The 3-dB Transcoding Penalty in Digital Cellular Communications

Jerry D. Gibson

Electrical & Computer Engineering, UC Santa Barbara
gibson@ece.ucsb.edu

Dedicated to Toby Berger on the occasion of his 70th birthday

Abstract- **In spite of the widespread attention to data and video, voice is still responsible for up to 75% of the revenue in wireless communications systems today. An unfortunate characteristic of 2nd and 3rd generation digital cellular systems has been the need to transcode at most network interfaces, since the voice codec at the other end of the call is usually unknown and cannot be negotiated. Fourth generation systems such as LTE also require transcoding when the call leaves the LTE network. Transcoding at network interfaces adds complexity, degrades quality, and increases latency, all of which directly impact the quality and cost of voice communications. We investigate the issues in voice communications over tandem connections of wireline and wireless communications links using rate distortion theoretic results and speech coding studies and show that each transcoding operation can incur a 3-dB penalty in source coding performance, in addition to increased latency and complexity. Suggestions for addressing this performance loss are presented.**

I. INTRODUCTION

As is well known, digital cellular calls must enter the wired backbone network to be delivered or connected to the receiving party. What is much less well known, at least outside the digital cellular industry, is that in 2nd and 3rd generation systems, a voice call is transcoded at the BSC/MSC interface to the wired circuit switched or packet switched network. Further, if the receiving party is another digital cellular handset or if the called party is a voice over wireless LAN access point user, another transcoding step may be necessary [1]. Transcoding at network interfaces adds complexity, degrades quality, and increases latency, all of which directly impact the quality and cost of voice communications [2]. Indeed, in many mobile-to-mobile calls, transcoding is the primary source of lost quality and poor user experience, even if the PHY layer connection is excellent.

The reason that transcoding is required is that a mobile placing a call does not know what codec is in the called party handset, and hence, the handset receiving the call may not be able to decode the bitstream created by the caller. Some GSM and CDMA systems now have provisions for negotiating the speech codec at the beginning of the call, but both the caller and the called party must be "in network" and both must have the same codec implemented in the two handsets. Any call that leaves either network must be transcoded. As will be elaborated later in the paper, fourth generation systems such as LTE do not provide much relief from this issue, since any call leaving the network will usually require transcoding [3].

Since voice calls are still the majority source of revenue for wireless service providers today, transcoding deserves careful analysis and proposed solutions. In this paper, we investigate the issues in voice communications over tandem connections of wireline and wireless communications links using rate distortion theoretic results and speech coding studies, and we demonstrate that it is not unusual for a mobile-to-mobile call or a voice over WLAN call to suffer a 3 dB loss in source coding quality, ignoring latency.

The paper is organized as follows. Section II provides additional details concerning how and why transcoding occurs in digital cellular, VoIP, and VoWLAN tandem network connections. Sections III and IV use the Shannon lower bound to explore transcoding losses for Gaussian sources with and without memory and weighted and unweighted squared error fidelity criteria. The loss incurred due to transcoding using today's standardized speech codecs is quantified in Section V, and further discussions of the increased latency and complexity resulting from transcoding are discussed. Section VI presents suggestions for reducing the impact of transcoding.

II. TRANSCODING IN WIRELESS VOICE CALLS

Figure 1 below represents a mobile to mobile call through a VoIP wired backbone, where an incomplete list of possible codecs used in each connection is shown. The key idea is that
when one mobile user places a call, the codec(s) implemented

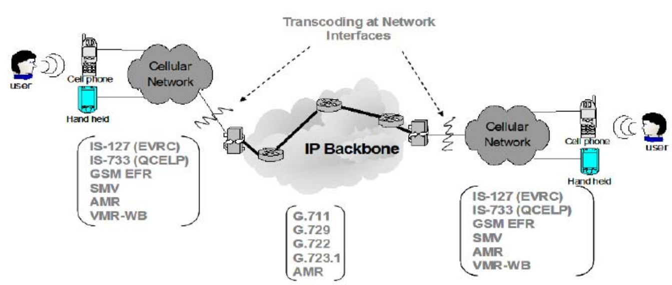

Figure 1. Tandem Voice Coding in Mobile-to-Mobile Calls

in the called party's handset is not known at the caller's handset. Further, unless the call is within a GSM/LTE network, there is no option for the codec to be negotiated so that the codec can be identified and a common codec used. This negotiation option is not widely implemented and certainly not widely implemented as yet in LTE systems [3]. Since the codec is unknown, the call is not just set up as a packet-based, end-to-end VoIP connection. Instead, the coded voice is decoded at the wireless to wired network interfaces (both of them), and then re-encoded. For the wired backbone, the codec used depends on the codec selected by the VoIP provider, often being G.711 at 64 kbps [4], but also possibly G.729 at 8 kbps [5] or AMR-NB at 12.2 kbps [6]. The final result is that the call is transcoded twice, with the attendant loss in quality, added latency, and the complexity of implementing transcoding at each network interface.

