

# A novel approach for artificial bandwidth extension of speech signals by LPC technique over proposed GSM FR NB coder using high band feature extraction and various extension of excitation methods

Ninad Bhatt · Yogeshwar Kosta

Received: 17 October 2013 / Accepted: 15 August 2014 / Published online: 24 August 2014  
© Springer Science+Business Media New York 2014

**Abstract** This research addresses an issue of wide band (WB) speech transmission (having cut-off frequency  $f_c = 8$  kHz) over standard narrow band (NB) communication link (supporting bandwidth of 300–3,400 Hz). A long transition time for technological up-gradation from NB to WB systems eventually lead to development of backward compatible techniques such as artificial bandwidth extension (ABE) which is capable of providing bandwidth of 50–7,000 Hz, in turn contributing toll quality recovered speech at receiving end. This paper investigates a novel approach to compute high band (HB) features using linear predictive coding (LPC) technique at transmitter from given input WB speech corpus. These encoded features are embedded into bitstream of proposed GSM Full Rate 06.10 NB speech coder using joint source coding and data hiding technique and then transmitted to receiver. At receiver, these HB features are extracted to reproduce HB recovered speech using watermark extraction algorithm and for the same different extension of excitation techniques have been adopted and implemented. An e-test bench is created to implement this proposed ABE coder in MATLAB and series of simulations are carried out using Subjective (mean opinion score—MOS) and Objective (perceptual evaluation of speech quality—PESQ) analysis. Obtained results for both analyses advocate performance improvement of proposed ABE coder over legacy GSM 06.10 FR NB coder for various extension of excitation techniques.

**Keywords** Proposed GSM Full Rate coder · Artificial bandwidth extension · Subjective analysis · Objective analysis · Linear predictive coding

## 1 Introduction

Today's wired and wireless communication networks suffer from limited NB acoustic bandwidth of 300–3,400 Hz. This inherent limitation reduces overall intelligibility and naturalness of recovered speech at receiving end. Effective and efficient techniques have been devised in last decade by implementing WB systems to overcome the limitations stated above. Still, bottleneck problem in the utilization of WB systems is the necessity of up-gradation of current NB systems to its counterpart WB system, in which, hardware and software up-gradation and backward compatibility is always an issue. Having known convincing fact about better perceived speech quality performance offered by WB coders, however, sudden replacement of entire NB coding and transmission systems is not feasible because of tremendous infrastructure expenses incurred to operators and also is a case with customers. Current speech transmission system is a mixture of traditional NB terminals and new WB terminals. The long transitional period between NB and WB system demands to enhance speech quality without significant modification in already existing network infrastructure which leads to development of artificial bandwidth extension (ABE) [Jax and Vary \(2006, 2003\)](#). Instead of transition to WB system, alternatively estimation of the missing high frequency components of the speech at receiving terminal from available NB speech is called as ABE. However, performance of stand-alone ABE, that artificially generates WB speech, is inherently limited ([Jax and Vary 2002](#); [Bhatt et al. 2012](#)). Further, in order to improve reconstructed speech

---

N. Bhatt (✉)  
EC Department, CKPCET, Surat, Gujarat, India  
e-mail: bhattninad@gmail.com

Y. Kosta  
Marwadi Education Foundation, Rajkot, Gujarat, India  
e-mail: ypkosta@gmail.com

quality while reducing the overall computation load, ABE with side information is surely one of the potential candidates of consideration. In a past decade, few data hiding techniques were devised to embed and hide HB feature bits into NB encoded bitstream before transmission and reverse operation was carried out at receiver for synthetic WB speech recovery.

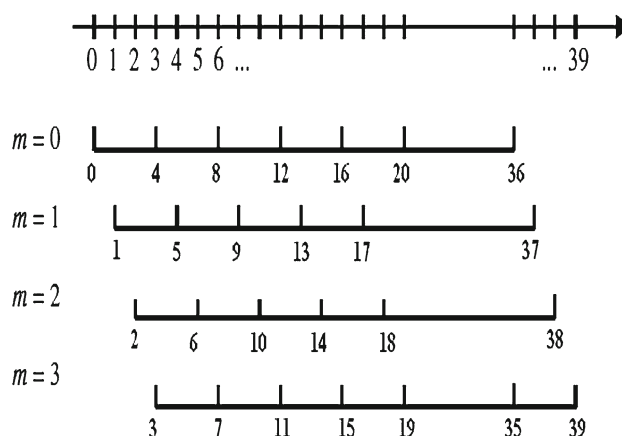
In this work, given a WB input speech, a novel approach is proposed to embed and transmit HB features (like envelop and gain information) as a stego signal over a carrier signal of proposed GSM FR encoded bitstream [Bhatt and Kosta \(2011\)](#). Watermark embedding algorithm embeds HB feature bitstream into encoded GSM FR coder bitstream at location of Class I<sub>b</sub> of GSM 05.03 channel coding standards as per [ETSI \(1999\)](#). Further, from a received NB bitstream, NB decoded speech can be reproduced from legacy GSM FR NB decoder at the same time watermark extraction section delivers embedded HB features. These HB features are then utilized for HB speech recovery and to perform the same, various extension of excitation methods like spectral translation (ST), spectral folding (SF), noise modulation (NM), non linear transformation (NLT), full wave rectification (FWR) etc. have been implemented in this research. Finally, both NB and HB versions of recovered speech have been summed up using over lap add (OLA) method to effectively reproduce the WB speech at receiving end.

