

# Group Sparsity Based Spectrum Estimation of Harmonic Speech Signals

Yimin D. Zhang<sup>†</sup> and Ben Wang<sup>‡</sup>

<sup>†</sup> Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA

<sup>‡</sup> College of Automation, Harbin Engineering University, Harbin, Heilongjiang 150001, China

## ABSTRACT

Spectrum analysis of speech signals is important for their detection, recognition, and separation. Speech signals are nonstationary with time-varying frequencies which, when analyzed by Fourier analysis over a short time window, exhibit harmonic spectra, i.e., the fundamental frequencies are accompanied by multiple associated harmonic frequencies. With proper modeling, such harmonic signal components can be cast as group sparse and solved using group sparse signal reconstruction methods. In this case, all harmonic components contribute to effective signal detection and fundamental frequency estimation with improved reliability and spectrum resolution. The estimation of the fundamental frequency signature is implemented using the block sparse Bayesian learning technique, which is known to provide high-resolution spectrum estimations. Simulation results confirm the superiority of the proposed technique when compared to the conventional STFT-based methods.

## 1. INTRODUCTION

Speech signals and, more generally, frequency modulated (FM) nonstationary signals, are widely encountered in radar, sonar, biomedicine, and speech signal processing. For such signals, because of their time-varying characteristics, traditional Fourier transform cannot provide high-resolution spectrum analysis and reveal time-varying instantaneous frequency signatures. Time-frequency analysis is an effective technique to process such nonstationary signals and thus has attracted significant interests [1, 2]. In particular, the short-time Fourier transform (STFT) is a commonly used linear time-frequency analysis method to analyze such nonstationary signals. STFT divides the signals into several overlapping segments with a fixed window, and the window width trades off between the time and frequency resolutions. A short window leads to a high time resolution but yields a poor frequency resolution, whereas a long window improves the frequency resolution but reduces the time resolution. The short-time Fan-Chirp transform (STFChT) is also commonly used for speech signal analysis [3, 4]. The STFChT can reach the limit of the time-frequency uncertainty principle when the signals are chirp-periodic.

On the other hand, compressed sensing and sparse reconstruction techniques have achieved great success in recent years [5, 6]. Numerous algorithms, such as orthogonal matching pursuit (OMP) [7], Lasso [8], and sparse Bayesian learning [9, 10], have been developed toward this purpose. Among different Sparse Bayesian learning or Bayesian compressive sensing techniques generally provide robust and high-resolution sparse signal reconstruction, and permit flexible treatment of sparse signal structures through prior designs (e.g., [11–13]).

In addition to sparsity, many real-world signals exhibit the so-called group sparsity, i.e., elements belong to the same group have the same sparse support whereas their values generally differ. Such group sparse (also referred to as block sparse) problems can be similarly solved by a number of methods, such as block OMP [14], group Lasso [15], multi-task compressive sensing (mt-CS) [16, 17], and block sparse Bayesian learning (BSBL) [18]. Group sparsity allows exploitation of multiple observations for better determination of the sparse signal support and thus achieve improved sparse reconstruction performance as compared to conventional methods without considering such group sparsity.

In this paper, we formulate the spectrum estimation of harmonic FM signals, with the focus on speech signals, as a group sparse signal reconstruction problem, which is then effectively solved using the BSBL methods.

---

Contact information: yimin.zhang@villanova.edu.

The group sparsity arises from the fact that the harmonic frequencies always appear as integer multiples of the fundamental frequency, which is also called the first harmonic [19]. In the BSBL framework, all harmonic components contribute to effective signal detection and spectrum estimation with improved reliability and spectrum resolution.

The remainder of this paper is organized as follows. In Section 2, we present the Gaussian-enveloped harmonic chirp signal of interest and a brief review of the STFT, then we derive the block sparse Bayesian reconstruction algorithm. Section 3 utilizes the group sparse property of the harmonic chirp signal in their time-frequency domain to estimate the fundamental frequency. Simulations in Section 4 demonstrate the effectiveness of the proposed method and finally, the conclusion is given in Section 5.

