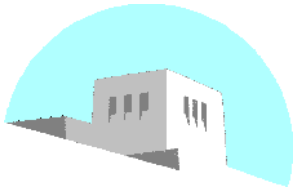
11-15-2014

Wideband-FM Demodulation for Large Wideband to Narrowband conversion factors Via Multirate Frequency Transformations

Wenjing Liu

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING



SCHOOL OF ENGINEERING
UNIVERSITY OF NEW MEXICO

Wideband-FM Demodulation for Large Wideband to Narrowband conversion factors Via Multirate Frequency Transformations

Wenjing Liu and Balu Santhanam[1]
Department of Electrical & Computer Engineering
University of New Mexico, Albuquerque, NM: 87131
Tel: 505 277-1611, Fax: 505 277 1439
Email: wenjing@unm.edu, bsanthan@unm.edu

UNM Technical Report: EECE-TR-14-0001

Report Date: November 15, 2014

[1]This research is supported by the Airforce Research Laboratory through FA9453-14-1-0234, Steven Lane, USAF
Email:steven.lane.1@us.af.mil

Abstract

Existing approaches for wideband FM demodulation are based on negative feedback, frequency tracking or multirate signal processing and heterodyning. Prior work that utilizes multirate frequency transformations for wideband-FM demodulation is impractical for large wideband to narrowband conversion factors such as those needed in DRFM systems. In this paper, we present a frequency transformation approach to wideband FM demodulation, using multirate systems, that can accommodate large conversion factors. We further apply the technique towards demodulation of wideband formants in speech to demonstrate its validity.

Keywords

Wideband signal, frequency demodulation, multirate systems, heterodying, adaptive linear prediction

1 Introduction

Frequency modulation signals are time-varying sinusoids of the form:

$$s(t) = A\cos\left(\int_{-\infty}^{t} \omega_i(\tau)\, d\tau + \theta_1\right), \tag{1}$$

where the instantaneous amplitude remains constant while the *instantaneous frequency* (IF) is given by

$$\omega_i(t) = \omega_c + \omega_m q_i(t), \tag{2}$$

where $q_i(t)$ is the normalized baseband modulated signal, and for sinusoidal FM it becomes

$$\omega_i(t) = \omega_c + \omega_m cos(\omega_f t + \theta_2). \tag{3}$$

Sinusoidal FM signals can be expressed via Bessel function as

$$s(t) = A\sum_{n=-\infty}^{+\infty} J_n(\beta)\cos(\omega_c t + n\omega_m t), \tag{4}$$

where J_n is the n^{th} order cylindrical Bessel function of the first kind. The modulation index of sinusoidal FM is defined as the ratio $\beta = \omega_m/\omega_f$ and the associated carson's bandwidth is given by

$$B = 2(\beta+1)\omega_f. \tag{5}$$

If $\beta \gg 1$, then it corresponds to the traditional wideband FM setting, where the *carrier-to-information-bandwidth ratio* (CR/IB) and the *carrier-to-frequency deviation ratio* (CR/FD) are defined respectively as

$$\frac{CR}{IB} = \frac{\omega_c}{\omega_f} \quad \frac{CR}{FD} = \frac{\omega_c}{\omega_m}. \tag{6}$$

Signals in this category are widely used in applications such as speech formant tracking and analysis [1], satellite communication [2] and DRFM system [3]. However, most conventional demodulation techniques only perform well under the narrowband assumptions where the two parameters CR/IB and CR/FD are usually small compared to the carrier freqeuncy, and fail in the wideband setting. For example, the Hilbert transform requires a relatively large carrier frequency of the signal to form an accurate analytic signal approximation. Larger ratios of CR/IB and CR/FD are beneficial for the analytic approximation and hence help to improve the demodulation. In this paper, a general approach that involves multirate systems as well as heterodyning is proposed for wideband FM demodulation that usually has a: 1) $\beta > 2$; 2) large information bandwidth. The adaptive linear predictive IF tracking technique as described in [4] is chosen as the demodulation method for implementing the proposed approach. In fact, it can be integrated with other existing FM demodulation methods, such as the zero-crossing approach according to [5], [6] and feedback demodulation as mentioned in [7], serving as a general framework for dealing with wideband FM. In this paper, we demonstrate that large wideband to narrowband conversion factors are feasible using the proposed system with designs that are realizable.

2 Multirate and Heterodyning System

The motivation for incorporating the multirate and heterodying systems into the demodulation framework is to apply *multirate frequency transformations* (MFT) that first compress the spectrum of wideband FM signals and then shift them into an optimal region in terms of CR/IB, CR/FD ratios, where existing demodulation techniques perform well.

2.1 Prior Work

Using the scaling property of the Fourier transform, compression in frequency domain is equivalent to expansion in the time domain expressed as

$$y(t) = x(at) \Longleftrightarrow Y(\omega) = X\left(\frac{\omega}{a}\right), \tag{7}$$

where $a = 1/R < 1$ is the factor of frequency compression. Then IF of the compressed signal becomes a scaled version of the input IF by a factor R expressed as

$$\tilde{\omega}_i(t) = \frac{\omega_i(t)}{R} = \frac{\omega_c}{R} + \frac{\omega_m}{R} q_i\left(\frac{t}{R}\right). \tag{8}$$

Note that for compressed signal, the carrier frequency is also scaled by the same factor R, which is undesirable since the ratios CR/FD and CR/IB that we wish to increase still remains invariant. Hence the heterodying operator is cascaded right after the compression process in order to upshift the carrier frequency to a higher level where we can attain larger CR/FD and CR/IB ratios. The compressed signal after frequency translation and bandpass filtering in the heterodying module is given by

$$y_{ush}(t) = \left[y(t)\cos(\omega_d t)\right] * h_{BPF}(t), \tag{9}$$

where $*$ denotes the convolution, ω_d refers to the amount of frequency translation and $h_{BPF}(t)$ represents the impulse response of the bandpass filter. Specific for the case of sinusoidal FM, it can be further simplified as

$$y_{ush}(t) \simeq \frac{1}{2} A\cos\left(\phi\left(\frac{t}{R}\right)\right), \tag{10}$$

where $\phi(t)$ denotes the phase of the original FM signal. The resultant signal that has a scaled information bandwidth with a higher CR/IB then passes through the demodulation block for IF extraction. Eventually, the IF estimate of the original signal is evaluated by

$$\omega_i^{\text{out}}(t) = R\left(\tilde{\omega}_i(Rt) - \omega_d\right), \tag{11}$$

where $\tilde{\omega}_i(t)$ is the IF of the compressed signal.

As for discrete-time signals, compression and expansion can be substituted by the corresponding multirate operations of interpolation and decimation as described in [8] with their properties carried over to their discrete counterparts. The block diagram of such a MFT framework in prior work [9] is illustrated by Fig. 1.

Interpolating the input signal will result in the reduction of both the frequency deviation and information bandwidth by a factor of R. Similar to increasing the sampling rate, the IF of the interpolated signal becomes slow-varying and the assumption that the message signal remains constant over the carrier period is more likely to hold, which will boost the performances of conventional demodulation algorithms. Meanwhile, heterodying serves the purpose of increasing the ratios of CR/FD and CR/IB by compensating for the scaled carrier frequency. By passing through the heterodying process, the CR/FD and the CR/IB of $y_{ush}(t)$ are given by

$$\left[\frac{CR}{FD}\right]_{out} = \left[\frac{CR}{FD}\right]_{in} + \frac{R\omega_d}{\omega_m} \tag{12}$$

$$\left[\frac{CR}{IB}\right]_{out} = \left[\frac{CR}{IB}\right]_{in} + \frac{R\omega_d}{\omega_f}. \tag{13}$$

2.2 Constraints of Prior MFT System

As we look further into this framework, an important question regarding the selection of the conversion factor R arises. Specifically prior work [9] only deals with small multirate compression factors. However, larger factors

over hundred or even thousand can be supported by current high-speed DSP with large memory, as in the case of DRFM system design. It is intuitive to expect a further reduction in the demodulation error since the gain brought by frequency compression should be extendable through the use of a larger factor. But for really large factors R, the passband of the lowpass filter in the multirate operation and that of the heterodyne-BPF operation will be scaled by R. For example, if $R = 1000$, we require a lowpass filter with cut-off frequency at $\pi/1000$ and a bandpass filter with a passband edge less than or equal to of $\pi/1000$. However, filters with such narrow passbands are unrealistic for direct implementation by any structure [1]. Thus the design of BPF within the previous MFT framework becomes the bottleneck that constrains the use of a very large factor.

