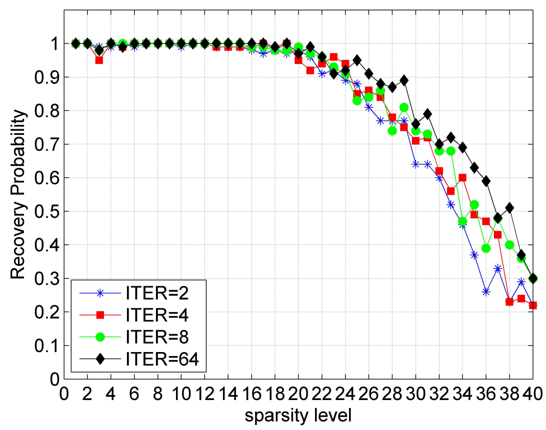

Fig. 1. Effect of iteration numbers ($ITER$) on recovery probability

3. NUMERICAL SIMULATIONS

To demonstrate the performance of our proposed method, we conduct several numerical simulations for both synthetic sparse signals and speech signals (none sparse).

3.1. Simulation for sparse signals

For sparse signals, we synthesise 100 different random vectors $\overline{\mathbf{x}}_i, i = 1, 2, ...100$ of length $n = 64$. Each vector is uniformly distributed in range $[-4, -3] \bigcup [3, 4]$. The support and the signal values are randomly selected for each simulation. The $N = 128$ point DFT of each signal is calculated, and their magnitude square is the measurement vectors $\mathbf{y}_i, i = 1, 2, ..., 100$. The GESPAR algorithm is tested for comparison purposes. The GESPAR method outperforms the conventional SDP-based and iterative Fienup-type algorithms in terms of recovery probability and the robustness to noise.

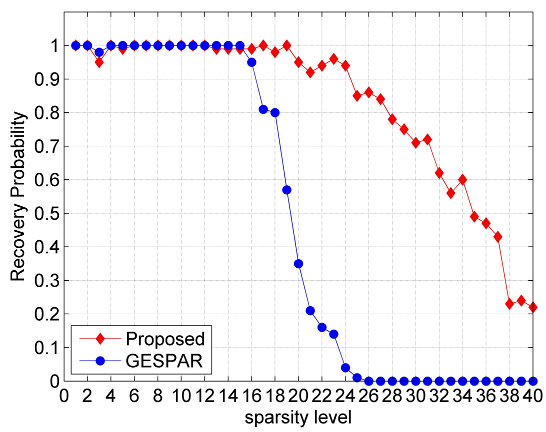

Fig. 2. Recovery probability vs. sparsity level

The number of iteration is an important factor (the input parameter $ITER$ in Algorithm 1) in the proposed algorithm. By increasing the number of iterations, we increase the probability to find the correct solution at the cost of increased computation time. The parameter $ITER$ is tested in the range $[2, 64]$. We report the recovery probability as a function of the number of iterations in Fig. 1, where the result is plotted for different sparsity levels. The success probability is defined as the ratio of correctly recovered signals $\mathbf{x}$ out of 100 simulations. The results show that by increasing the number of