Notations: We use lower-case (upper-case) bold characters to denote vectors (matrices). In particular,  $\mathbf{I}_L$  denotes the  $L \times L$  identity matrix.  $(\cdot)^T$  and  $(\cdot)^H$  respectively denote the transpose and conjugate transpose of a matrix or vector.  $\text{diag}(\mathbf{x})$  represents a diagonal matrix that uses the elements of  $\mathbf{x}$  as its diagonal elements, and  $\text{bdiag}\{\mathbf{A}_1, \dots, \mathbf{A}_U\}$  denotes a block diagonal matrix with principal diagonal blocks  $\mathbf{A}_1, \dots, \mathbf{A}_U$  being in turn.  $\mathcal{CN}(\mathbf{a}, \mathbf{B})$  represents multivariate complex Gaussian distribution with mean vector  $\mathbf{a}$  and covariance matrix  $\mathbf{B}$ .

## 2. PROBLEM FORMULATION

In this section, we introduce the time-frequency analysis of a harmonic nonstationary signal using the STFT, and then formulate the spectrum estimation of nonstationary harmonic signals as a group sparse signal reconstruction problem.

### 2.1 Short-time Fourier transform of harmonic nonstationary signals

Consider a general model of a harmonic nonstationary signal in the following form:

$$s(t) = \sum_{h=1}^H \sum_{k=1}^K a_{h,k}(t) \exp(jk\phi_h(t)), \quad (1)$$

where  $H$  is the number of the fundamental frequency components,  $K$  is the maximum order of the harmonics,  $a_{h,k}(t)$  is a slowly time-varying coefficient of the  $k$ th harmonic corresponding to the  $h$ th fundamental frequency, and  $\phi_h(t)$  is the time-varying phase of the  $h$ th fundamental frequency component. In this paper, we only consider a single signal component, i.e.,  $H = 1$ . In this case, the above formulation is simplified as

$$s(t) = \sum_{k=1}^K a_k(t) \exp(jk\phi(t)). \quad (2)$$

We sample the signal with a sample frequency  $f_s$ , and the measurement time is  $T$ . Thus, we can obtain  $N = f_s T$  discrete-time samples as  $s[1], s[2], \dots, s[N]$ . The noisy observations can be expressed as

$$x[i] = s[i] + v[i], \quad i = 1, \dots, N, \quad (3)$$

where  $v[i]$  is additive white Gaussian noise vector with zero mean and variance  $\sigma$ .

Let us now consider the STFT of a harmonic FM signal. A time-domain sequence  $s[i]$ ,  $i = 1, \dots, N$ , is divided into  $D$  overlapping segments and each segment has  $M$  elements. The segment-to-segment shift (frame hop) is  $S$  samples, where  $1 \leq S < M$ . Thus, the number of overlapping samples between two adjacent segments is  $M - S$ . As such, we obtain an  $M \times D$  matrix of the observed data, expressed as:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_D] = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_D] + [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_D], \quad (4)$$

where  $\mathbf{x}_d = [x[(d-1)S+1], x[(d-1)S+2], \dots, x[(d-1)S+M]]^T$  is the  $d$ th segment,  $1 \leq d \leq D$ . The STFT basis matrix  $\mathbf{\Psi}$  is expressed as a weighted Fourier transform matrix, given by

$$\mathbf{\Psi} = \mathbf{W} \mathbf{\Pi}, \quad (5)$$

where  $\mathbf{\Pi}$  is the  $M$ -point Fourier transform matrix, and  $\mathbf{W} = \text{diag}\{\mathbf{w}\}$  with  $\mathbf{w} = [w_1, w_2, \dots, w_M]^T$  denoting an  $M$ -length window vector. Thus, the STFT vector  $\mathbf{y}_d$  is related with  $\mathbf{x}_d$  through the weighted Fourier transform matrix, expressed as

$$\mathbf{y}_d = \mathbf{\Psi}\mathbf{x}_d. \quad (6)$$

We can also obtain the following relationship between  $\mathbf{Y}$  and  $\mathbf{X}$  with all  $D$  segments accounted for:

$$\mathbf{Y} = \mathbf{\Psi}\mathbf{X}. \quad (7)$$

Consider an example of harmonic FM signal which consists of two successive Gaussian-enveloped harmonic chirps with duration 0.8 second and are spaced 0.4 second apart [21]. Both harmonic chirps have  $K=10$  harmonic components. The fundamental frequency of the first harmonic chirp starts from 200 Hz and rises to 230 Hz, whereas that of the second harmonic chirp varies from 245 Hz to 210 Hz. The sampling frequency is  $f_s = 8$  KHz, and a Hamming window with a duration of 2048 samples is used. The 2048-point fast Fourier transform (FFT) is used, and the frame hop is set to 128 samples. Fig. 1(a) shows the time-domain waveforms of this signal, and Fig. 1(b) shows the STFT magnitude.