2.3 Proposed System

In order to reduce the burden placed on the practical implementation of the bandpass filter, we first consider a different MFT framework where the order of the interpolation operator and the heterodyning operator are exchanged. Due to the switch of interpolation and heterodying, the CR/FD and CR/IB parameters under this MFT framework are given by

$$\left[\frac{CR}{FD}\right]_{out} = \left[\frac{CR}{FD}\right]_{in} + \frac{\omega_d}{\omega_m} \tag{14}$$

$$\left[\frac{CR}{IB}\right]_{out} = \left[\frac{CR}{IB}\right]_{in} + \frac{\omega_d}{\omega_f}. \tag{15}$$

By comparing these ratios with (12) and (13), note that the upshift frequency ω_d in this case needs to be large enough such that the ratios of CR/FD and CR/IB still stay at high level. However, ω_d cannot be too large such that the resultant carrier frequency after heterodying exceeds one half of the sampling rate, we otherwise will need to interpolate the signal first by an appropriate factor in order to perform discrete-time bandpass filtering after heterodyning. Hence the practical implementation of MFT framework for a large conversion factor is not as simple as just exchanging the order of interpolation and heterodyning. Actually, an interpolation operation is still required prior to the heterodying with an appropriate factor that depends on the upshift frequency ω_d and the sampling frequency of the original wideband FM signal. This implies that the overall interpolation factor can be split into two with the first one prior to the frequency translation and the other one right after. Then upshifting by a frequency ω_d that is not too large would result in a relatively small factor for the first interpolation, thereby lessening the burden of the heterodyne-BPF.

In this paper, we propose a MFT framework for large conversion factors illustrated in Fig. 2(a) that achieves a practical design of the heterodyne-BPF. As previously discussed, the interpolation module of prior framework is separated into two with one in fornt of the heterodyning module and the other one right after in the proposed framework. The first interpolation module has a relatively small upsampling rate of R_1 which is appropriately chosen such that the discrete BPF can be implemented within the range of half the sampling rate after heterodyning the signal with a frequency translation of ω_d. The relatively small R_1 would result in a wider passband for the discrete bandpass filter, thus reducing the its design of complexity. In general, there is a sacrifice in terms of achievable CR/IB and CR/FD ratios for the proposed MFT framework, however, the system suggested in prior work does not realize large conversion factors, due to the placement of impractical constraints on the BPF design.

To further relax the constraint imposed on filter design, we first note that the lowpass filters in the multirate structure can be implemented efficiently using the multirate noble identities mentioned in [8]. As shown by Fig. 2(b), the original one stage multirate structure is equivalent to multistage implementation with R factorized into smaller integer factor corresponding to each stage. As a result, the lowpass filter corresponding to each stage will have much larger cut-off frequency and thus a wider passband. Therefore we can conclude that the use of a large conversion factor is primarily constrained by unrealistic requirements on the heterodyne-BPF, which can not be easily relaxed as in the case of the lowpass filter within the multirate structure.

[1] Narrow passband implies clustered poles and zeros that result in sensitivity and stability issues of digital filters as described in [10].

The heterodyne-BPF also plays a crucial role when we take noise into consideration. Since the spectrum of the wideband FM has an infinite number of spectral components, the passband for the heterodyne-BPF is required to cover the scaled spectrum as much as possible in order to reduce the distortion caused by loss of spectrum when the noise is inconsiderable. In the presense of observable noise, however, the noise introduced by covering wider ranges may significantly corrupt the IF estimate. Due to this fact, the passband width of the heterodyne-BPF needs to be optimized around the Carson bandwidth to appropriately trade-off harmonic distortion and noise related distortion [11]. Therefore, the noise free and noisy environments need to be treated separately in terms of the passband range for the heterodyne-BPF.

In addition, a binomial smoothing module is incorporated into the proposed framework as shown in Fig. 2(a), to further reduce the effects of noise. Even though the FM signal is wideband, the IF waveform itself is not necessarily wideband in nature. In many cases, the wideband FM is primarily generated by a large modulation index while the IF still remains in the narrowband range. Under this assumption, by applying the binomial smoothing we can efficiently filter out the high frequency noise in the corrupted IF estimation. When the SNR is high, the improvement becomes extremely evident as we shall see later. Usually we would expect a gain between 5dB and 10dB in the scenario of relatively high SNR.

3 Adaptive Linear Predictive IF Tracking

According to the Wiener-Hopf equations [12], the optimal coeffcients of a linear predictor are given by

$$\mathbf{w}^{opt} = \mathbf{R}_{xx}^{-1}\mathbf{r}_{dx}, \tag{16}$$

where $\mathbf{w}^{opt}$ denotes the optimal tap weight vector, $\mathbf{R}_{xx}$ denotes the input correlation matrix and $\mathbf{r}_{dx}$ denotes the cross-correlation between input vector and desired signal. As summarized in [4], the prediction error filter is given by:

$$E(z) = 1 - \sum_{l=1}^{L} g_l^{opt} z^{-l}, \tag{17}$$

where $\left\{g_l^{opt}\right\}_{l=1}^{L}$ are the coefficients of the corresponding optimal predictor. Conditioned on the small prediction error assumption and the further assumption that the message signal remains essentially invariant over the sampling range of the linear prediction filter, we end up with the approximation in [13] through (17) given by

$$\sum_{l=1}^{L} g_l(k) \exp\{-jl[w_c + m(k)]\} \simeq 1, \tag{18}$$

where $g_l(k)$ is the weight coefficient of tap l at time index k and $m(k)$ is the sample of the message signal at time index k. Then the IF of the signal of interest can be estimated by executing the following steps: 1) Compute the coefficients of the prediction error filter; 2) Obtain the roots of the coefficient polynomial; 3) Calculate the phase argument of the complex conjugate pole location of the corresponding roots.

In the previous work of Asilomar 2005, adaptive algorithms such as AS-LMS and AF-RLS have been incorporated into the structure of linear predictor and compared with each other based on the demodulation error. However, for both algorithms the step-size or the forgetting factor need to be truncated to remain within certain range . In this paper, we choose the generalized normalized gradient descent (GNGD) described in [14] for coefficients update of the linear predictor, which avoids truncation of the adaptively adjusted step-size. The algorithm for this GNGD linear predictive filter is summrized via

$$e(k) = x(k+1) - \sum_{l=1}^{L} g_l(k)\, x(k-L+1) \tag{19}$$

$$\beta(k) = \beta(k-1) - \rho\alpha \frac{e(k)e(k-1)\mathbf{x}^T(k)\mathbf{x}(k-1)}{(\|\mathbf{x}(k)\|_2^2 + \beta(k))^2} \tag{20}$$

$$\eta(k) = \frac{\alpha}{\|\mathbf{x}(k)\|_2^2 + \beta(k)} \tag{21}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \eta(k)e(k)\mathbf{x}(k), \tag{22}$$

where $\mathbf{x}(k)$ and $\mathbf{g}(k)$ denote the vectors of input and tap weights at time index k respectively, α is the step-size parameter and ρ is the offset learning rate parameter. The merit of the GNGD algorithm lies in that the adaptation of its learning rate provides compensation for the assumptions in the derivation of NLMS. On account of its robustness and stability, the GNGD is well-suited for narrowband nonstationary signal environments.