Paper is organized in the following way: proposed modifications in GSM FR coder are explained in Sect. 2. Section 3 describes proposed ABE coder with HB feature extraction as side information using LPC technique. Section 4 touches upon overall performance analysis of proposed ABE coder for all different extension of excitation methods. Concluding remarks are given in last section.

## 2 Proposed modifications in GSM 06.10 FR NB coder

ETSI GSM 06.10 Full Rate coder consists of three major blocks LPC section, long term Predictive (LTP) section and regular pulse excitation (RPE) section [ETSI \(2005–2006\)](#). The proposed modifications are suggested in RPE section in the selection of grid positions. In RPE section, selection of grid position and samples is modified such a way that no samples repeat in multiple grids which is the case of standard GSM FR coder in first and forth grid where except sample number 0 and sample number 39 don't repeat whereas all other samples in both grids repeat. A new proposed grid selection strategy is as shown in Fig. 1.

As can be highlighted in Fig. 1, if the weighting filtered prediction-error sequence is down-sampled by a ratio of 4 instead of 3, it results into four interleaved sequences with regularly spaced pulses [Bhatt and Kosta \(2011\)](#). These are defined with



**Fig. 1** Sampling grids used in position selection for proposed GSM FR 11.2 kbps coder

**Table 1** Bit allocation for proposed GSM FR speech coder [Bhatt and Kosta \(2011\)](#)

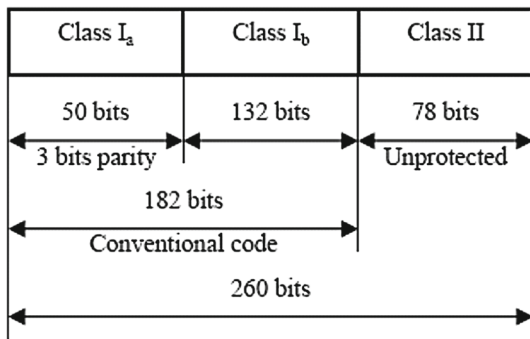
Parameter	No. per frame	Resolution	Total bits/frame
LPC	8	6,6,5,5,4,4,3,3	36
Pitch Period	4	7	28
Long term gain	4	2	8
Grid position	4	2	8
Peak magnitude	4	6	24
Sample amplitude	4*10	3	120
Total			224

$$x_m(k) = x(m + 4k) \quad m = 0, 1, 2, 3; \quad k = 0, 1, 2, \dots, 9; \quad (1)$$

Where, m = no. of grids per sub segment and k = no. of samples per grid.

Main benefit in this sampling grid position selection is that, there is no repetition of any sample in multiple grids whereas now the total number of samples per grid reduces from 13 to 10 so ultimately as an outcome there is a reduction in overall bit-rate of 1.8 kbps (3 samples per grid \* 3bits per sample \* 4 sub frames = 36 bits/20 ms frame) compared to actual bit rates of 13 kbps for GSM 06.10 FR coder which can be useful for steganographic data transmission [Bhatt et al. \(2011\)](#) or for HB feature transmission over NB bitstream in each frame transmission. The proposed modification in GSM FR offers a new bit allocation as shown in Table 1.

The parameters produced by GSM FR encoder like short term filter parameters, long term prediction parameters and regular pulse excitation parameters have their unequal importance with respect to their recovered speech quality. As shown in Fig. 2, the GSM FR encoded 260 bits are rearranged according to its subjective importance as mentioned in GSM 05.03 [ETSI \(1999\)](#).