If the loss in quality is not too great and the added latency is small, then this is not a serious issue. Unfortunately, both the quality loss and latency involved in transcoding are significant. First, any speech coding method suffers a loss in perceptual quality as the bit rate is reduced. Table 1 presents the PESQ-MOS [7], see also Appendix, of three common speech codecs, including the narrowband adaptive multirate (AMR-NB) codec at three commonly used rates expected to be used in LTE (rates in kbps are shown in parentheses in the table). The G.711 codec at 64 kbps is the benchmark for toll quality for telephony (narrowband speech). Lathe and We were away are two speech utterances described more fully in the appendix. We observe a drop in quality as the rate is decreased, and it can be expected that a change in MOS of 0.5 point will be audible in this range.

Table 1 PESQ-MOS of single codec

Single codec (kbps)	G.711 (64)	AMR-NB (12.2)	AMR-NB (7.95)	AMR-NB (5.9)	G.729 (8)
Lathe	4.495	3.932	3.698	3.403	3.777
We were away	4.315	4.146	3.867	3.780	3.919

PESQ-MOS values for tandem connections of codecs that can commonly occur in today's cellular networks, and in the near future as well, are presented in Table 2.

Table 2 Tandem Results

Tandem	Tandem 1	Tandem 2	Tandem 3	Tandem 4	Tandem 5
Lathe	3.737	3.647	3.354	3.125	3.427
We were away	3.902	3.800	3.619	3.207	3.479

Tandem 1: AMR-NB (12.2 kbps) -> G.711 -> AMR-NB (12.2 kbps)
Tandem 2: AMR-NB (12.2 kbps) -> G.729 -> AMR-NB (12.2 kbps)
Tandem 3: AMR-NB (7.95 kbps) -> G.711 -> AMR-NB (7.95 kbps)
Tandem 4: AMR-NB (7.95 kbps) -> G.729 -> AMR-NB (7.95 kbps)
Tandem 5: AMR-NB (7.95 kbps) -> G.711 -> G.729

Tandem 1 shows a drop of only about 0.2 in MOS compared to 12.2 AMR-NB alone. However, if the backbone VoIP codec is G.729 and there is no codec negotiation possible, Tandem 2 can also occur and another 0.1 point decrease is evident. Tandem 3 is a connection that is more representative of the quality that occurs today in mobile-to-mobile calls, and the total drop in PESQ-MOS compared to G.711 or 12.2 AMR-NB alone is more than 0.5 with tandeming. If G.729 is used as the VoIP codec as in Tandem 4, the performance loss becomes greater than 0.5 due to tandeming alone, and there is audible quality loss evident to almost all users for all sentences. Tandem 5 is a connection that will most likely occur for a mobile to wireless access point VoIP user, and the performance loss is more than 0.5 point in MOS compared to G.711 alone [2].

Latency is another issue and the encoding/decoding latency of each end-to-end tandem connection will consist of one encoding/decoding per codec. The latency associated with the G.711 codec is negligible, so for Tandems 1 and 3, the total latency attributable to the two stages of encoding and decoding required would be about 70 msec, or 35 msec per AMR encoding/decoding. Since the total one way delay for high quality voice communications is 150 msec [2] and the desired end-to-end delay in LTE is 155 msec, [3] the additional 35 msec of latency can prove problematical. However, reducing latency for voice communications is receiving considerable emphasis for fourth generation systems, so latency, in the future, is less of an issue than the loss of quality per tandem transcoding operation. For 2nd and 3rd generation systems, latency is still a major problem and transcoding delay further compromises performance.