4 Simulation Results

In this section, we present simulation results of the proposed approach under both noise free and noisy environments as well as corresponding results of the prior approach for comparison. Note that the performance is judged by the *normalized RMS IF demodulation error* (NRMSE) throughout this section.

We first look at the case of a wideband sinusoidal FM signal that has a modulation index of 10 and the CR/IB of 20. Under a noise free environment, the performance of the previous MFT framework is illustrated by Fig. 3(a). Note that the performance associated with $R = 1$, i.e, the origin of the plot corresponds to GNGD demodulation without MFT, while ω_d is the normalized upshift radian frequency translation in the range of $[0, \pi]$. By applying a large conversion factor of $R = 128$, a reduction of around 40 dB in the demodulation error over GNGD demodulation is attained. The result of Fig. 3(a) confirms the claim that a large conversion factor strengthens the effect achieved by frequency compression thus leading to significant reduction in the demodulation error. Moreover, it reflects the fact that the use of a larger factor requires a very high order FIR bandpass filter with a satisfactory frequency response. For example, $R = 128$, demands the order of FIR bandpass filter to be as high as 4096, which results in unrealistic parameters for the narrow passband. This constraint seriously limits the implementation in a practical system for large factors.

To relax the constraint, an alternative method is to apply the proposed MFT framework, whose performance is illustrated by 3(b). In comparison to Fig. 3(a), it can be observed that the required order for the FIR bandpass filter is effectively reduced at the cost of sacrificing a small amount of improvement in demodulation error. For instance, when $R = 128$, the order for the FIR bandpass filter drops significantly from 4096 to 512 with just 4dB loss in performance, suggesting no observable difference in performance between both MFT frameworks except for the dramatic reduction of the order for the heterodyne-BPF. Note that the frequency response of the heterodyne-BPF in the proposed MFT framework for the case $R = 128$ is illustrated by Fig. 4(b), which has much wider passband compared to the frequency response of the BPF in the previous MFT framework shown in Fig. 4(a) and thus practical for implementation.

For noisy environments, the performance of the previous MFT framework is illustrated by Fig. 5(a) for the same sinusoidal FM signal corrupted by additive white Gaussian noise of different SNR. Based on the observation, improvement over GNGD demodulation without MFT varies according to the different SNRs. For example, when the SNR is 20dB, the improvement is only aournd 6dB, and when the SNR increases to 40dB, improvement increases to around 20dB. In addtion, Fig. 5(b) summarizes the results of the proposed MFT framework in the presense of noise, indicating better performance due to the binomial smoothing. For the case of 20dB SNR, the improvement increases to 15dB compared with prior system. Note that the NRMSE gradually becomes saturated as R increases, due to LTI filtering induced harmonic distortion of the FM signals[2]. In contrast, the proposed MFT approach results in a better performance, and further provides a more practical approach to the BPF design.

We further investigate an extreme wideband scenario under the noise free environment, where the modulation index β is as large as 50 and the frequency deviation is equal to the carrier frequency with the IF varying over the entire carrier range as illustrated by Fig. 6(a). For the signal of interest, the IF estimates of both the prior and proposed MFT frameworks are illustrated by Fig. 6(b) and Fig. 6(c) respectively. It can be observed

[2] Since they are only approximate eigenfunctions as described in [15].

that the GNGD demodulation alone fails in this extreme wideband scenario, while both the prior and proposed frameworks that exploit MFT maintain tracking. The observation implies that both MFT frameworks guarantee the demodulation with acceptable performance even in the worst senarios where conventional algorithms would normally fail.

To quantify the performance of the porposed MFT approach, we explore another scenario where the signal is a wideband linear chirp instead of a sinusoidal FM. The short-time spectrum of the chirp signal is illustrated by Fig. 7(a). To validate the performance of the proposed MFT approach , we can compare the variance of error with respect to the chirp rate estimate with its *Cramér Rao lower bound*(CRLB). The chirp rate can be obtained via demodulated IF followed by a least square estimator. In the presence of noise with different SNRs, the result is summarized in Fig. 7(b). Improvement over the GNGD alone is more apprent with respect to lower SNR, indicating satisfactory performance of proposed MFT approach in noisy condition. Also note that gap between the error variance estimate of the proposed MFT and the corresponding CRLB is nearly a constant, which can be explained directly via the loss of spectrum incurred by filtering the FM signal. The proposed approach is applied to the problem of demodulating the first formant of a speech phoneme which tends to be a wideband formant.

5 Applications to the demodulation of wideband speech formants

Formant analysis is at the heart of various speech applications including speech synthesis, encoding and transmission. Formants that are natural resonant frequencies determined by the vocal tract have been successfully modeled as damped sinusoidal signals according to [16], [17]. Traditionally, by approximating the speech signal as the superposition of corresponding damped sinusoidal signals, each formant is extracted out from the speech signal via bandpass filtering with center frequencies and bandwidths determined by the *linear predictive coding* (LPC) estimates. However, specific formants of many vowels and semi-vowels are wideband in nature. Their corresponding spectrums spread out with significant energy in the sidelobes. In particular, the first formant usually has a CR/IB ratio less than 10 due to its relatively small center frequency in comparison to tis bandwidth. For example, the first formant of the phoneme /iy/ has a bandwidth estimate of 54Hz and its formant frequency centered around 226Hz as illustrated in [18], which results in a CR/IB ratio between 4 to 5 that would normally be classified as wideband. Since LPC itself relies upon the narrowband assumption, bandpass filtering using parameters from LPC estimates is not appropriate for the extraction of such wideband formants.

For multicomponent wideband signals, separation techniques such as the multiband filtering, empirical mode decomposition (EMD), periodicity based algebaraic separation (PASED) [19] may be employed to separate out the components prior to applying the wideband demodulation approach.

In this report, we employ the *empirical mode decomposition* (EMD) [20] as the approach to extract the first formant instead of bandpass filtering in order to perform the first formant analysis. Because each formant is modeled as a damped sinusoidal signal, the proposed approach via GNGD cannot be applied directly to the formant since it only deals with pure FM signals. Therefore, the Hilbert transform is first applied to the first formant extracted via the EMD to obtain its instantaneous amplitude estimate. Then we normalize the formant using the instantaneous amplitude calculated by the Hilbert transform to obtain a pure FM equivalent. The IF of the first formant can be eventually demodulated from this pure FM equivalent by GNGD with a large conversion factor as we propose.

Here we consider the example of a voiced phoneme /ae/ for the first formant analysis, whose waveform and short-time spectrum are depicted by Fig. 8(a) and Fig. 8(b) respectively. Fig. 9(a) illustrates the short-time spectrum for the first formant extracted by bandpss filtering using parameters from LPC estimates, while Fig. 9(b) shows the short-time spectrum for the first formant extracted by summming up corresponding IMF's calculated via the EMD. By comparing these spectrums, we can observe that the output from EMD matches the spectrum of the phoneme /ae/ better than the output from bandpass filtering. The EMD is able to retain the sidelobes that characterize the wideband nature of the first formant while the bandpass filter merely cuts off the sidelobes outside its passband and makes the extracted formant narrowband due to the nature of filtering. The demodulation

results based on the different methods are summarized in Fig. 10. Line 1 represents the IF estimate of the first formant extracted via bandpass filtering and demodulated by the Hilbert transform, while Line 2 represents the IF estimate via the EMD and demodulated directly by the Hilbert transform. As we can see, the IF demodulated via the EMD (Line 2) has higher peaks and deviates more frequently from its center frequency compared to the IF demodulated via filtering, indicating a larger bandwidth of the first formant. However, the Hilbert transform also assumes the modulated signal to be narrowband and suffers from the large error caused by wideband formants as in this case. Line 3 and Line 4 depicts the IF estimates using GNGD with a large conversion factor of 64 based on the previous and proposed MFT frameworks respectively. As analyzed in previous sections, the MFT approach converts the wideband signal into a narrowband signal and hence, improves the performances of most demodulation algorithms under the narrowband assumption. In this example, the IF estimates incorporating the MFT approach are actually smoother and more reliable in the region where the corresponding IF estimate from the direct Hilbert tranform incurs a large error and even becomes negative [3] as Line 2 shows. Large spikes and negative IF's at the pitch boundaries are a direct result of violation of narrowband assumptions at these instants made by the underlying demodulation algorithm. Besides, the proposed MFT framework obtains a gain over the previous MFT framework in terms of implementation of the FIR bandpass filter with order reduced from 3072 to 768, meanwhile exhibiting very close demodulation performances as shown by Line 3 and Line 4 that almost overlap with each other.