**Fig. 2** GSM FR frame structure with bit locations according to their subjective importance [Licai and Shuozhong \(2006\)](#)

**Table 2** Quantization and bit allocation of HB parameters

Parameter	No. per frame	Resolution	Total bits/frame
LPC coefficients	08	6,5,5,4,4,3,3,3	33
Gain	01	3	3
Total			36 bits

The rearranged bits are classified into class Ia, Ib and II which contains total no. of 50, 132 and 78 bits respectively. Error protection provided by channel coding in different classes uses different methods like CRC and convolution coding. Class Ia contains highest priority 50 bits so a single bit error in these bits may result into severe degradation in recovered speech quality and hence addition of HB feature data in this class is not feasible. Class II contains the lowest priority 78 bits in which if all or some of the bits are replaced by HB feature data then quality of speech does not suffer much but as there is no error protection added in this class of bits, it is hence not feasible to transmit HB feature data in this class because recovery of that steganographic data at receiver may be suffered. Thus, class Ib which contains 132 bits can be chosen for hiding and transmitting steganographic HB feature data over wireless link [Licai and Shuozhong \(2006\)](#).

Table 3 of [Bhatt et al. \(2011\)](#) demonstrates the proposed modifications in GSM 05.03 (Table 2) [ETSI \(1999\)](#) to pro-

vide room for steganographic data addition. As after proposed modifications in GSM FR, new bit rate is 11.2 kbps (224 bits/20ms frame), classification of bits in different class Ia, Ib and II now contains total no. of 50, 96 and 78 bits respectively. So, ultimately 36 bits (per each frame) from class Ib can be spared and adjusted in order to utilize them for HB feature data transmission with error protection.

### 3 Implementation of proposed ABE coder

#### 3.1 Band splitting and recombining

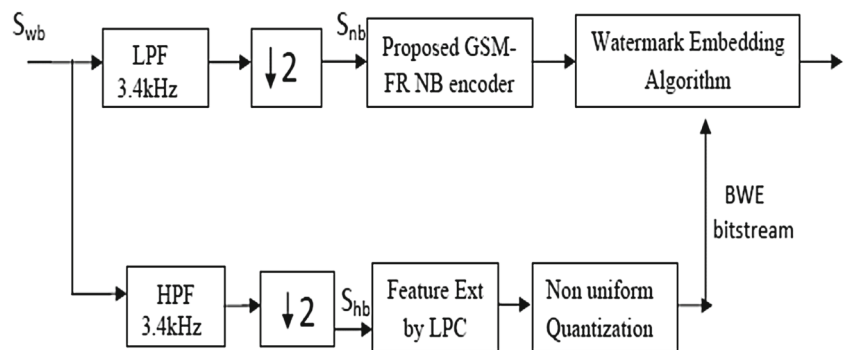
As depicted from Fig. 3, WB speech corpus (having bandwidth up to 7 kHz and sampling frequency of 16 kHz) is provided as an input to the ABE coder. This WB signal is separated into NB and HB signal using LPF and HPF respectively (both having cut-off freq. equal to 3.4 kHz) followed by down sampling both of them with the factor of two. Framing of 160 samples/20ms time frame is applied on the NB down sampled speech signal as per GSM 06.10 FR coder standards. Such each frame is supplied to proposed GSM FR coder [Bhatt and Kosta \(2011\)](#). An encoded bitstream of 224 bits/frame are produced as an outcome and such bitstream are then frame-wise inserted to watermark embedding algorithm. HB features are produced from HB down sampled speech using LPC technique and then encoded, embedded with NB coded signal by watermarking and transmitted over NB wireless link.

At receiver, NB synthetic speech is reproduced using proposed GSM FR NB decoder and up-sampled by factor of two which is then added using OLA method with synthetically generated and interpolated (with a factor of two) HB version of signal to finally regenerate synthetic WB speech signal.

#### 3.2 HB feature extraction

Further, HB down sampled speech is then transferred to LPC block having 8th order filter to produce LP coefficients (1, a<sub>1</sub>, . . . , a<sub>7</sub>) per frame. In order to derive LP coefficients, Levinson Durbin algorithm is introduced as per ([Chen and](#)

**Fig. 3** Transmitter section of proposed ABE coder



Leung 2005; Fuemmeler et al. 2001; Lee and Choi 2013). Also, frame-wise gain parameter is computed from the given HB signal. Thus, eventually HB features are frame-wise represented by LP coefficients and gain co-efficient. Encoding of these HB LP parameters are carried out using non-linear quantization depending upon the subjective importance of individual parameters in accordance with Table 2.

While parameters of the HB speech are estimated using LP analysis filter, and original speech signal is regenerated by the LP synthesis filter from these HB parameters. Therefore, the shape of the HB speech signal's spectral envelope is represented in the LPCs. Speech is processed on frame by frame basis of 20 ms for LP analysis considering the fact that characteristics of speech remains constant for the time interval of one speech frame. This process is followed by pre-emphasizing the speech and then auto regressive (AR) coefficients are computed using LPC of order 8 for each frame.