From this result, the following observations are made for the direct application of STFT to the harmonic FM signals:

- (a) The STFT of the high-order harmonic signal components becomes broadened with smeared time-frequency distribution. It makes difficult to analysis the spectrum of harmonic signal components for their detection, recognition and separation.
- (b) As each harmonic signal component is independently analyzed and estimated, the intrinsic relationship that the harmonic frequencies are integer multiples of the fundamental frequency is not used in the spectrum estimation.

In the next subsection, we formulate the spectrum estimation of harmonic FM signals as a group sparse signal reconstruction problem. As a result, all the harmonic signal components are exploited in the spectrum estimation of each harmonic component, and a high-resolution spectrum estimation is achieved.

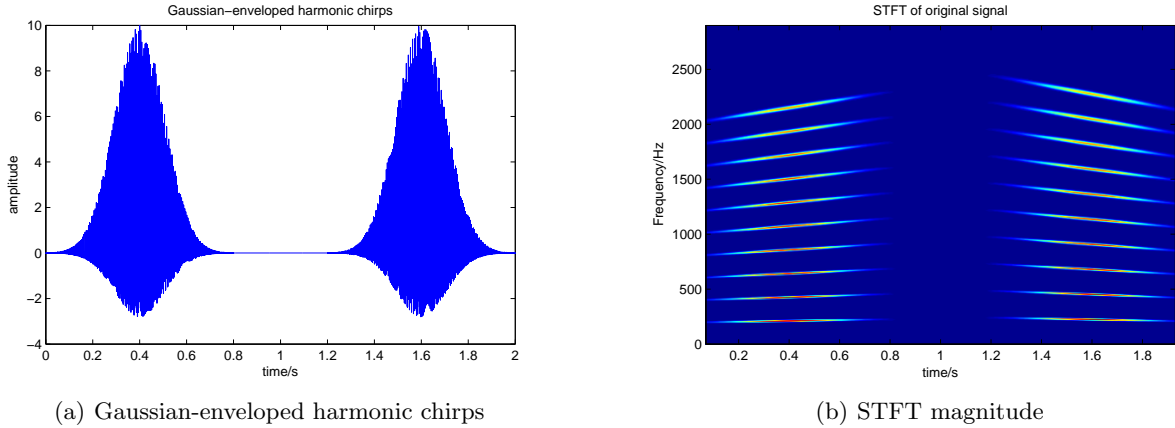


Figure 1. Two successive Gaussian-enveloped harmonic chirps.

## 2.2 Group-sparse representation of harmonic nonstationary signals

In Eq. (6), vector  $\mathbf{y}_d$  describes the frequency spectrum of the signal in the  $d$ th segment. Without considering the noise, most of the elements in  $\mathbf{y}_d$  are zeroes. In addition, the spectra of the harmonic signal components occur in positions of the integer multiples of the fundamental frequency. These facts enable the formulation of

the harmonic spectrum estimation as a group sparsity based signal reconstruction. Toward this end, we first express the  $d$ th noise-free signal vector,  $\mathbf{s}_d$ , as

$$\mathbf{s}_d = \mathbf{\Gamma}_d \boldsymbol{\lambda}_d, \quad (8)$$

where

$$\mathbf{\Gamma}_d = \begin{bmatrix} 1 & \cdots & 1 \\ e^{j\omega_d} & \cdots & e^{jK\omega_d} \\ \vdots & \ddots & \vdots \\ e^{j(M-1)\omega_d} & \cdots & e^{j(M-1)K\omega_d} \end{bmatrix}, \quad (9)$$

$$\boldsymbol{\lambda}_d = [a_1(d), a_2(d), \dots, a_K(d)]^T. \quad (10)$$

Note that both the magnitude,  $a_k(d)$ , and the normalized fundamental radian frequency,  $\omega_d$ , are considered time-invariant within the  $d$ th segment, for  $k = 1, \dots, K$ , because of the short time period considered in each segment.

