ETSI TS 146 022 V9.0.0 (2010-02)

Technical Specification

Digital cellular telecommunications system (Phase 2+);
Half rate speech;
Comfort noise aspects for the half rate speech traffic channels
(3GPP TS 46.022 version 9.0.0 Release 9)

GLOBAL SYSTEM FOR MOBILE COMMUNICATIONS

Reference
RTS/TSGS-0446022v900

Keywords
GSM

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00   Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or
perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF).
In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive
within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status.
Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Contents

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The present document gives the detailed requirements for the correct operation of the background acoustic noise evaluation, noise parameter encoding/decoding and comfort noise generation within the digital cellular telecommunications system. The present document is part of a series covering the half rate speech traffic channels as described below:

GSM 06.02 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech processing functions".

GSM 06.06 "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec".

GSM 06.07 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Test sequences for the GSM half rate speech codec".

GSM 06.20 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech transcoding".

GSM 06.21 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Substitution and muting of lost frames for half rate speech traffic channels".

GSM 06.22 **"Digital cellular telecommunications system (Phase 2+); Half rate speech; Comfort noise aspects for half rate speech traffic channels".**

GSM 06.41 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Discontinuous Transmission (DTX) for half rate speech traffic channels".

GSM 06.42 "Digital cellular telecommunications system (Phase 2+); Half rate speech; Voice Activity Detector (VAD) for half rate speech traffic channels".

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document gives the detailed requirements for the correct operation of the background acoustic noise evaluation, noise parameter encoding/decoding and comfort noise generation in GSM Mobile Stations (MS)s and Base Station Systems (BSS)s during Discontinuous Transmission (DTX) on half rate speech traffic channels.

The requirements described in the present document are mandatory for implementation in all GSM MSs capable of supporting the half rate speech traffic channel.

The receiver requirements are mandatory for implementation in all GSM BSSs capable of supporting the half rate speech traffic channel, the transmitter requirements are only for those where downlink DTX will be used.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1]	GSM 01.04: "Digital cellular telecommunication system (Phase 2+); Abbreviations and acronyms".
[2]	GSM 06.20: "Digital cellular telecommunications system (Phase 2+); Half rate speech transcoding".
[3]	GSM 06.41: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Discontinuous Transmission (DTX) for half rate speech traffic channels".
[4]	GSM 06.42: "Digital cellular telecommunications system (Phase 2+); "Half rate speech; Voice Activity Detector (VAD) for half rate speech traffic channels".
[5]	GSM 06.06: "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply.

frame: time interval of 20 ms corresponding to the time segmentation of the half rate speech transcoder, also used as a short term for a traffic frame.

H(Z): combination of the short term (spectral) filter A(z) and the spectral weighting filter W(z).

SID codeword: fixed bit pattern for labelling a traffic frame as a SID frame.

SID field: bit positions of the SID codeword within a SID frame.

SID frame: frame characterized by the SID (Silence Descriptor) codeword. It conveys information on the acoustic background noise.

SP flag: speech flag.

speech frame: traffic frame that cannot be classified as a SID frame.

VAD flag: Voice Activity Detector flag.

W(Z): spectral weighting filter of the GSM half rate speech codec.

Other definitions of terms used in the present document can be found in GSM 06.20 [2] and GSM 06.41 [3]. The overall operation of DTX is described in GSM 06.41 [3].

3.2 Symbols

For the purposes of the present document, the following symbols apply:

GS	Energy tweak parameter.
R0	Frame energy value.
R(i)	Unquantised (normalized) autocorrelation sequence.
r_j	Optimal reflection coefficient.
b	
SUM (x(n))	$= x(a) + x(a+1) + ... + x(b-1) + x(b);$ (Accumulation).
n=a	
GSP0 codeword	Vector quantization index, joint vector quantization of the parameters GS and P0.
P0	Power contribution of the first excitation vector as a fraction of the total excitation power at a subframe.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AFLAT	Autocorrelation Fixed Point LAttice Technique (used in the GSM half rate speech codec for the vector quantization of the LPC coefficients)
BSS	Base Station System
DTX	Discontinuous Transmission
ETS	European Telecommunication Standard
GSM	Global System for Mobile communications
MS	Mobile Station
SID	SIlence Descriptor
RX	Receive
TX	Transmit
VAD	Voice Activity Detector
VQ	Vector Quantization

For abbreviations not given in this subclause, see GSM 01.04 [1].

4 General

A problem when using DTX is that the background acoustic noise, which is transmitted together with the speech, would disappear when the radio transmission is switched off, resulting in a modulation of the background noise. Since the DTX switching can take place rapidly, it has been found that this effect may be annoying for the listener, especially in a car environment with high background noise levels. In bad cases, the speech may be hardly intelligible.