The proposed method presents the generation of HB excitation signal at receiving end only with the use of NB excitation signal and different extension of excitation techniques. In stark contrast with the other methods, the proposed coder needs not transmit HB excitation information from the transmitter which also incorporates pitch period parameter transmission from transmitting end to the receiver. The major criteria is to reduce the overhead data bits of HB so that embedding of HB feature information distorts the NB coded speech up to certain tolerable limits only (Fig. 4).

### 3.3 Watermark embedding and data hiding

As discussed earlier and also demonstrated in Table 3 of Bhatt et al. (2011), for proposed GSM FR coder, initially Class  $I_b$  contains 96 bits (bit location d50-d145) and the quantized bits of HB information, as shown in Table 2, are embedded at the bit location of d146-d181 (36 bits of Class  $I_b$ ) that finally generates 260 bits/frame as per standard ETSI GSM FR coder. It is to be noted that proposed GSM FR coder produces 224 bits/frame and other 36 bits of quantized HB feature bits are

embedded by watermark embedding algorithm using joint embedding and data hiding approach.

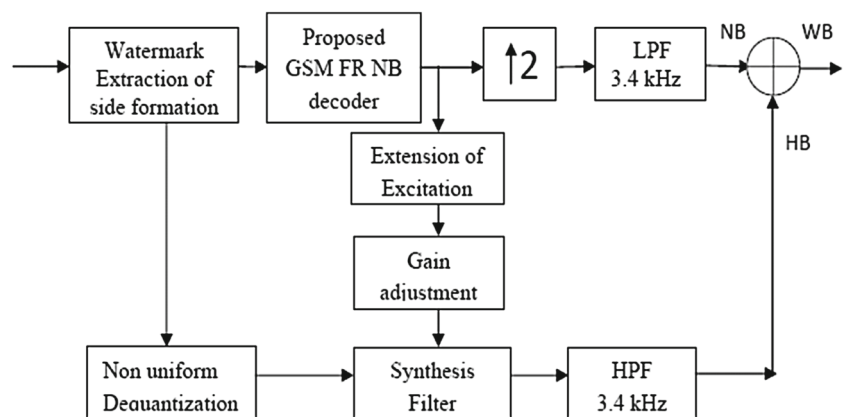
### 3.4 WB speech regeneration at receiver

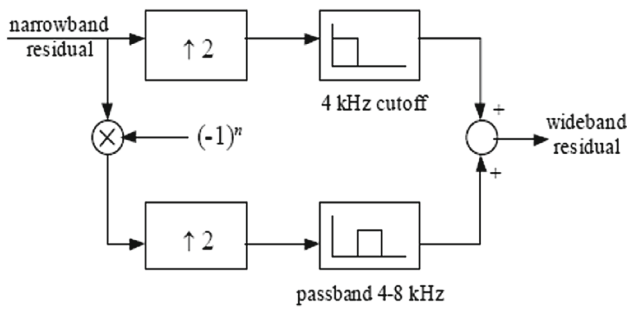
Receiving end accepts NB bitstream on the frame-wise basis (260bits/frame). Watermark extraction algorithm separates out 36 bits (from class  $I_b$  of Table 3 of Bhatt et al. (2011)) for HB speech recovery and passes on the bitstream of 224 bits/frame to legacy proposed GSM FR NB coder for decoding and reproduction of NB version of speech. The produced NB speech can then be up-sampled by the factor of two and subsequently passed through a LPF having cut-off frequency of 3.4 kHz to ultimately regenerate final NB speech. Frame-wise separated 36 bit coded bitstream of HB features by watermark extraction are then dequantized to produce LP coefficients. The excitation signal is produced from the available NB excitation signal using different excitation techniques as will be discussed further. Along with generated HB excitation, the reproduced features are then supplied to their respective blocks like gain adjustment and synthesis filter to effectively reconstruct HB speech. Thereafter, HB speech is processed through interpolation by the factor of two followed by passing it to HPF having cut-off frequency of 3.4 kHz. Finally, both NB and HB recovered signals are summed up using OLA with length equalization to reconstruct WB speech at receiver.

### 3.5 High frequency excitation signal

An important aspect of any wideband enhancement scheme is the generation of the HB / WB speech excitation. Ultimate aim of the excitation (residual) extension process is to double the sampling rate, from 8 to 16 kHz, as well as the whole spectrum should be kept flat. The harmonics contained in the NB residual should also be carried forwarded in the WB residual. Followings are few of the known techniques which are explored in this research for generating HB/WB exci-

**Fig. 4** Receiver section of proposed ABE coder





**Fig. 5** Spectral translation method

tation signal from the given input NB excitations (residual signal).