To solve this problem under the group sparsity framework, we consider a fundamental frequency grid  $[\omega^{(1)}, \omega^{(2)}, \dots, \omega^{(U)}]^T$  over the maximum possible fundamental frequency range  $[\omega_{\min}, \omega_{\max}]$ . In this case, the harmonic frequency components can be expressed as

$$\mathbf{s}_d = \tilde{\mathbf{\Gamma}} \tilde{\boldsymbol{\lambda}}_d, \quad (11)$$

where

$$\tilde{\mathbf{\Gamma}} = [\tilde{\mathbf{\Gamma}}_1, \tilde{\mathbf{\Gamma}}_2, \dots, \tilde{\mathbf{\Gamma}}_U] \quad (12)$$

and

$$\tilde{\boldsymbol{\lambda}}_d = [\tilde{\boldsymbol{\lambda}}_1^T(d), \tilde{\boldsymbol{\lambda}}_2^T(d), \dots, \tilde{\boldsymbol{\lambda}}_U^T(d)]^T \quad (13)$$

respectively represent the measurement (dictionary) matrix and the sparse vector of the group sparse reconstruction problem. In particular,  $\tilde{\mathbf{\Gamma}}_u$  and  $\tilde{\boldsymbol{\lambda}}_u$  corresponding to the hypothetical frequency component at the  $u$ th grid entry are respectively expressed as

$$\tilde{\mathbf{\Gamma}}_u = \begin{bmatrix} 1 & \cdots & 1 \\ e^{j\omega_u} & \cdots & e^{jK\omega_u} \\ \vdots & \ddots & \vdots \\ e^{j(M-1)\omega_u} & \cdots & e^{j(M-1)K\omega_u} \end{bmatrix}, \quad (14)$$

and

$$\tilde{\boldsymbol{\lambda}}_u(d) = [\tilde{a}_u^{(1)}(d), \tilde{a}_u^{(2)}(d), \dots, \tilde{a}_u^{(K)}(d)]^T. \quad (15)$$

Note that vector  $\tilde{\boldsymbol{\lambda}}_d$  in (13) is group sparse in the sense that (a) most of its elements take zero values; and (b) elements in  $\tilde{\boldsymbol{\lambda}}_u(d)$  tend to be present or absent as a group. That is, for some  $u_0$ , if  $\tilde{a}_{u_0}^{(1)}(d)$  takes a nonzero value, i.e.,  $\omega_{u_0}$  is a valid fundamental frequency, then  $\tilde{a}_u^{(k)}(d)$  for other values of  $k$  are also likely to take nonzero values, as they are the coefficients of the associated harmonic signal components at frequencies  $k\omega_{u_0}$  for  $k = 2, \dots, K$ . In the next section, we will use the BSBL method to effectively recover the positions and values of the sparse frequency coefficients, thus achieving harmonic spectrum estimation based on all harmonic signals components.

### 3. GROUP SPARSE RECONSTRUCTION ALGORITHM

We write the formulation of  $\tilde{\boldsymbol{\lambda}}_d$  in the below:

$$\tilde{\boldsymbol{\lambda}}_d = \underbrace{[\tilde{a}_1^{(1)}(d), \tilde{a}_1^{(2)}(d), \dots, \tilde{a}_1^{(K)}(d)]}_{\tilde{\boldsymbol{\lambda}}_1^T(d)}, \underbrace{[\tilde{a}_2^{(1)}(d), \tilde{a}_2^{(2)}(d), \dots, \tilde{a}_2^{(K)}(d)]}_{\tilde{\boldsymbol{\lambda}}_2^T(d)}, \dots, \underbrace{[\tilde{a}_U^{(1)}(d), \tilde{a}_U^{(2)}(d), \dots, \tilde{a}_U^{(K)}(d)]}_{\tilde{\boldsymbol{\lambda}}_U^T(d)}^T. \quad (16)$$

As such,  $\tilde{\boldsymbol{\lambda}}_d$  clearly show the group sparsity as depicted in [20]. There are a number of methods that solve such group sparse problems. In this paper, we use the BSBL algorithm [20] due to its superior performance when

compared to block orthogonal matching pursuit (BOMP) [14] and group Lasso (gLasso) [15]. Note that the original BSBL was developed for real-valued problems, and is extended to complex-valued formulation to solve our problem.