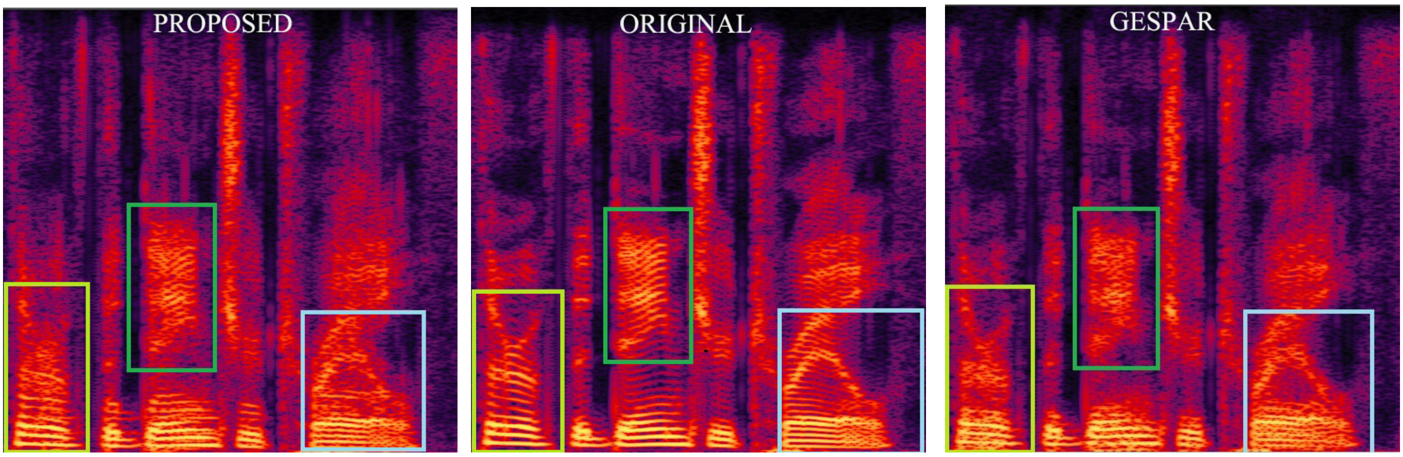

Fig. 4. Spectrograms of signal reconstructed by the proposed method (left), the original signal (middle) and signal reconstructed by the GESPA method (right). We emphasize three regions for comparison.

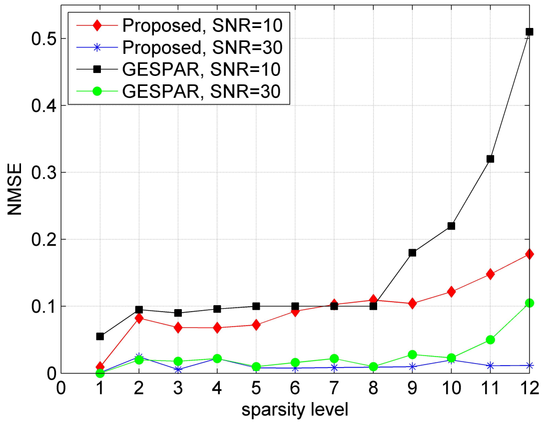

Fig. 3. Normalized MSE vs. sparsity level

Table 1. The NMSE and Quality Score of recovered speech signals.

	NMSE		Quality Score	
	Our method	GESPAR	Our method	GESPAR
1	0.38	0.42	4.2	3.8
2	0.40	0.42	4.1	3.8
3	0.41	0.45	4.2	3.9
4	0.40	0.43	4.1	3.6
5	0.40	0.43	3.9	3.6
6	0.42	0.45	4.0	3.8
7	0.42	0.46	3.8	3.5
8	0.36	0.39	4.2	4.0
9	0.37	0.39	4.1	3.7
10	0.39	0.41	4.0	3.6
Average	0.395	0.423	4.06	3.73

iterations, we improve the recovery probability. The signal recovery results of our method ($ITER = 4$) and the GESPAR method are shown in Fig. 2. It is clear that our proposed algorithm outperforms GESPAR. Let's define signal to noise ratio as $SNR = 20log\frac{\|\mathbf{y}\|}{\|\mathbf{v}\|}$, where $\mathbf{y}$ is the Fourier measurement and $\mathbf{v}$ is a white Gaussian noise, and the normalized mean squared reconstruction error (NMSE) as $NMSE = \frac{\|\mathbf{x}-\widehat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}$, where $\widehat{\mathbf{x}}$ is the reconstructed signal. We have the NMSE for different SNR values in Fig. 3. We can see that our method is more robust to noise.

3.2. Experiment for speech signals

We use 10 mono, 16-bit, 16kHz sampled speech signals from CMU-ARCTIC database to see how the proposed method works for none sparse signals. The speech signal is divided into frames with length $n = 64$ and overlap $o = 32$. The Hanning windowed speech frames make up a matrix, and its Fourier transform magnitude is the input of the proposed method. We apply the proposed method ($ITER = 4$) to get an approximately recovered matrix. By overlapping and adding each column of the recovered matrix, we can reconstruct a speech signal. We apply some minor revisions to the GESPAR algorithm, and test it with the same speech signal and same procedure to get a result too.

In Fig. 4, we compare the spectrograms of one signal reconstructed by our approach, the original speech signal, and the signal reconstructed by GESPAR. We can see that the phase retrieval result of our method recovers more details about pitch and formants. The NMSE of reconstructed signals of our method is also smaller than that of GESPAR as shown in Table 1. We conduct a listening test to compare the voice quality of the reconstructed signals. A score of range 1 to 5 (the default score of original signal is 5) is marked by 10 different listeners according to the voice quality. We also report the average score of the reconstructed signals in Table 1. We can see that listeners consistently favor our method.