6 Conclusion

A system that combines multirate processing and heterodyning to accomplish wideband FM demodulation for large wideband to narrowband factors was proposed. Prior work combining these systems was shown to produce impractical designs for large factors, needing bandpass filters of very high orders and very narrow passbands. Interchanging the order of the heterodyne and multirate modules was shown to reduce the computational burden placed on the bandpass filter for for large conversion factors. Application of the proposed wideband demodulation approach to speech first formant demodulation, using EMD extracted formants, was shown to result in larger IF excursions than those produced by bandpass filtering and Hilbert transform demodulation,without the IF becoming negative.

References

[1] P. Rathore and R. Pachori, "Instantaneous fundamental frequency estimation of speech signals using desa in low-frequency region," in *Signal Processing and Communication (ICSC), 2013 International Conference on*, pp. 470–473, Dec 2013.

[2] M. Fitch, B. Exact, I. Boyd, B. Exact, K. Briggs, B. Exact, and F. Stentiford, "Gaussian multi-level fm for high-bandwidth satellite communications," *BT Exact*, 2004.

[3] W. Zongbo, G. Meiguo, L. Yunjie, and J. Haiqing, "Design and application of drfm system based on digital channelized receiver," in *Radar, 2008 International Conference on*, pp. 375–378, 2008.

[4] M. Gupta and B. Santhanam, "Adaptive linear predictive frequency tracking and cpm demodulation," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1, pp. 202–206, 2003.

[5] H. Kwon and K.-B. Lee, "A novel digital fm receiver for mobile and personal communications," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1466–1476, 1996.

[3] A negative IF for a speech formant is not physically meaningful

[6] R. G. Wiley, H. Schwarzlander, and D. Weiner, "Demodulation procedure for very wide-band fm," *IEEE Trans. Commun.*, vol. 25, no. 3, pp. 318–327, 1977.

[7] A. Noga and T. Sarkar, "A discrete-time method of demodulating large deviation fm signals," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1194–1200, Aug 1999.

[8] A. V. Oppenheim and R. W. Schafer, "Discrete-time signal processing," *Prentice Hall, New York*, 1999.

[9] B. Santhanam, "Generalized energy demodulation for large frequency deviations and wideband signals," *IEEE Signal Process. Lett.*, vol. 11, no. 3, pp. 341–344, 2004.

[10] F. F. Kuo and J. F. Kaiser, *System analysis by digital computer*. Wiley, 1966.

[11] A. Bovik, P. Maragos, and T. Quatieri, "Am-fm energy detection and separation in noise using multiband energy operators," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3245–3265, 1993.

[12] S. S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2005.

[13] L. J. Griffiths, "Rapid measurement of digital instantaneous frequency," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 23, no. 2, pp. 207–222, 1975.

[14] D. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 115–118, 2004.

[15] B. Santhanam, "Orthogonal modes of frequency modulation and the sturm-liouville frequency modulation model," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3486–3495, July 2012.

[16] P. Maragos, T. Quatieri, and J. Kaiser, "Speech nonlinearities, modulations, and energy operators," in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, vol. 1, pp. 421–424 , Apr 1991.

[17] A. Potamianos and P. Maragos, "Speech analysis and synthesis using an am–fm modulation model," *Speech Communication*, vol. 28, no. 3, pp. 195–209, 1999.

[18] L. Rabiner and R. Schafer, *Theory and Applications of Digital Speech Processing*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

[19] B. Santhanam and P. Maragos, "Multicomponent am-fm demodulation via periodicity-based algebraic separation and energy-based demodulation," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 473–490, Mar 2000.

[20] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

List of Figures

1 Block diagram of the previous MFT framework in [9]. The wideband signal is first sampled above the Nyquist rate, interpolated by a factor R and then heterodyned via multiplying $cos(\omega_d n)$, followed by a discrete FIR bandpass filter with a scaling module based on (10) to achieve MFT. Then it goes through a demodulation block to obtain the IF estimation of the compressed heterodyned signal. To reconstruct IF estimation of the original signal, the compressed heterodyned IF is then shifted back by subtracting ω_d, decimated by R and scaled back approapriately according to (8), followed by the DAC module. 11

2 (a) Block diagram of the proposed MFT framework for large wideband to narrowband conversion factors. The interpolation module of the prior framework is separated into two with one in fornt of the heterodyning module and the other one right after. The first interpolation module has a relatively small upsampling rate of R_1 which is appropriately chosen such that the discrete BPF can be implemented within the range of half the sampling rate after heterodyning the signal with a frequency translation of ω_d. The relatively small R_1 would result in a wider passband for the discrete bandpass filter, thus reducing the its design of complexity. (b) Multistage implementation for interpolation modules in 2(a) based on Noble identity. Note that multistage implementation for the corresponding decimation modules can be realized in a similar fashion. Multistage structures relax the constraints of narrow passbands for the lowpass filters with large multirate factors by decomposing the one-stage lowpass filters into multistages each with a much wider passband range. 12

3 Comparison between performances of both MFT frameworks under noise free environments. (a) Performance of the previous MFT framework with R specifying the multirate factor and L specifying the order of BPF. (b) Performance of the proposed MFT framework with R specifying the multirate factor and L specifying the order of BPF. Both MFT frameworks reduce the demodulation error significantly with large conversion factors. 13

4 Comparison between the previous and proposed MFT frameworks in terms of the requirements on the heterodyne-BPF with a conversion factor $R = 128$. (a) Frequency response of BPF in the previous MFT framework with order $L = 4096$. (b) Frequency response of BPF in the previous MFT framework with order $L = 512$. Note that the proposed framework succesfully reduced the complexity of the design of heterodyne-BPF. 14

5 Comparison between performances of both MFT frameworks in environments corrupted by AGWN. (a) Performance of the previous MFT framework with R specifying the multirate conversion factor. (b) Performance of the proposed MFT framework with R specifying the multirate conversion factor. 15

6 Demodulation performances of both MFT frameworks with conversion factor $R = 128$ and normalized radian frequency shift $w_d = 0.1\pi$ under the extreme senario. (a) Wideband sinusoidal FM signal with modulation index $\beta = 50$ and CR/IB $= 50$. (b) IF estimates (dashed line) of GNGD under the previous MFT framework with the heterodyne-BPF order $L = 4096$, (dasheddotted line) the GNGD, and (solid line) actual IF. (c) IF estimates (dashed line) of GNGD under the proposed MFT framework with the heterodyne-BPF order $L = 512$, (dasheddotted line) the GNGD, and (solid line) actual IF. Note that the GNGD alone fails under this large deviation FM signal with a modulation index as large as 50 while the GNGD combined with both MFT frameworks maintain tracking with similar performances. 16

7 Performance of the proposed MFT framework with a multirate conversion factor of 128 in wideband linear chirp scenario. (a) Short-time spectrum of the wideband linear chirp. (b) Comparison of the error variance with respect to the estimate of chirp rate. Note that there is only a small gap less than 20dB between the error variance estimate of the proposed MFT and the corresponding *Cramér Rao lower bound* (CRLB). 17

8 Speech waveform and short-time spectrum of the phoneme /ae/. (a) Time domain waveform of the phoneme /ae/ sampled at 16kHz. (b) Short-time spectrum of the phoneme /ae/. Note that the first formant of phoneme /ae/ has a widespread spectrum around its formant frequency with a relatively large CR/IB ratio, which indicates its wideband nature. 18