Rewrite (11) as follows with the noise vector  $\mathbf{v}_d$  added:

$$\mathbf{x}_d = \mathbf{s}_d + \mathbf{v}_d = \tilde{\mathbf{\Gamma}}\tilde{\boldsymbol{\lambda}}_d + \mathbf{v}_d. \quad (17)$$

Assume that each block  $\tilde{\boldsymbol{\lambda}}_i(d) \in \mathbb{R}^K$  satisfies the following parameterized multivariate Gaussian distribution:

$$p(\tilde{\boldsymbol{\lambda}}_i(d); \gamma_i, \mathbf{B}_i) \sim \mathcal{CN}(\mathbf{0}, \gamma_i \mathbf{B}_i), i = 1, \dots, U, \quad (18)$$

where  $\gamma_i$  and  $\mathbf{B}_i$  are unknown parameters. Here,  $\gamma_i$  is a nonnegative parameter controlling the block-sparsity of vector  $\tilde{\boldsymbol{\lambda}}_i(d)$ . In particular, it becomes zero when  $\gamma_i = 0$ . In addition,  $\mathbf{B}_i \in \mathbb{R}^{K \times K}$  is a positive definite matrix that captures the covariance matrix structure of vector  $\tilde{\boldsymbol{\lambda}}_i(d)$ . In our paper, the blocks are considered mutually uncorrelated and, as such, the prior of  $\tilde{\boldsymbol{\lambda}}_d$  can be represented as  $p(\tilde{\boldsymbol{\lambda}}_d; \gamma_i, \mathbf{B}_i) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_0)$ , where  $\boldsymbol{\Sigma}_0 = \text{bdiag}\{\gamma_1 \mathbf{B}_1, \dots, \gamma_U \mathbf{B}_U\}$ . Let the noise vector satisfies  $p(\mathbf{v}_d; \sigma) \sim \mathcal{N}(\mathbf{0}, \sigma \mathbf{I}_M)$ , where  $\sigma$  is a positive scalar representing the noise variance. Thus, the posterior of  $\tilde{\boldsymbol{\lambda}}_d$  can be obtained as

$$p(\tilde{\boldsymbol{\lambda}}_d | \mathbf{x}_d; \sigma, \{\gamma_i, \mathbf{B}_i\}_{i=1}^U) = \mathcal{CN}(\boldsymbol{\mu}_{\lambda_d}, \boldsymbol{\Sigma}_{\lambda_d}) \quad (19)$$

where

$$\boldsymbol{\mu}_{\lambda_d} = \boldsymbol{\Sigma}_0 \tilde{\mathbf{\Gamma}}^H (\sigma \mathbf{I}_M + \tilde{\mathbf{\Gamma}} \boldsymbol{\Sigma}_0 \tilde{\mathbf{\Gamma}}^H)^{-1} \mathbf{x}_d, \quad (20)$$

$$\boldsymbol{\Sigma}_{\lambda_d} = \left( \boldsymbol{\Sigma}_0^{-1} + \frac{1}{\sigma} \tilde{\mathbf{\Gamma}}^H \tilde{\mathbf{\Gamma}} \right)^{-1}. \quad (21)$$

Utilizing the Expectation Maximization (EM) method, we can derive the learning rules for  $\mathbf{B}_i$ ,  $\gamma_i$  and  $\sigma$ :

$$\mathbf{B}_i \leftarrow \frac{1}{U} \sum_{i=1}^U \frac{\boldsymbol{\Sigma}_{\lambda_d}^i + \boldsymbol{\mu}_{\lambda_d}^i (\boldsymbol{\mu}_{\lambda_d}^i)^T}{\gamma_i} \quad (22)$$

$$\gamma_i \leftarrow \frac{1}{K} \text{Tr} \left[ \mathbf{B}_i^{-1} \left( \boldsymbol{\Sigma}_{\lambda_d}^i + \boldsymbol{\mu}_{\lambda_d}^i (\boldsymbol{\mu}_{\lambda_d}^i)^T \right) \right], i = 1, \dots, U \quad (23)$$

$$\sigma \leftarrow \frac{\left\| \mathbf{x}_d - \tilde{\mathbf{\Gamma}} \boldsymbol{\mu}_{\lambda_d} \right\|_2^2 + \sum_{i=1}^U \text{Tr} \left( \boldsymbol{\Sigma}_{\lambda_d}^i (\tilde{\mathbf{\Gamma}}^i)^T \tilde{\mathbf{\Gamma}}^i \right)}{M} \quad (24)$$