4. CONCLUSIONS

We propose an effective phase retrieval algorithm for recovering a signal from its Fourier transform magnitude. The proposed method leads to accurate recovery of sparse signals with very high probability, and also works for none sparse signals with an approximate solution. Numerical simulations show that our method outperforms alternative approaches in terms of success probability and robustness to noise. Experiments for speech signals show that the voice quality of the reconstructed signal is almost as good as the original speech. Although our proposed algorithm has good empirical performance, we shall carry out theoretical analysis regarding the convergence of the iterative optimization for the phase retrieval problem in a non-convex setting.

5. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 61175018).

6. REFERENCES

[1] R. P. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.

[2] R. W. Harrison, "Phase problem in crystallography," *JOSA A*, vol. 10, no. 5, pp. 1046–1055, 1993.

[3] J. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, pp. 231–275, 1987.

[4] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, "Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes," *Annu. Rev. Phys. Chem.*, vol. 59, pp. 387–410, 2008.

[5] L. R. Rabiner and B.-H. Juang, *Fundamentals of speech recognition*. PTR Prentice Hall Englewood Cliffs, 1993, vol. 14.

[6] *INTERSPEECH 2014 Special Session: Phase Importance in Speech Processing Applications*, 2014.

[7] K. Paliwal, K. Wójcicki, and B. Shannon, "The importance of phase in speech enhancement," *speech communication*, vol. 53, no. 4, pp. 465–494, 2011.

[8] R. Schluter and H. Ney, "Using phase spectrum information for improved speech recognition performance," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 133–136.

[9] G. Degottex, A. Roebel, and X. Rodet, "Phase minimization for glottal model estimation," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 5, pp. 1080–1090, 2011.

[10] R. Maia, M. Akamine, and M. J. Gales, "Complex cepstrum as phase information in statistical parametric speech synthesis," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4581–4584.

[11] M. K. Watanabe and P. Mowlaee, "Iterative sinusoidal-based partial phase reconstruction in single-channel source separation." in *INTERSPEECH*, 2013, pp. 832–836.

[12] 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.

[13] J. L. Sanz, "Mathematical considerations for the problem of fourier transform phase retrieval from magnitude," *SIAM Journal on Applied Mathematics*, vol. 45, no. 4, pp. 651–664, 1985.

[14] Y. Shechtman, Y. C. Eldar, A. Szameit, and M. Segev, "Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing," *Optics express*, vol. 19, no. 16, pp. 14 807–14 822, 2011.

[15] K. Jaganathan, S. Oymak, and B. Hassibi, "Recovery of sparse 1-d signals from the magnitudes of their fourier transform," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium On*. IEEE, 2012, pp. 1473–1477.

[16] H. Ohlsson, A. Y. Yang, R. Dong, and S. S. Sastry, "Compressive phase retrieval from squared output measurements via semidefinite programming," *arXiv preprint arXiv*, vol. 1111, 2011.

[17] I. Waldspurger, A. dAspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[18] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, p. 237, 1972.

[19] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization," *JOSA A*, vol. 19, no. 7, pp. 1334–1345, 2002.

[20] V. Elser, "Phase retrieval by iterated projections," *JOSA A*, vol. 20, no. 1, pp. 40–55, 2003.

[21] S. Marchesini, "Invited article: A unified evaluation of iterative projection algorithms for phase retrieval," *Review of Scientific Instruments*, vol. 78, no. 1, p. 011301, 2007.

[22] R. Fan, Q. Wan, F. Wen, H. Chen, and Y. Liu, "Iterative projection approach for phase retrieval of semi-sparse wave field," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 24, 2014.

[23] Y. Shechtman, A. Beck, and Y. C. Eldar, "Gespar: Efficient phase retrieval of sparse signals," *Signal Processing, IEEE Transactions on*, vol. 62, no. 4, pp. 928–938, 2014.

[24] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 2796–2804.

[25] S. Mukherjee and C. S. Seelamantula, "An iterative algorithm for phase retrieval with sparsity constraints: application to frequency domain optical coherence tomography," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 553–556.

[26] A. Beck and Y. C. Eldar, "Sparsity constrained nonlinear optimization: Optimality conditions and algorithms," *SIAM Journal on Optimization*, vol. 23, no. 3, pp. 1480–1509, 2013.

[27] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.

[28] A. Björck, *Numerical methods for least squares problems*. Siam, 1996.

[29] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Advances in neural information processing systems*, 2006, pp. 801–808.