9 Short-time spectrums of the first formant extracted via filtering and EMD respectively. (a) Short-time spectrum of the first formant extracted by narrowband filtering. (b) Short-time spectrum for the first formant extracted by the EMD. Note that the spectrum of the first formant extracted by filtering results in a narrowband spectrum since harmonic components outside the passband are cut off due to bandpass filtering, while the first formant extracted via EMD has a widespread spectrum that preserves the harmonic components revealing the characteristics of wideband formants. 19

10 IF estimates of the first formant (Line 1) extracted by narrowband filtering via the Hilbert transform, (Line 2) extracted by the EMD via the Hilbert transform, (Line 3) extracted by the EMD via multirate GNGD based on the previous MFT framework with conversion factor $R = 64$, normalized radian frequency shift $w_d = 0.2\pi$ and a order of 3072 for heterodyne-BPF, (Line 4) extracted by the EMD via multirate GNGD based on the proposed MFT framework with conversion factor $R = 64$, normalized radian frequency shift $w_d = 0.2\pi$ and a order of 768 for equation-BPF. Note that Line 3 and Line 4 almost overlap with each other, indicating very close demodulation performances of both MFT frameworks without the IF becoming negative as of the demodulation by direct Hilbert transform shown by Line 2. 20

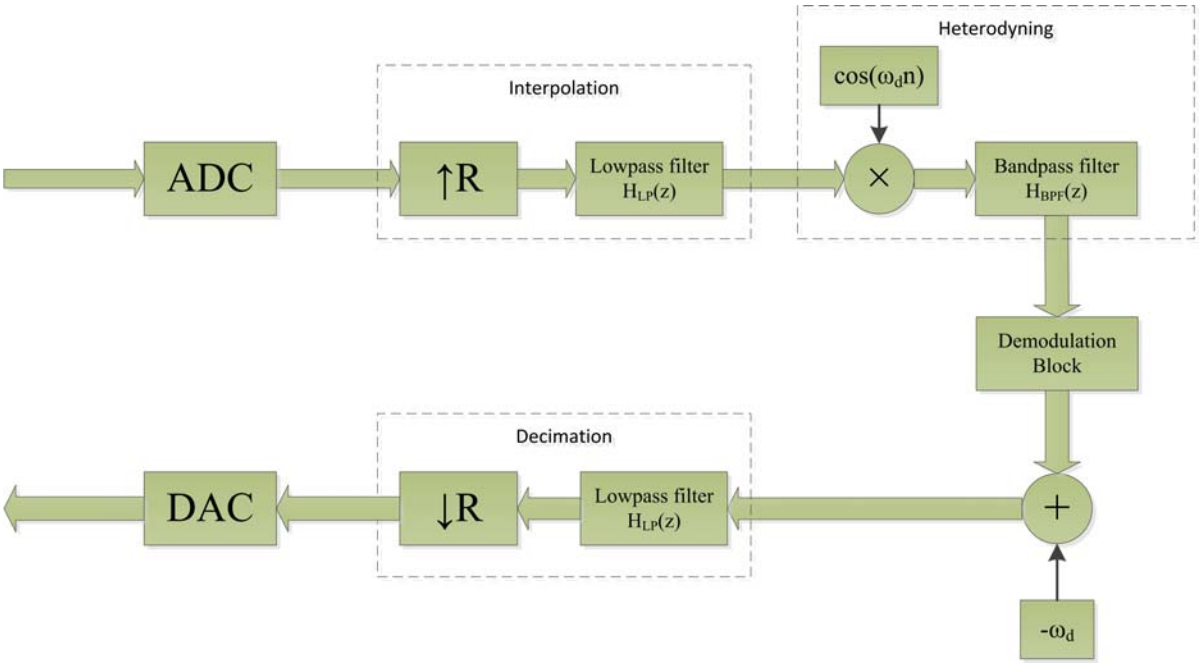
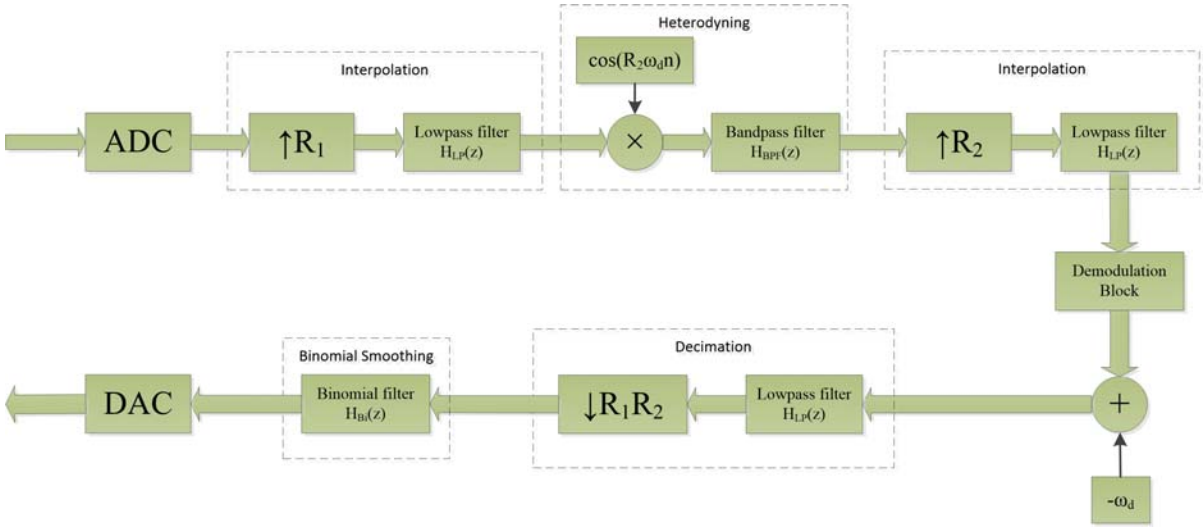
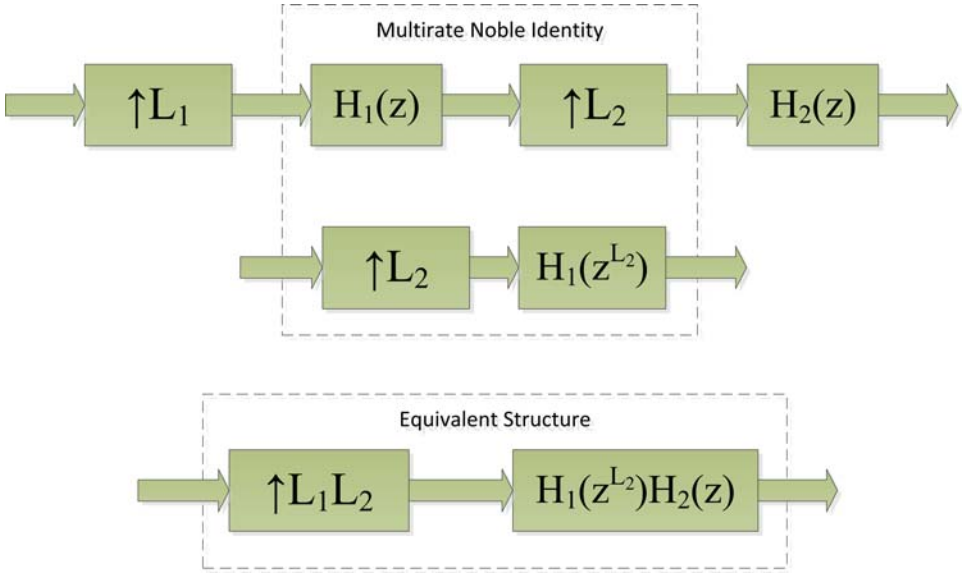


Figure 1: Block diagram of the previous MFT framework in [9]. The wideband signal is first sampled above the Nyquist rate, interpolated by a factor R and then heterodyned via multiplying $cos(\omega_d n)$, followed by a discrete FIR bandpass filter with a scaling module based on (10) to achieve MFT. Then it goes through a demodulation block to obtain the IF estimation of the compressed heterodyned signal. To reconstruct IF estimation of the original signal, the compressed heterodyned IF is then shifted back by subtracting ω_d, decimated by R and scaled back approapriately according to (8), followed by the DAC module.