where  $\boldsymbol{\mu}_{\lambda_d}^i \in \mathbb{R}^{K \times 1}$  is the corresponding  $i$ th block in  $\boldsymbol{\mu}_{\lambda_d}$ ,  $\boldsymbol{\Sigma}_{\lambda_d}^i \in \mathbb{R}^{K \times K}$  is the corresponding  $i$ th principal diagonal block in  $\boldsymbol{\Sigma}_{\lambda_d}$ , and  $\tilde{\mathbf{\Gamma}}^i \in \mathbb{R}^{M \times K}$  is the submatrix of  $\tilde{\mathbf{\Gamma}}$  which corresponds to the  $i$ th block of  $\mathbf{x}_d$ .

After the parameters  $\sigma, \{\gamma_i, \mathbf{B}_i\}_{i=1}^U$  are calculated, the maximum *a posteriori* (MAP) estimate of  $\tilde{\boldsymbol{\lambda}}_d$  can be obtained from the mean of the posterior, expressed as

$$\hat{\boldsymbol{\lambda}}_d \leftarrow \boldsymbol{\Sigma}_0 \tilde{\mathbf{\Gamma}}^H (\sigma \mathbf{I}_M + \tilde{\mathbf{\Gamma}} \boldsymbol{\Sigma}_0 \tilde{\mathbf{\Gamma}}^H)^{-1} \mathbf{x}_d. \quad (25)$$

#### 4. SIMULATION RESULTS

In this section, we provide simulation results to demonstrate the effectiveness of the proposed technique. We use the successive harmonic chirps as described in Section 2.1 and depicted in Fig. 1 as example.

Fig. 2 shows the time-frequency domain representations of the recovery result of the first 0.8 second of the data covering the first harmonic chirp signal utilizing the proposed method and the STFChT, respectively. Complex Gaussian noise is added to yield an average signal-to-noise ratio (SNR) of 5 dB. The fundamental frequency grid used in the dictionary matrix  $\tilde{\mathbf{\Gamma}}$  is between 180 Hz and 270 Hz with a 1 Hz increment, yielding  $U = 91$  group entries in the sparse spectrum. We assume 15 hypothetic harmonic components in the group sparse reconstruction, which is higher than the actual harmonic order of 10. In both figures, we divide the signal into multiple segments where each segment has 512 hamming-windowed samples. The number of overlapped samples between two adjacent segments is 292. Compared with the STFChT method, the proposed method has a much higher frequency resolution.