III. RATE DISTORTION ANALYSIS OF TANDEM SOURCE CODING

A general tandem communication system can be represented as shown in Fig. 2, wherein source signal X is transmitted through N different links connected in tandem to the destination node where the source is reconstructed as $\hat{X}$. Each of these links may contain a source coder, a noiseless channel, or a noisy channel. We examine each of these situations in the following. The intermediate outputs are expressed as Y_i where $i=1, \ldots, N$. We consider this tandem link connection as a Markov chain such that there is no unmodeled forward side information and there are no feedback channels.

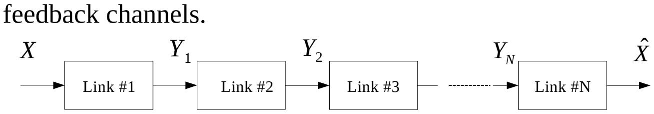

Figure 2. Generalized Tandem System

Case 1:

With respect to Fig. 2, we consider the source vector X to be transmitted over a cascade of communication links, each with a known channel capacity, $C_i, i = 1, 2, \ldots N$. We assume that we operate optimally at rates below capacity for each link, so that no distortion is generated by the channels.

Further, we assume that the source is encoded at the input of Link #1 and decoded only at the output of Link #*N*, so that the only distortion contributed is due to a single source encoding/decoding operation. As a result, for a Gaussian source subject to a mean squared error (MSE) fidelity criterion and given channel capacities, the average distortion is D [8].

Case 2:

We again consider encoding the Gaussian source X subject to the MSE fidelity criterion. Although we still assume no distortion is contributed by any of the channels, we require that the compressed source be decoded at the output of each link and then re-encoded for transmission over the subsequent link as shown in Fig. 2. For an average distortion constraint D_1, the optimally encoded output of Link #1, $\hat{X}_1$, satisfies the Shannon backward channel condition represented as [8]

$$X = \hat{X}_1 + Z_1 \tag{1}$$

where if X is zero mean, Gaussian with variance σ^2, then $\hat{X}$ is zero mean, Gaussian with variance $\mathrm{var}(\hat{X}) = \sigma^2 - D_1$, with Z_1 a zero mean Gaussian random variable with variance D_1 that is statistically independent of $\hat{X}_1$. If we now pass $\hat{X}_1$ on as input to tandem link 2 and it is to be encoded subject to the MSE fidelity criterion with average distortion D_2, then according to the Shannon backward channel condition, we have

$$\hat{X}_1 = \hat{X}_2 + Z_2 \tag{2}$$

where $\hat{X}_2$ and Z_2 are statistically independent with variances, $\mathrm{var}(\hat{X}_2) = \mathrm{var}(\hat{X}_1) - D_2$ and $\mathrm{var}(Z_2) = D_2$, respectively. Using $\mathrm{var}(\hat{X}_1) = \sigma^2 - D_1$, we find that $\mathrm{var}(\hat{X}_2) = \sigma^2 - D_1 - D_2$. If we then pass $\hat{X}_2$ on as input to the third link, subject to MSE encoding at average distortion D_3, we again use the Shannon backward channel result to find that

$$\hat{X}_2 = \hat{X}_3 + Z_3 \tag{3}$$

where $\hat{X}_3$ and Z_3 are statistically independent with variances $\mathrm{var}(\hat{X}_3) = \sigma^2 - \sum_{i=1}^{3} D_i$ and $\mathrm{var}(Z_3) = D_3$, respectively. Now, $\mathrm{var}(X - \hat{X}_3) = \sum_{i=1}^{3} \mathrm{var}(Z_i) = \sum_{i=1}^{3} D_i$, since the Z_i's are statistically independent, we see that the distortion in encoding the original source X is accumulating.

This is in contrast to Case 1 wherein the average distortion is equivalent to only one source encoding/decoding operation since we did not require decoding at the output of each link and re-encoding. Case 2 is the situation that occurs now in some network interconnections and is expected to occur more often in the future.

IV. AUTOREGRESSIVE SOURCE WITH WEIGHTED DISTORTION MEASURE

The tandem coding analyses in the prior section considers memoryless Gaussian sources and the squared error fidelity criterion; however, real speech signals are often modeled by autoregressive (AR) sources and the distortion measures used in practice are frequency weighted squared error. Therefore, in this section we extend the analyses in Sec. III to these more general conditions. The Shannon backward channel formulation is key to obtaining meaningful results.