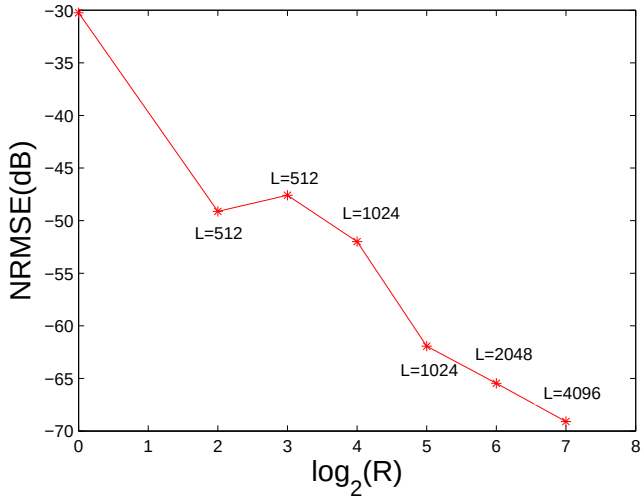


(a)

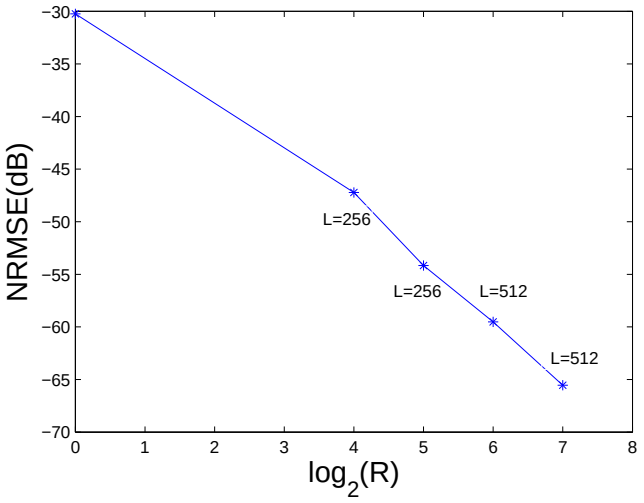


(b)

Figure 2: (a) Block diagram of the proposed MFT framework for large wideband to narrowband conversion factors. The interpolation module of the prior framework is separated into two with one in fornt of the heterodyning module and the other one right after. The first interpolation module has a relatively small upsampling rate of R_1 which is appropriately chosen such that the discrete BPF can be implemented within the range of half the sampling rate after heterodyning the signal with a frequency translation of ω_d. The relatively small R_1 would result in a wider passband for the discrete bandpass filter, thus reducing the its design of complexity. (b) Multistage implementation for interpolation modules in 2(a) based on Noble identity. Note that multistage implementation for the corresponding decimation modules can be realized in a similar fashion. Multistage structures relax the constraints of narrow passbands for the lowpass filters with large multirate factors by decomposing the one-stage lowpass filters into multistages each with a much wider passband range.

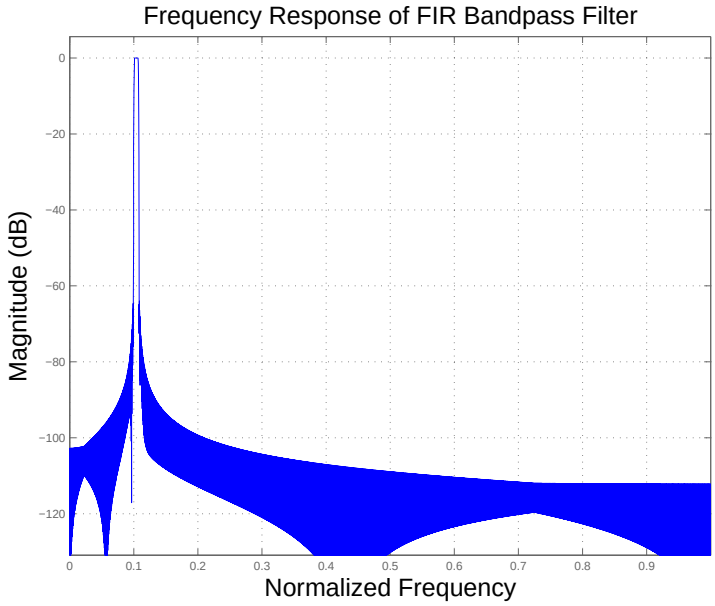


(a)

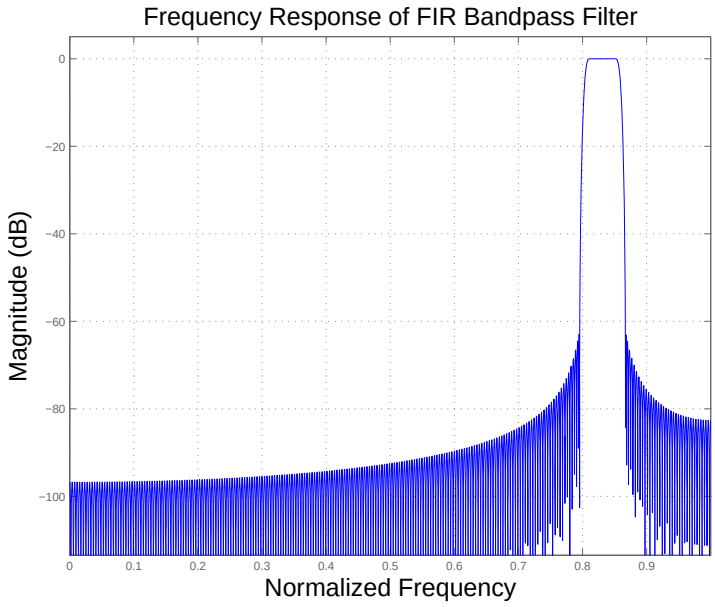


(b)

Figure 3: Comparison between performances of both MFT frameworks under noise free environments. (a) Performance of the previous MFT framework with R specifying the multirate factor and L specifying the order of BPF. (b) Performance of the proposed MFT framework with R specifying the multirate factor and L specifying the order of BPF. Both MFT frameworks reduce the demodulation error significantly with large conversion factors.

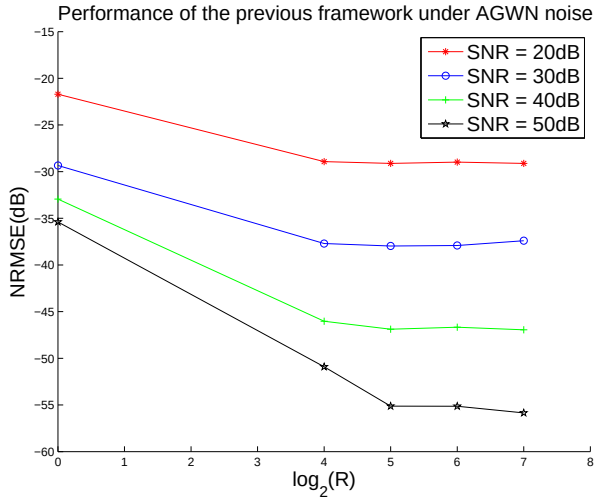


(a)

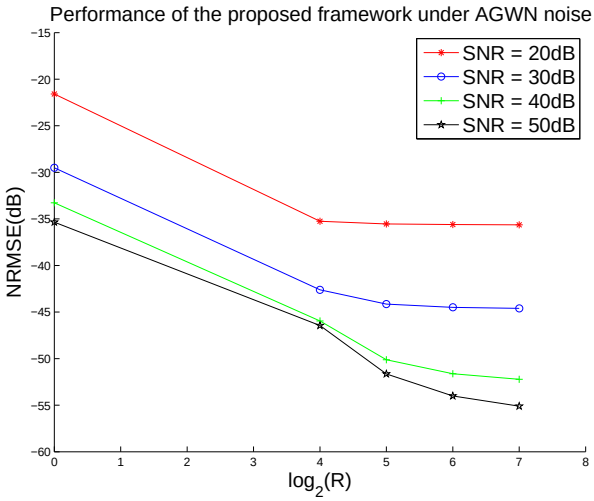


(b)

Figure 4: Comparison between the previous and proposed MFT frameworks in terms of the requirements on the heterodyne-BPF with a conversion factor $R = 128$. (a) Frequency response of BPF in the previous MFT framework with order $L = 4096$. (b) Frequency response of BPF in the previous MFT framework with order $L = 512$. Note that the proposed framework succesfully reduced the complexity of the design of heterodyne-BPF.