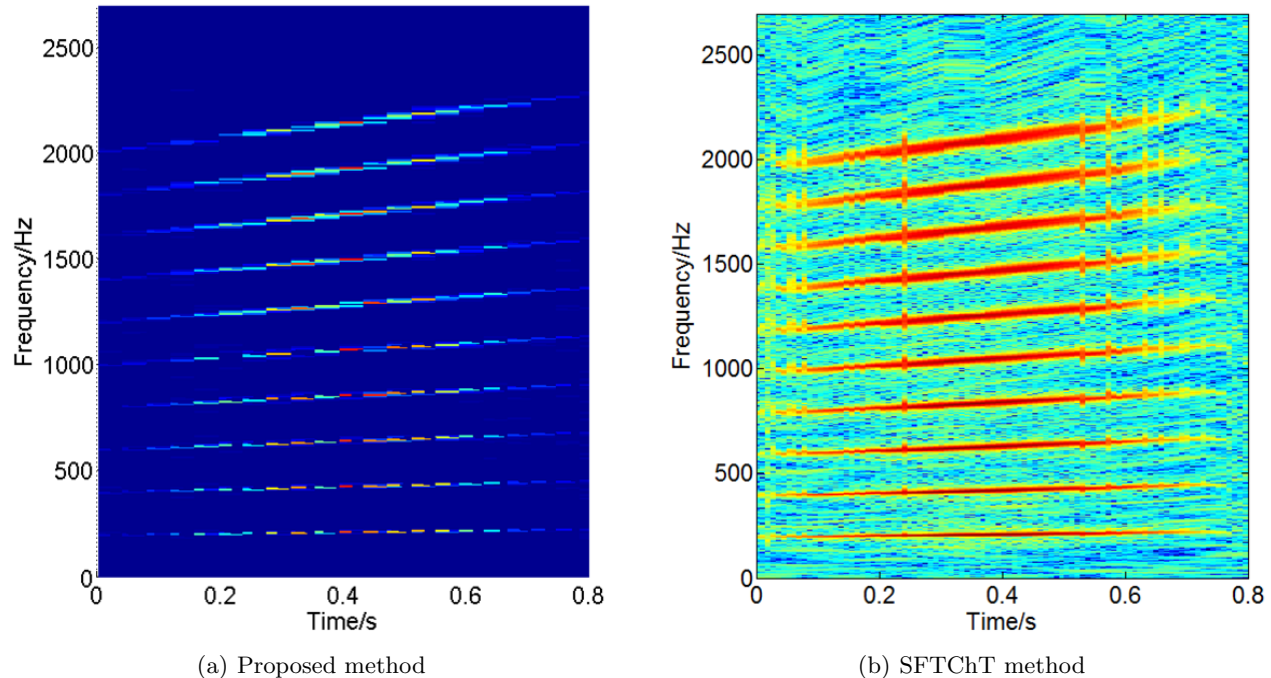


Figure 2. Spectrum estimation utilizing proposed method and STFChT method.

## 5. CONCLUSION

In this paper, we proposed a novel spectrum estimation technique for harmonic frequency modulated signals with time-varying fundamental frequencies. The proposed method casts the spectrum estimation as a group sparse signal reconstruction problem where all the harmonic spectra are jointly estimated using all available harmonic signal components. Compared with the conventional STFT and STFChT methods, the proposed method achieves improved frequency resolution and performance.

## ACKNOWLEDGMENT

The authors would like to thank Scott Wisdom, University of Washington in Seattle, for generously providing the MATLAB code for performing the STFChT algorithm for the Gaussian-enveloped harmonic chirps.

## REFERENCES

- [1] L. Cohen, *Time Frequency Analysis: Theory and Applications*. Upper Saddle River, NJ: Prentice Hall, 1994.
- [2] B. Boashash (ed.), *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Amsterdam, Netherlands: Elsevier, 2003.
- [3] M. Képesi and L. Weruaga, “Adaptive chirp-based time-frequency analysis of speech signals,” in *Speech Communication*, vol. 48, no. 5, pp. 474–492, 2006.
- [4] L. Weruaga and M. Képesi, “The fan-chirp transform for non-stationary harmonic signals,” in *Signal Processing*, vol. 87, no. 6, pp. 1504–1522, 2007.
- [5] D. L. Donoho, “Compressed sensing,” in *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [6] E. Candes and M. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–C30, 2008.

- [7] J. A. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [8] R. Tibshirani, "Regression shrinkage and selection via the Lasso," in *Journal of the Royal Statistical Society*, vol. 56, no. 1, pp. 267–288, 1996.
- [9] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [10] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.
- [11] Q. Wu, Y. D. Zhang, and M. G. Amin, "Continuous structure based Bayesian compressive sensing for sparse reconstruction of time-frequency distributions," in *Proceedings of International Conference on Digital Signal Processing*, Hong Kong, China, pp. 831–836, Aug. 2014.
- [12] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5736C–5750, 2014.
- [13] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Multi-task Bayesian compressive sensing exploiting intra-task correlation," *IEEE Signal Processing Letters*, vol. 22, no. 4, pp. 430–434, April 2015.
- [14] Y. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signal: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing* vol. 58, no. 6, pp. 2009–2015, June 2010.
- [15] L. Jacob, G. Obozinski, and J. Vert, "Group Lasso with overlap and graph Lasso," in *Proceedings of International Conference on Machine Learning*, Montreal, QC, June 2009.
- [16] S. Ji, D. Dunson and L. Carin, "Multitask compressive sampling," in *IEEE Transactions on Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.
- [17] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Complex multitask Bayesian compressive sensing," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Florence, Italy, pp. 3375–3379, May 2014.
- [18] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 57, no. 1, pp. 92–106, 2009.
- [19] L. J. Raphael, G. J. Borden, and K. S. Harris, *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*, Fifth Edition. Baltimore, MD: Lippincott Williams & Wilkins, 2007.
- [20] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2009–2015, April 2013.
- [21] S. Wisdom, T. Powers, L. Atlas, and J. Pitton, "Enhancement of reverberant and noisy speech by extending its coherence," in *Proceedings of REVERB Challenge Workshop*, Florence, Italy, pp. 1–8, 2014.