An *m*th-order, time-discrete AR source can be expressed as

$$X_t = -\sum_{k=1}^{m} a_k X_{t-k} + Z_t \tag{4}$$

where $a_1, \ldots, a_m$ are the AR coefficients, and $\{Z_t\}$ is a sequence of iid random variables, and X_r and Z_s are statistically independent if $s>r$. We analyze the optimal encoding of this source subject to a weighted squared error distortion measure by employing a diagonalizing transform and imposing the Shannon backward channel condition [8], so that after some manipulation we find that the reconstructed source has the *z*-domain power spectral density (psd) [9]

$$\Phi_Y(z) = \Phi_X(z) - \frac{D}{|W(z)|^2} \tag{5}$$

where the average distortion is D, $\Phi_X(z)$ is the psd of the source, and $W(z)$ is the frequency weighting of the distortion. This is for one encoding, and among the conclusions available from this result is the fact that the reconstructed output is no longer purely AR, but it is now an autoregressive moving average (ARMA) sequence even if $W(z) = 1$. The implication being that if the resulting output of the first stage is the input to another codec, the codec in the second stage should not be designed for an AR process matched to the original source.

Continuing to analyze the tandem connection of codecs as indicated in Fig. 1, we find that for a tandem connection of n codecs perhaps using different distortion measures and for different average distortions with the output of the previous stage serving as input to the next stage, we have for the psd of the output of the *n*th stage,

$$\Phi_{Y_n}(z) = \Phi_{Y_{n-1}}(z) - \frac{D_n}{|W_n(z)|^2} = \Phi_X(z) - \sum_{i=1}^{n} \frac{D_i}{|W_i(z)|^2} \tag{6}$$

Using the fact that the input source is an AR sequence, the output psd becomes [9]

$$\Phi_{Y_n}(z) = \frac{\sigma^2 - |A(z)|^2 \sum_{i=1}^{n} \frac{D_i}{|W_i(z)|^2}}{|A(z)|^2} \tag{7}$$

which is valid for the small distortion condition

$$\sum_{i=1}^{n} \frac{D}{|W_i(z)|^2} < \Phi_X(z) \tag{8}$$

The results in Eqs. (5) and (7) are particularly illuminating and the Shannon backward channel formulation is a very effective tool in the tandem coding analysis. For the AR source in Eq. (4), the source psd is

$$\Phi_X(z) = \frac{\sigma^2}{|A(z)|^2} \tag{9}$$

So when we compare to Eqs. (5) and (7), we see that if we desired the output power spectral density to be the same as the source, the tandem coding and perceptual weighting is causing some apparently serious degradation.

The results are particularly interesting since the distortion spectral shaping is due to the perceptual weighting inside the CELP analysis-by-synthesis loop, and also since the distortion is not purely multiplicative as in the common postfiltering problem. It is here that the Shannon backward channel formulation has been extremely helpful in allowing us to gain insight into the effects of the perceptual weighting that has not been exposed by prior methods.

The significance of the additional spectral distortion due to tandem coding is illustrated in Fig. 3 below where we show one voiced frame from a two stage tandem coding for AMR-NB at 6.7 kbps [9]. After one stage the 2nd and 3rd formants are offset relative to the input speech, and for two stages, the 2nd formant is almost entirely lost and the 3rd formant is offset and 10 dB down from the correct input formant.

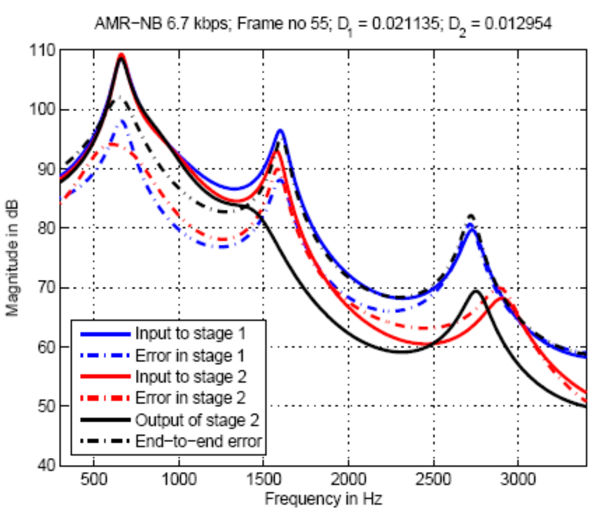

Figure 3. Spectra for Two Stage Tandem Coding with AMR-NB at 6.7 kbps

One reason for the severe distortion in this case may be that the small distortion requirement in Eq. (8) might not be satisfied. In fact, if the codec is designed to achieve small distortion at the lowest possible bit rate, more than a single encoding/decoding will violate this condition.