### 3.5.1 Spectral folding (SF)

This time-domain method for extension of excitation is one of the most popular methods because of its simplicity and wide usage in high frequency excitation regeneration in ABE. Inserting zero between each sample of NB residual signal folds the baseband spectrum to higher frequencies. Thus, the up-sampling extends the signal bandwidth from 4 to 8 kHz. Low-frequency spectrum is flat and consisting harmonics, therefore, the extended HB excitation will also have a flat spectrum and harmonics. Finally, generated WB excitation signal can be used to synthesize WB speech.

Spectrum Folding results in the mirrored image of the NB spectrum at the spectrum of HB but the disadvantage lies in this method is that the harmonic structure leaves a spectral gap at 4 kHz. A possible solution to above issue is to lower the sampling frequency. Hence, NB excitation signal can be first down-sampled to 7 from 8 kHz. Other possible solution is to band pass filter the spectrum between 2 and 3 kHz and it is repeatedly folded up to 3 to 8 kHz. Although having band gap in the middle of the spectrum, this does not cause perceivable artifact in the output speech.

### 3.5.2 Spectral translation (ST)

Spectral Translation produces a shifted spectral version of the NB residual at the high frequency band. One technique for Spectral Translation is depicted in Fig. 5. However, in the

spectral shifting method the input NB residual is mixed with  $(-1)^n$ , where  $n$  is the index of each sample. This is similar to modulation with frequency equal to Nyquist frequency. This modulated signal is up-sampled and high pass filtered and added to the up-sampled and low pass filtered NB residual to produce the WB residual at the output.

Inherent limitation of this approach is that it does not preserve the original NB residual information and this problem can be eliminated by high pass filtering output of the LP synthesis filter, removing the NB portion of the regenerated signal. However, with minor complexity, the performance is still an improvement over the spectral folding method.

### 3.5.3 Non linear distortion (NLD)

In this method, first the NB residual is up-sampled by interpolation (by factor two) and then transferred to a nonlinear function as demonstrated in Eq. 2. The desired bandwidth and harmonic structure can be obtained over the whole spectrum from the resulted distorted signal. After the whitening filter, the resulting signal spectrum is further flattened so that the excitation does not affect the overall spectral shape. Finally, the output will be the WB residual. A simple nonlinear function can be expressed by:

$$y(t) = [(1 + \alpha)|x(t)| + (1 - \alpha)x(t)] / 2 \tag{2}$$

Where,  $x(t)$  is the input signal,  $y(t)$  is the distorted output signal, and  $\alpha$  is a parameter between 0 and 1. When  $\alpha = 1$ , it becomes the absolute value function. In this work, while varying the values of  $\alpha$  within a specific range on the trial and error basis, it is observed that value of  $\alpha = 0.7$  offers overall better results and hence it is chosen and utilized and the obtained results are cited in Tables 3 and 4.

### 3.5.4 Noise modulation (NM)

According to the human auditory system, perceiving capabilities of human ear gradually reduces with the increase in the frequency above 4 kHz. Even the harmonic structure of speech signal is also lost by itself at higher frequencies and becomes more noise-like. This is the reason behind using

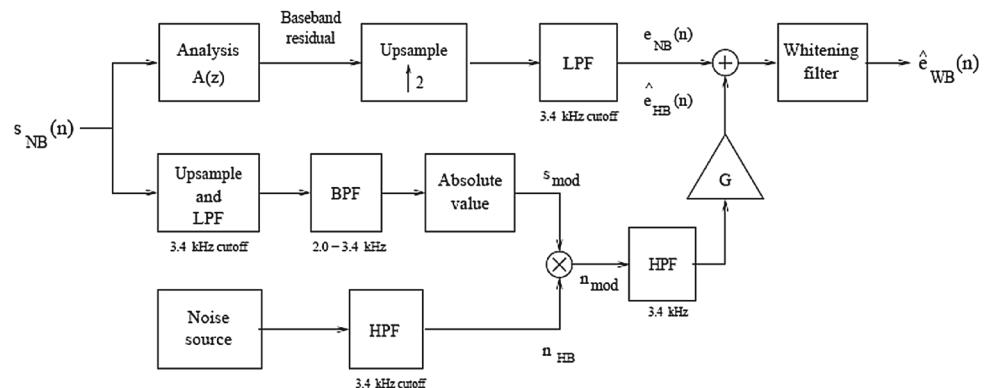
**Table 3** MOS analysis of the proposed ABE coder including different excitation generation methods