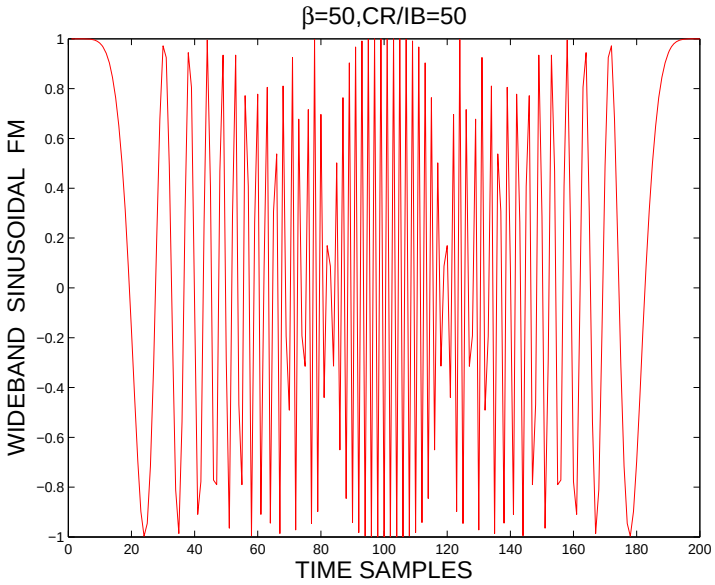


(a)

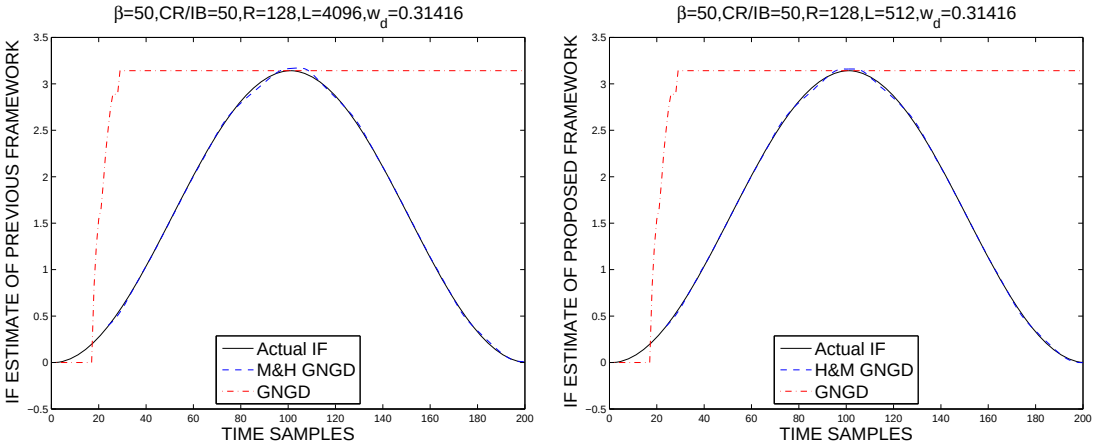


(b)

Figure 5: Comparison between performances of both MFT frameworks in environments corrupted by AGWN. (a) Performance of the previous MFT framework with R specifying the multirate conversion factor. (b) Performance of the proposed MFT framework with R specifying the multirate conversion factor.

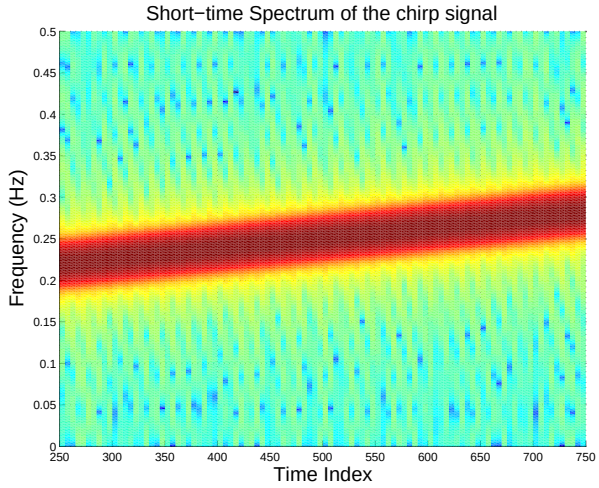


(a)

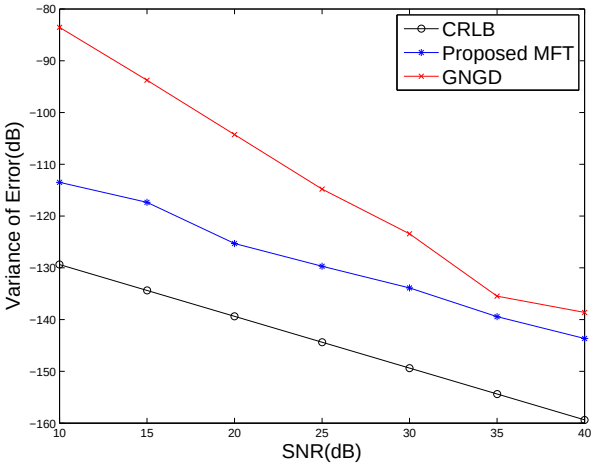


(b) (c)

Figure 6: Demodulation performances of both MFT frameworks with conversion factor $R = 128$ and normalized radian frequency shift $w_d = 0.1\pi$ under the extreme senario. (a) Wideband sinusoidal FM signal with modulation index $\beta = 50$ and CR/IB $= 50$. (b) IF estimates (dashed line) of GNGD under the previous MFT framework with the heterodyne-BPF order $L = 4096$, (dasheddotted line) the GNGD, and (solid line) actual IF. (c) IF estimates (dashed line) of GNGD under the proposed MFT framework with the heterodyne-BPF order $L = 512$, (dashed-dotted line) the GNGD, and (solid line) actual IF. Note that the GNGD alone fails under this large deviation FM signal with a modulation index as large as 50 while the GNGD combined with both MFT frameworks maintain tracking with similar performances.