The present document specifies a solution to overcome this problem by generating synthetic noise similar to the transmit (TX) side background noise on the receive (RX) side. The comfort noise parameters are estimated on the TX side and transmitted to the RX side before the radio transmission is switched off and at a regular low rate afterwards. This allows the comfort noise to adapt to the changes of the noise on the TX side.

5 Functions on the transmit (TX) side

The comfort noise evaluation algorithm uses the following parameters of the GSM half rate speech encoder, defined in GSM 06.20 [2]:

- the unquantized frame energy value R0;

- the unquantized (normalized) autocorrelation sequence R(i) derived from the optimal reflection coefficients r_j;

- the quantized energy tweak parameter GS.

These parameters give information on the level (R0 and GS) and the spectrum (R(i)) of the background noise.

Two of the evaluated comfort noise parameters (R0 and R(i)) are encoded into a special frame, called a SIlence Descriptor (SID) frame, for transmission to the RX side. While the energy tweak parameter GS can be evaluated in the encoder and decoder in the same way as given in subclause 5.1, therefore no transmission of GS is necessary.

The SID frame also serves to initiate the comfort noise generation on the RX side, as a SID frame is always sent at the end of a speech burst, i.e. before the radio transmission is terminated.

The scheduling of SID or speech frames on the radio path is described in GSM 06.41 [3].

5.1 Background acoustic noise evaluation

The comfort noise parameters to be encoded into a SID frame are calculated over 8 consecutive frames marked with Voice Activated Detector (VAD) flag = "0", as follows:

The frame energy values shall be averaged according to the equation:

$$\text{mean}(R0[j]) = 1/8 \sum_{n=0}^{7} R0[j-n];$$

where:

R0[j]   is the frame energy value of the current frame j (n=0);

R0[j-n]   is the frame energy of the previous frames (n=1,...,7);

n   is the averaging period index n=0,1,...,7;

j   is the frame index.

The averaged value mean(R0[j]) is encoded using the same encoding table that is also used by the GSM half rate speech codec for the encoding of the non-averaged R0 values in ordinary speech encoding mode.

The (normalized) autocorrelation sequence R(i) shall be averaged according to the equation:

$$\text{mean}(R[j](i)) = 1/8 \sum_{n=0}^{7} R[j-n](i) \quad i = 0,1,2...,10;$$

where:

R[j](i)   is the i'th autocorrelation value of the current frame j (n=0);

R[j-n](i)   is the i'th autocorrelation value of one of the previous frames (n=1,...,7);

n   is the averaging period index n=0,1...,7;

j   is the frame index.

The averaged values mean(R[j](i)) are used as input parameters of the Autocorrelation Fixed Point LAttice Technique (AFLAT) recursion algorithm which calculates the Vector Quantization (VQ) indices of the reflection coefficients, see GSM 06.20 [2].

The SID frame containing the quantization index of mean(R0[j]), the VQ indices of mean(R[j](i)) and the SID codeword is passed to the radio subsystem instead of frame number j (see subclause 5.3, SID-frame encoding).

The averaging of the energy tweak parameters GS is made on the basis of the quantized GS parameters. The quantized GS parameters can be derived from the GSP0 indices. These indices are used as pointers to the GSP0 vector quantization codebook. The GS components of the selected GSP0 vectors are the quantized GS values which will be averaged.

The quantized energy tweak parameters GS shall be averaged according to the equation:

$$\text{mean}(GS[j]) = \frac{1}{28} \sum_{n=1}^{7} \left(\sum_{i=1}^{4} GS[j-n](i) \right);$$

where:

GS[j](i)   is the quantized energy tweak parameter in subframe i of the current frame j   (n=0);

GS[j-n](i) is the quantized energy tweak parameter in subframe i of one of the last frames   (n=1,...7);

n   is the averaging period index n=1,2,...,7;

i   is the subframe index i=1,2,3,4;

j   is the frame index.

NOTE: The averaging of GS is made over 7 frames only.

For each comfort noise insertion period, the averaging of the GS parameters is done only once before sending the first SID frame to the decoder and for the rest of the comfort noise insertion period, the averaged value mean(GS[j]) will be frozen.

Under normal conditions, the averaging of the GS parameters is done during the hangover period, but in case of short speech bursts handling, the hangover period can be skipped under certain conditions, see GSM 06.41 [3]. In such cases, the GS parameters of the last seven speech frames marked with SP flag="1" are averaged.