V. THE 3-dB TRANSCODING LOSS

The loss in tandem coding for real speech codecs is characterized in terms of PESQ-MOS in Sec. II, while the rate distortion analysis presented in Sec. III utilizes the MSE distortion measure and the development in Sec. IV utilizes the frequency weighted squared error distortion measure employed in today's standardized CELP codecs. In order to obtain a quantitative indicator to characterize the loss due to tandem coding of real codecs, we conducted an experiment to relate PESQ-MOS and MSE for some well known waveform-following codecs for which MSE is a meaningful performance indicator. In particular, we tabulated PESQ-MOS/MSE pairs for G.726 and G.727 voice codecs at rates of 40, 32, 24, and 16 kbps. Since G.727 is an embedded coder, there are multiple choices of quantization in the prediction loop for each codec rate, so in G.727 combinations are referred to by (x,y) pairs where x refers to the total of both enhancement and core bits, which sets the transmitted bit rate, and y refers to the number of core bits used in the predictor coefficient adaptation loop. ITU-T G.727 Recommendation provides coding rates of 40 kbps for the 3 combinations (5,4), (5,3), and (5,2), 32 kbps for 3 combinations (4,4), (4,3), and (4,2), 24 kbps for 2 combinations (3,3) and (3,2), and 16 kbps for one combination (2,2), resulting in nine pairs of coding rates.

Therefore with the 4 coding rates for G.726 and the 9 coding rates for G.727, we have 13 MSE and PESQ pairs to generate a mapping function. We then fit these pairs with a mapping function that allows us to map PESQ-MOS into MSE, and the result is shown in Fig. 4 below [10].

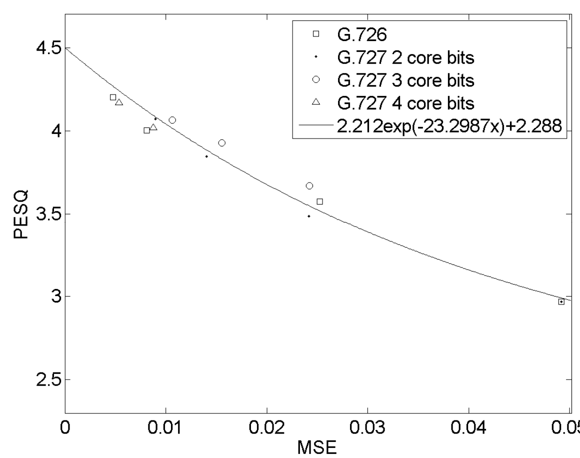

Figure 4. MSE to PESQ-MOS Mapping Function

Since the commonly occurring tandem connection of codecs labeled Tandem 3 in Table 2 has a total drop in PESQ-MOS from 4.0 to 3.5 (roughly), we see from the mapping function that this corresponds to an increase in MSE by at least a factor of 2, or a loss in SNR of 3 dB. The conclusion is thus that a mobile to mobile call commonly suffers a 3 dB penalty compared to toll quality performance. This loss varies depending upon the quality of the codecs at each end of the call, the codec in the VoIP backbone, and the codec chosen as the reference point, but with respect to toll quality achieved by G.711, this loss is at least 3 dB.

VI. PROPOSED IMPROVEMENTS

Aside from keeping all mobile calls within a network that allows codec negotiation, no response is planned by the industry to the tandeming problem. Furthermore, any idea that one codec will become the de facto standard throughout the World is not realistic. So, what can be done?

There are several possible approaches. One approach would be to conduct research to obtain a new speech codec that has high quality at low bit rates, low encoding/decoding delay, low complexity, and good tandeming performance (for backward compatibility until the networks incorporate the new codec). A second approach would be to perform the transcoding between codec pairs without full decoding. That is, rather than decode back to reconstructed speech, the codec parameters from the first codec could be mapped into the codec parameters of the following codec. This is the approach studied in [11] for the tandem connection of G.729 at 8 kbps with the IS-641 codec at 7.4 kbps (IS-641 was standardized for the North American TDMA systems). Since both of these codecs are CELP based, it is necessary to map fixed codebooks, adaptive codebooks, and linear prediction parameters. The effort involved in the design is far from trivial but the resulting mappings did improve the speech quality over a direct tandem and both complexity and delay were reduced as well. However, the quality was not uniformly better and latency was only about 5 msec less than direct decoding. Of course, a further difficulty with this mapping approach is that mappings between all possible codec pairs that can occur at the network interfaces must be developed and implemented at each interface.