Wave file	MOS score for various approaches						
	No. of samples	Legacy GSM FR NB decoder	Proposed ABE coder using LPC technique with different extension of excitation methods				
			ST	SF	NM	NLD	FWR
cc-17	64,768	3.35	3.56	3.55	3.57	3.54	3.58
cc-21	52,992	3.45	3.68	3.66	3.71	3.70	3.73
cc-24	106,240	3.30	3.52	3.53	3.58	3.53	3.60
cc-27	125,440	3.44	3.60	3.62	3.65	3.62	3.64

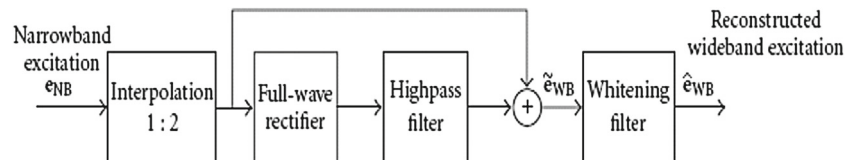
**Table 4** Objective analysis of the proposed method of ABE with side information including the different excitation generation methods

Wave file	PESQ score for various methods						
	No. of samples	Legacy GSM FR NB decoder	Proposed ABE coder using LPC technique with different extension of excitation methods				
			ST	SF	NM	NLD	FWR
cc-17	64,768	2.8089	3.1139	3.0468	3.1332	3.0879	3.1568
cc-21	52,992	3.0052	3.3092	3.2328	3.3160	3.2880	3.2896
cc-26	106,240	2.8240	3.0450	3.0873	3.0229	3.0742	3.1263
cc-38	125,440	2.8966	3.0910	3.0894	3.1481	3.1471	3.1379

**Fig. 6** Noise modulation technique



**Fig. 7** Full wave rectification technique Sagi and Malah (2007)



the Noise Modulation for HB excitation regeneration. This suggests that in HB speech model, the 3–4 kHz band of NB speech is utilized to extract its time domain envelope and the HB excitation is produced by modulating HB noise using this envelope. Figure 6 illustrates the process of noise modulation for HB excitation generation. For the same, white gaussian noise is generated with a positive random sequence having mean equal to 2 and variance of 0.5. This modulated signal is then high pass filtered with cut-off frequency of 3.4 kHz and also the energy is scaled of the resulting HB excitation signal using NB excitation. Further, the up-sampled and low pass filtered NB residual is added in the high frequency modulated noise i.e. HB excitation to regenerate WB excitation. Finally, WB excitation is passed through the whitening filter to have flat spectrum of resulting excitation signal.

This Noise Modulation method is computationally efficient and is motivated from human perception, but it has the drawback of being dependent on the assumption that the time domain envelope of the 3–4 kHz speech band is identical to that of the 4–8 kHz band.

### 3.5.5 Full wave rectification (FWR)

This method demonstrates how WB excitation signal is evolved from NB input speech signal. Initially, the received NB speech is applied to LP analysis filter (inverse filter) which in turn generates NB excitation signal. As can be depicted from Fig. 7, NB excitation which is denoted by  $e_{NB}$  is then passed through an interpolator of 1:2 followed by LPF to convert the NB excitation signal to WB sampling frequency rates. Interpolated and low pass filtered NB speech signal then undergoes FWR followed by HPF. As obvious, the goal of FWR is to typically expand the bandwidth of the given signal. Finally, WB excitations are produced by summing up interpolated NB excitation with HB portion of full wave rectified and high pass filtered excitation signal. Due to the rectification process, WB excitation signal has downward tilt at higher frequency; and in order to compensate for the same, inverse filtering using whitening filter is performed on it Sagi and Malah (2007).

### 4 Performance evaluation & results obtained

In this paper, performance of proposed ABE coder is evaluated using both Objective and Subjective analysis. Each analysis is performed on recovered WB speech generated after different extension of excitation techniques and also on NB legacy GSM FR decoder.

#### 4.1 Subjective analysis

In Subjective analysis, Mean Opinion Score (MOS) rating is carried out for four different clean wave files chosen from WB corpus [http://www.repository.voxforge1.org/downloads/SpeechCorpus/Trunk/Audio/Original/16kHz\\_16bit/](http://www.repository.voxforge1.org/downloads/SpeechCorpus/Trunk/Audio/Original/16kHz_16bit/). MOS analysis is conducted in quiet environment and with high quality headphones. For this analysis, twenty un-trained listeners are chosen to participate in MOS rating. Out of which ten listeners are men and ten are women. Each listener is offered with total of twenty four (4 wave files of legacy NB decoded speech and other 20 wave files of proposed ABE WB recovered speech, between five extension of excitation methods; each having four wave files) wave files. Ratings

given by all twenty listeners (for each wave files and for all cases of excitation extension methods) are then averaged to produce final MOS ratings.