The hangover period is defined in GSM 06.41 [3]. It is a period added at the end of a speech burst in which no voice activity is detected (VAD flag="0"), but the speech encoder stays for the processing of 7 speech frames in speech encoding mode (SP flag= "1"). This hangover period and the first SID frame are used for averaging the comfort noise parameters contained in the first SID frame.

mean(GS[j]) can be evaluated at the decoder in the same way as in the encoder, because in both the encoder and decoder, the GSP0 indexes of the last 7 speech frames shall be kept in memory. In case of an error free transmission, the GSP0 indexes are identical at the encoder and decoder.

5.2 Modification of the speech encoding algorithm during SID frame generation

When the SP flag is equal to "0", the speech encoding algorithm is modified in the following way:

- the non-averaged reflection coefficients which are used to derive the filter coefficients of the filters H(z) and W(z) of the speech encoder are not quantized;

- the unvoiced speech encoding mode is forced. This simplifies the open loop long term prediction processing: only the integer lags have to be calculated, no determination of fractional lags is necessary and the frame lag trajectory derivation can be avoided;

- no fixed codebook search is made. In each subframe, the indices of both fixed codebooks (CODE1_1, ...,CODE1_4 and CODE2_1, ...,CODE2_4) are replaced by pseudo random numbers uniformly distributed in [0,127] (7 bit random numbers);

- no GSP0 determination is made. The GSP0 codeword is selected as follows:

  - at the beginning of a comfort noise insertion period, mean(GS[j]) is calculated as defined in subclause 5.1. Then mean(GS[j]) is quantized, using only the GS component of the GSP0 vector quantization codebook of the unvoiced speech encoding mode as quantization table. The P0 parameter is not averaged. For this parameter, the value is used which is associated with the quantized mean(GS[j]) value in the GSP0 codebook of the unvoiced speech encoding mode. For the rest of the comfort noise insertion period, the GSP0 indices are frozen.

A simplified block diagram of the GSM half rate speech encoder in comfort noise insertion mode is shown in figure 1.
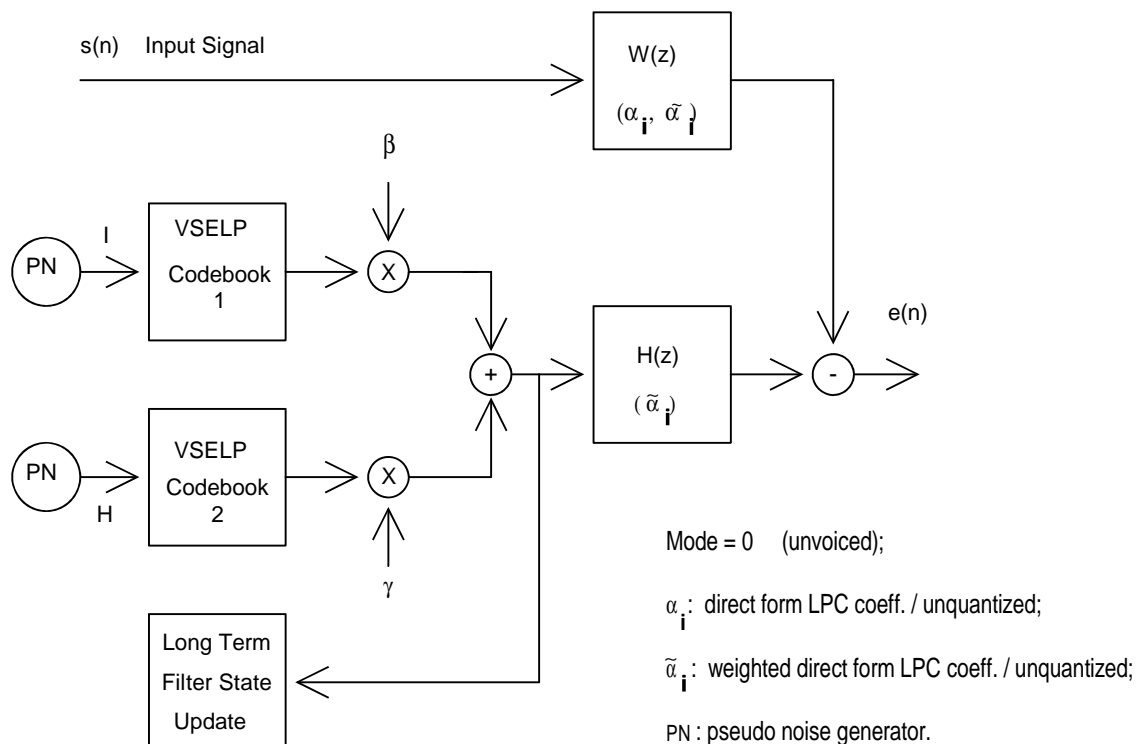

Figure 1: GSM half rate speech encoder in comfort noise insertion mode

5.3 SID-frame encoding

The SID frame encoding algorithm exploits the fact that only some of the 112 bits in a frame are needed to code the comfort noise parameters. The other bits can then be used to mark the SID frame by means of a fixed bit pattern, called the SID codeword.