There are other possible approaches as well. Further insights into addressing the tandem coding problem can be obtained by considering Fig. 5 below where we show one voiced frame from a two stage tandem coding for AMR-NB at 12.2 kbps [9]. Here we see that while the third formant is being attenuated and shifted, the distortion is not nearly so substantial as in the figure for the 6.7 kbps case. Thus, an approach to reducing distortion accumulation due to tandeming would be to maximize the bit rate of the voice codec. Note, however, that this can reduce the distortion but not eliminate it.

One reason for the increasing spectral distortion with multiple tandems is that the coefficients for the perceptual weighting filter inside the analysis by synthesis loop in CELP methods are calculated based upon the input to the current stage. While calculating parameters for encoding the current input based on the current input would appear well-motivated, we may ask what would happen if the perceptual weighting is kept the same for every stage of tandem coding (were that possible). We see from Eq. (6) that if $W_i(z) = W(z)$ for all i,

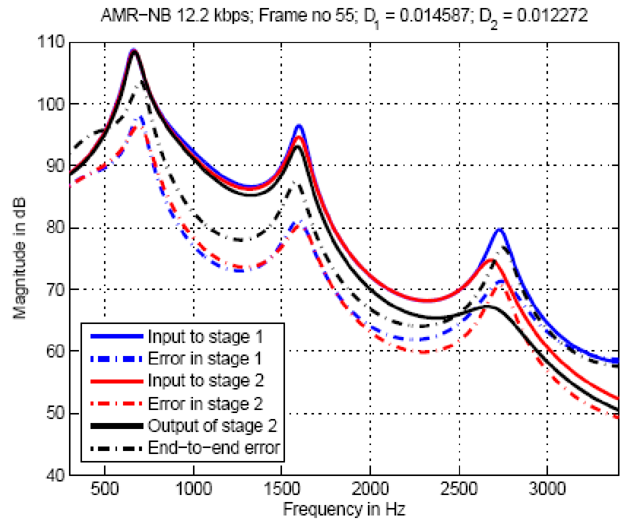

Figure 5. Spectra for Two Stage Tandem Coding with AMR-NB at 12.2 kbps

then the psd of the output of the nth stage is

$$\Phi_{Y_n}(z) = \Phi_X(z) - \frac{\sum_{i=1}^{n} D_i}{|W(z)|^2} \tag{10}$$

Therefore, that the spectral shaping does not change, although the gain of the shaping function will change with each tandem coding stage. An indication that maintaining the same perceptual weighting function at each tandem coding stage

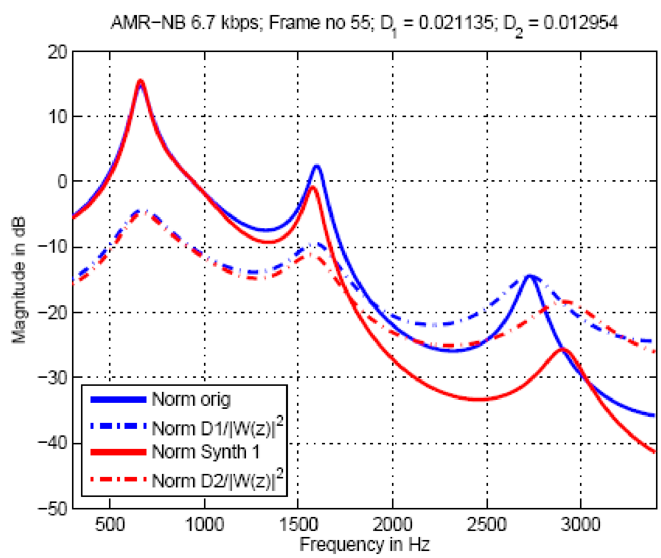

Figure 6. Normalized Spectra for Two Stage Tandem Coding with AMR-NB at 6.7 kbps

would reduce the relocation of formants is evident in Fig. 6 [9]. Here it is seen that the error due to the weighting after one stage of encoding (top dashed spectrum) retains the correct formant location, whereas the error spectral shape after the second encoding (the lower dashed spectrum) has shifted the third formant. Note that in the figures all gains have been normalized so the accumulated distortion in Eq. (10) due to the summation is normalized out.