As observed from Table 3 and Fig. 8, obtained results for MOS scores advocate the significant improvement in performance of proposed ABE coder (for all wave files) over legacy GSM FR NB decoder. It is evident from the results obtained that MOS scores for all extension of excitation methods are quite comparable and FWR technique performs slightly better compared to others. It can also be highlighted that for each of the excitation methods, MOS scores are significantly higher in comparison with legacy GSM FR NB recovered speech at receiver.

#### 4.2 Objective analysis

The Objective tool for Perceptual Evaluation of Speech Quality (PESQ) in its wideband version could be used for quality evaluation as per P.862 standards (ITU-T 2000; Kwon et al. 2012; Park 2011). As depicted from Table 4 and Fig. 9, the PESQ reported for the case of different extension of excitation methods of proposed ABE coder are found quite superior

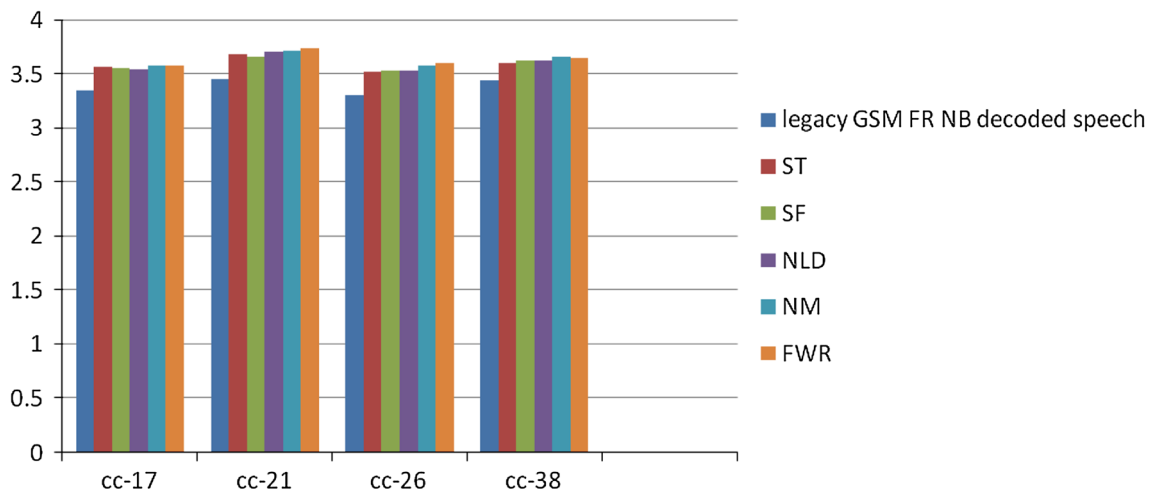


Fig. 8 Comparison chart for MOS score between all methods

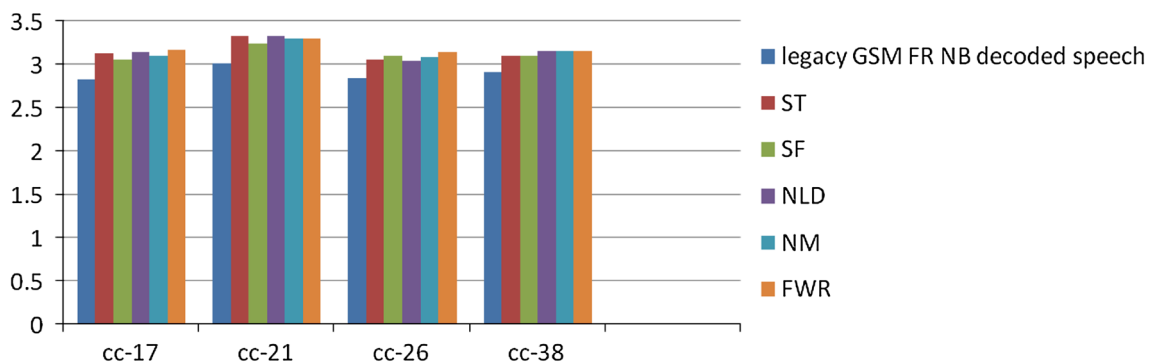


Fig. 9 Comparison chart for PESQ score between all methods

in comparison with PESQ score obtained from legacy GSM FR NB coder for all offered WB speech utterances. While comparing performances between various excitation techniques, in this analysis also, FWR technique claims slightly better results in comparison with the other methods. Results obtained for both Subjective and Objective analysis also advocate that both MOS and PESQ score for all cases and for all utterances are quite comparable for both legacy NB and proposed ABE WB coders.