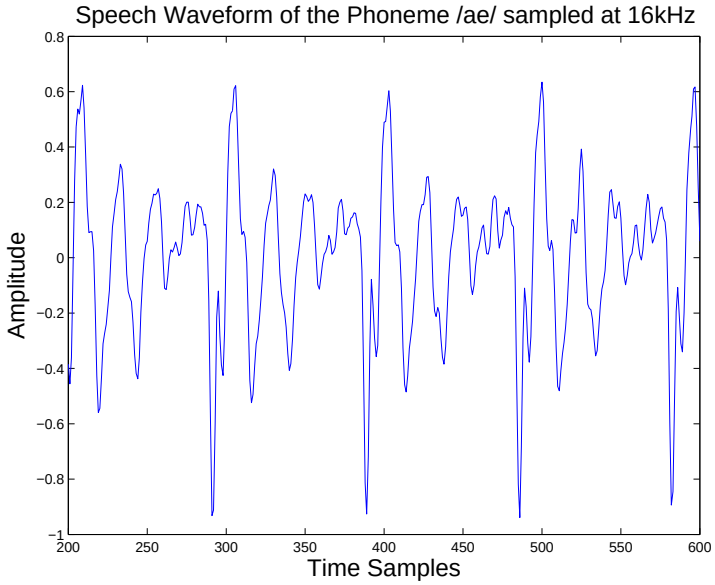


(a)

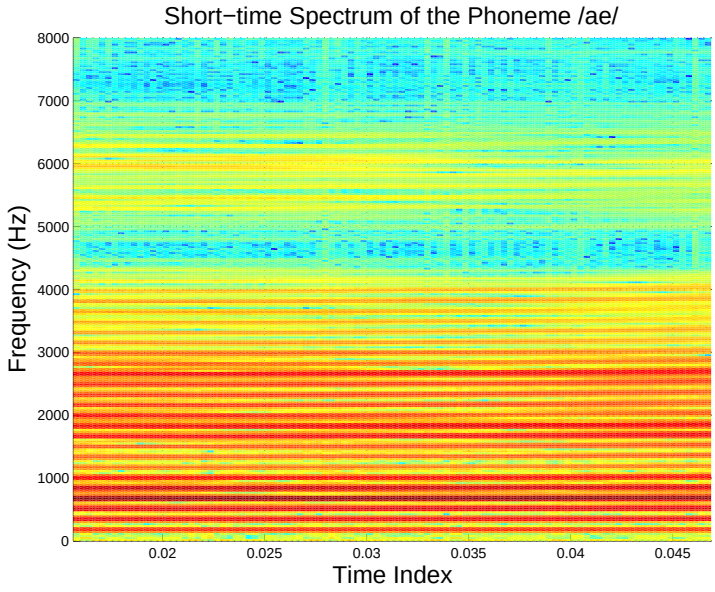


(b)

Figure 7: Performance of the proposed MFT framework with a multirate conversion factor of 128 in wideband linear chirp scenario. (a) Short-time spectrum of the wideband linear chirp. (b) Comparison of the error variance with respect to the estimate of chirp rate. Note that there is only a small gap less than 20dB between the error variance estimate of the proposed MFT and the corresponding *Cramér Rao lower bound* (CRLB).

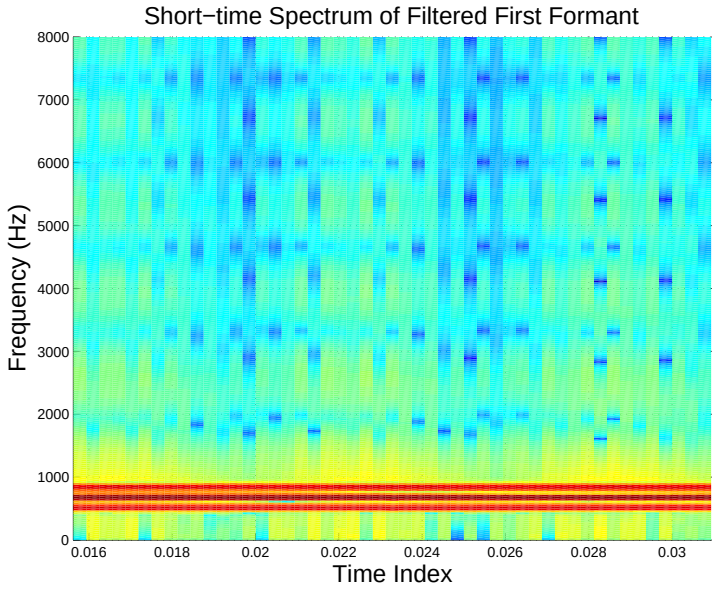


(a)

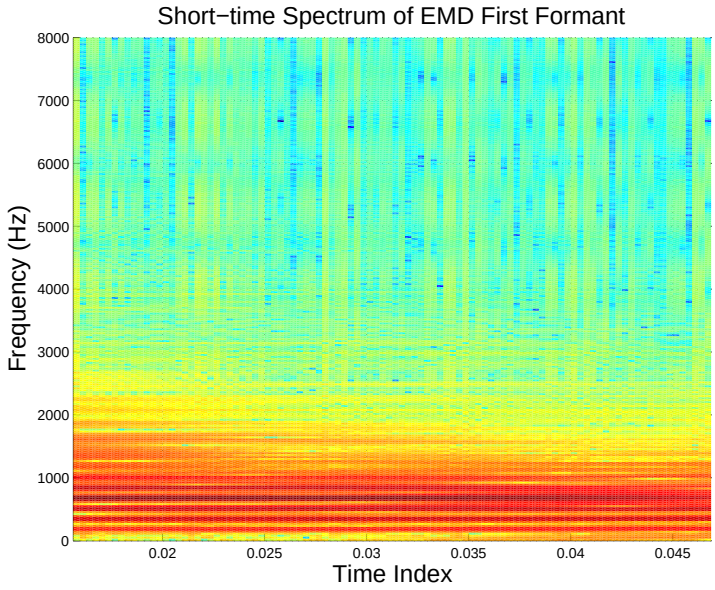


(b)

Figure 8: Speech waveform and short-time spectrum of the phoneme /ae/. (a) Time domain waveform of the phoneme /ae/ sampled at 16kHz. (b) Short-time spectrum of the phoneme /ae/. Note that the first formant of phoneme /ae/ has a widespread spectrum around its formant frequency with a relatively large CR/IB ratio, which indicates its wideband nature.



(a)



(b)

Figure 9: Short-time spectrums of the first formant extracted via filtering and EMD respectively. (a) Short-time spectrum of the first formant extracted by narrowband filtering. (b) Short-time spectrum for the first formant extracted by the EMD. Note that the spectrum of the first formant extracted by filtering results in a narrowband spectrum since harmonic components outside the passband are cut off due to bandpass filtering, while the first formant extracted via EMD has a widespread spectrum that preserves the harmonic components revealing the characteristics of wideband formants.

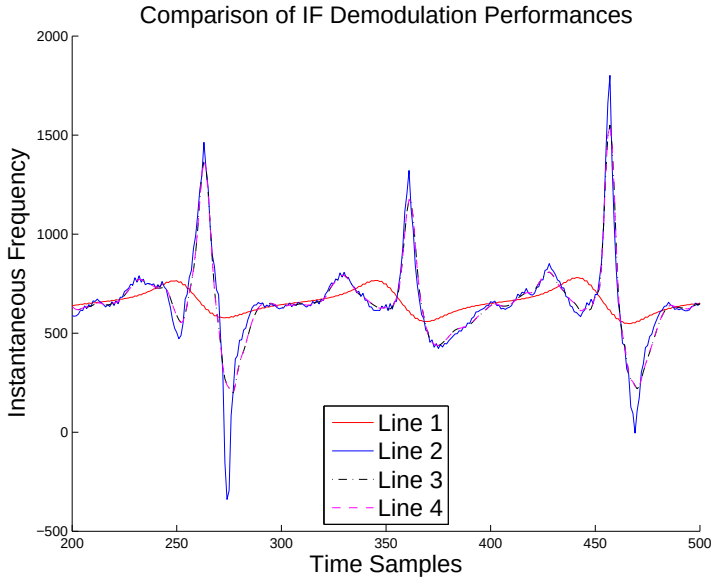


Figure 10: IF estimates of the first formant (Line 1) extracted by narrowband filtering via the Hilbert transform, (Line 2) extracted by the EMD via the Hilbert transform, (Line 3) extracted by the EMD via multirate GNGD based on the previous MFT framework with conversion factor $R = 64$, normalized radian frequency shift $w_d = 0.2\pi$ and a order of 3072 for heterodyne-BPF, (Line 4) extracted by the EMD via multirate GNGD based on the proposed MFT framework with conversion factor $R = 64$, normalized radian frequency shift $w_d = 0.2\pi$ and a order of 768 for equation-BPF. Note that Line 3 and Line 4 almost overlap with each other, indicating very close demodulation performances of both MFT frameworks without the IF becoming negative as of the demodulation by direct Hilbert transform shown by Line 2.