In order to reuse the first stage weighting function, the coefficients calculated at the first stage would have to been passed forward at each interface and the code implementing the CELP based codecs would need to be modified so that the coefficients in the perceptual weighting and those used in the encoding could be different. This last approach has the appeal of being much less invasive than the other methods suggested.

Notice that the two-stage tandem coding results presented in Figs. 3 and 5 each are for a codec tandemed with the same codec at the same rate in the following stage. If different CELP codecs are tandemed, the results can be expected to be poorer than those shown. However, the last proposed approach of retaining the same perceptual weighting function at every stage would still be possible.

VII. Conclusions

The significance of the requirement to transcode at network interfaces in digital cellular communications is highlighted with respect to performance loss, latency, and complexity. The need to transcode for mobile calls outside of a heterogeneous wireless network will persist even with the widespread introduction of the fourth generation LTE systems. Several approaches to addressing this problem are presented, with an emphasis on the insights provided by a rate distortion analysis using the Shannon backward channel theorem for autoregressive sources and the weighted squared error distortion measure.

ACKNOWLEDGMENT

The author gratefully acknowledges his student, Ying-Yi Li, for her assistance with Tables 1 and 2 and Figure 4. This paper is dedicated to Toby Berger on the occasion of his 70th birthday and in recognition of his profound contributions to rate distortion theory, and in particular, for his landmark textbook [8]. This research has been supported by NSF Grant Nos. CCF-0728646 and CCF-0917230.

APPENDIX
PESQ-MOS

Perceptual evaluation of speech quality (PESQ) is an objective method for end-to-end speech quality assessment of narrowband speech codecs [7]. The distance between the original and degraded speech signal, called the PESQ score, is calculated based on the PESQ perceptual model. The PESQ score is mapped to a MOS-like scale by a monotonic function. The MOS-like PESQ (PESQ-MOS) is a single number in the range of -0.5 and 4.5, although for most cases the output range will be between 1.0 and 4.5, the normal range of MOS values found in an ACR listening quality experiment. Even though PESQ-MOS is not the same as MOS, and it has known limitations, it is a standardized objective measure for evaluating the perceptual performance of speech codecs that is widely used and quoted. PESQ-MOS is described in detail in the ITU-T P.862 Recommendation.

References

[1] B. Wei and J. D. Gibson, "Tandem Voice Communications: Digital Cellular, VoIP, and Voice Over Wi-Fi," 2004 IEEE Globecom, Dallas, TX , Nov. 29-Dec. 3, 2004.

[2] J. D. Gibson, “Speech Coding Methods, Standards, and Applications,” *IEEE Circuits and Systems Magazine*, vol. 5, no. 4, pp. 30 – 49, 2005.

[3] H. Holma and A. Toskala, *LTE for UMTS-OFDMA and SC-FDMA Based Radio Access*, Wiley, 2009.

[4] J. D. Gibson, T. Berger, T. Lookabaugh, D. Lindbergh, R. Baker, *Digital Compression for Multimedia: Principles & Standards*, Morgan Kaufmann Publishers, Inc., 1998.

[5] R. Salami, et al, “Design and Description of CS-ACELP: A Toll Quality 8 kb/s Speech Coder,” *IEEE Trans. Speech and Audio Processing*, vol. 6, pp. 116-130, March 1998.

[6] B. Bessette et al., “The Adaptive Multirate wideband speech codec (AMR-WB)”, *IEEE Trans. Speech and Audio Processing*, pp. 620-636, Nov. 2002.

[7] ITU-T Recommendation P.862, “Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end Speech Quality Assessment of Narrow-band telephone networks and Speech Codecs,” Feb.2001.

[8] T. Berger, *Rate-Distortion Theory: A Mathematical Basis for Data Compression*, Prentice Hall, Englewood Cliffs, NJ, 1971.

[9] N. A. Shetty, “Tandeming in Multi-Hop Voice Communication,” Ph. D. Dissertation, Dec.2007.

[10] Y.-Y. Li and J. D. Gibson, “Rate Distortion Bounds for Speech Coding Based on A Perceptual Distortion Measure (PESQ-MOS),” submitted to ICME, Barcelona, July 2011.

[11] H.-G. Kang, H. K. Kim, and R. V. Cox, “Improving the Transcoding Capability of Speech Coders,” *IEEE Trans. Multimedia*, vol. 5, pp. 24-33, March 2003.