## 5 Concluding remarks

Inherent limitation imposed by the NB wired and wireless networks lies in its limited quality of recovered speech. Remedy to this problem is improvement in the intelligibility and naturalness of recovered speech which in turn enhances overall speech quality. Without major modifications and up-gradation in current NB transmission channels and end devices, ABE is an emerging technique to reproduce WB comparable speech quality at receiver. Paper demonstrates proposed ABE coder (with side information) using LPC technique in which HB feature parameters are embedded and transmitted into proposed GSM FR NB coder bitstream. These HB parameters, when decoded at receiver, are being utilized to artificially produce synthetic WB speech. In this research, various extension of excitation techniques are implemented and analyzed to compare the overall performance of WB recovered speech produced for all of the above techniques with legacy GSM NB decoded speech. In order to judge the performance of developed ABE coder, Subjective (MOS score) and Objective (PESQ score) analysis are conducted on four selected WB speech corpuses. As witnessed from Tables 3 and 4 and Figs. 8 and 9, both MOS and PESQ scores offer satisfactory and comparable values for all offered cases. It is also evident from the results that in comparison with legacy GSM FR NB decoder, proposed ABE WB decoded speech offer significantly improved MOS and PESQ scores for all speech corpuses and for all excitation techniques. Results for all five excitation techniques are also quite comparable. Still, it remains the fact that FWR approach performs marginally better in contrast with other approaches as witnessed from obtained values for both MOS and PESQ scores.

## References

- Bhatt, N., et al. (2011). "Proposed modifications in ETSI GSM 06.10 full rate speech coder for high rate data hiding and its Objective evaluation of performance using simulink", In *International Conference on Communication Systems and Network Technologies*, IEEE Computer Society, Katra, June 2011.
- Bhatt, N., et al. (2012). "Artificial bandwidth extension of speech & its applications in wireless communication systems: A review, In *Proceedings of IEEE International Conference Communication Systems and Network Technologies*, (Rajkot, India), pp. 563–568, May 2012.
- Bhatt, N., & Kosta, Y. (2011). Proposed modifications in ETSI GSM 06.10 Full Rate speech codec and its overall evaluation of performance using MATLAB. *International Journal of Speech Technology (Springer)*, 14, 157–165.
- Chen, S., & Leung, H. (2005) "Artificial bandwidth extension of telephony speech by data hiding", In *Proceedings of International Symposium on Circuits and Systems (ISCAS)*, Kobe, Japan, May 2005.
- ETSI, Channel coding [GSM 05.03 version 8.9.0 (2005–01), release 1999], pp. 12–19 & 98.
- ETSI, Digital cellular telecommunications system (Phase 2+), Full Rate speech, transcoding, (GSM 06.10 version 8.2.0 Release 2005–2006), pp. 10–59.
- Fuemmeler, J., Hardie, R., & Gardner, W. (2001). Techniques for the regeneration of wideband speech from narrowband speech. *EURASIP Journal on Applied Signal Processing*, 4, 266–274.
- ITU-T. (2000). Perceptual evaluation of speech quality (PESQ), and objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs. ITU-T Rec. P., 862.
- Jax, P., & Vary, P. (2002). "An upper bound on the quality of artificial bandwidth extension of narrowband speech signals", In *Proceedings of ICASSP*, vol. 1. Orlando, FL, USA, 2002, 237–240.
- Jax, P., & Vary, P. (2003). On artificial bandwidth extension of telephone speech. *Signal Processing*, 83(8), 1707–1719.
- Jax, P., & Vary, P. (2006). Bandwidth extension of speech signals: A catalyst for the introduction of wideband speech coding? *IEEE Communication Magazine*, 44(5), 106–111.
- Kwon Y., et al. (2012). "Bandwidth Extension of G.729 speech coder using search-free codebook mapping", 35th IEEE International Conference on Telecommunications and Signal Processing (TSP), Prague, July 2012, pp. 437–440.
- Lee, Y.H., & Choi, S.H. (2013). "Super wideband Bandwidth Extension using normalized MDCT coefficients for scalable speech and audio coding", *Journal of Advances in Multimedia, Hindwai*, Vol. 2013, June 2013.
- Licai, H., & Shuozhong, W. (2006). "Information hiding based on GSM Full Rate speech coding", In *Proceedings of WiCOM, IEEE conference*, Wuhan, Sept 2006.
- Park, N. I., et al. (2011). Artificial bandwidth extension of narrowband speech signals for the improvement of perceptual speech communication quality. *International Journal of Communications in Computer and Information Science, Springer*, 266, 143–153.
- Sagi, A., & Malah, D. (2007). Bandwidth Extension of Telephone Speech Aided by Data Embedding. *EURASIP Journal of Advances in Signal Processing*, vol. 2007. Article no. 64921.