SID frames are encoded in the encoder output format for voiced frames (MODE = 3), because the two voicing mode bits are part of the SID codeword.

The index of the frame energy value R0 is replaced by the quantization index derived from mean(R0[j]). mean(R0[j]) is defined in subclause 5.1 and is encoded as described in GSM 06.20 [2].

The VQ indices of the reflection coefficients are replaced by VQ indices derived from mean(R[j](i)). mean(R[j](i)) is defined in subclause 5.1 and the VQ of the reflection coefficients is described in GSM 06.20 [2].

The SID codeword consists of 79 bits which are all "1". To mark a frame as a SID frame, the parameters in table 1 have to be set as shown.

Table 1: SID codeword

Parameter	Number of bits	Value (Hex)
MODE	2	0x0003
INT_LPC	1	0x0001
LAG_1	8	0x00ff
LAG_2	4	0x000f
LAG_3	4	0x000f
LAG_4	4	0x000f
CODE_1	9	0x01ff
CODE_2	9	0x01ff
CODE_3	9	0x01ff
CODE_4	9	0x01ff
GSP0_1	5	0x001f
GSP0_2	5	0x001f
GSP0_3	5	0x001f
GSP0_4	5	0x001f

The parameters in table 1 are defined in GSM 06.20 [2].

6 Functions on the receive (RX) side

The situations in which comfort noise shall be generated on the RX side are defined in GSM 06.41 [3] and may be started or updated whenever a valid SID frame is received.

6.1 Averaging of the GS parameters

When speech frames are received by the decoder, the GS parameters of the last seven speech frames shall be kept in memory. As soon as a SID frame is received, these stored GS parameters shall be averaged. The averaged GS value will be frozen and used for the actual comfort noise insertion period.

The averaging procedure works as follows:

- when a speech frame is received, the GSP0 indices are decoded and the decoded GS parts of these parameters are stored in memory;

- when the first SID frame is received, the stored GS values are averaged in the same way as in the speech encoder as follows (see also subclause 5.1):

$$\text{mean} (GS[j]) = \frac{1}{28} \sum_{n=1}^{7} \left(\sum_{i=1}^{4} GS[j-n](i) \right);$$

where:

GS[j](i) is the quantized energy tweak parameter in subframe i of the current frame j;

GS[j-n](i) is the quantized energy tweak parameter in subframe i of one of the last frames;

n is the averaging period index n=1,2,...,7;

i is the subframe index i=1,2,3,4;

j is the frame index;

- then mean(GS[j]) is quantized, using the GS component of the GSP0 vector quantization codebook for the unvoiced speech encoding mode as quantization table. The resulting index of this quantization is used for one complete comfort noise insertion period as GSP0 codeword. The P0 parameter is not averaged. For this parameter, the value is used which is associated with the quantized mean(GS[j]) value in the GSP0 codebook of the unvoiced speech encoding mode.

6.2 Comfort noise generation and updating

The comfort noise generation procedure uses the GSM half rate speech decoder algorithm defined in GSM 06.20 [2].

When comfort noise is to be generated, then the various encoded parameters are set as in table 2.

Table 2: Comfort noise encoded parameters

Parameter	Value
MODE	0
R0	interpolation of the values received in the last two valid SID frames
LPC1 LPC2 LPC3	interpolation of the values received in the last two valid SID frames
INT_LPC	1
CODE1_1 CODE1_2 CODE1_3 CODE1_4 CODE2_1 CODE2_2 CODE2_3 CODE2_4	pseudo random numbers uniformly distributed in [0,127] (7 bit numbers)
GSP0_1 GSP0_2 GSP0_3 GSP0_4	index of the averaged GS parameter (calculated at the beginning of each comfort noise insertion period and frozen for the rest of the period)

With these parameters, the speech decoder now performs the standard operations described in GSM 06.20 [2] and thereby synthesizes comfort noise.

Updating of the comfort noise parameters (frame energy and LPC coefficients) occurs each time a valid SID frame is received, as described in GSM 06.41 [3].

NOTE: The GSP0 codewords are not updated, they are frozen during each comfort noise insertion period.

When updating the comfort noise parameters (frame energy and LPC coefficients), these parameters shall be interpolated over the SID update period to obtain smooth transitions.

7 Computational details

A low level description has been prepared in form of an ANSI C source code which is part of GSM 06.06 [5].

Annex A (informative): Change Request History

Change history					
SMG No.	**TDoc. No.**	**CR. No.**	**Section affected**	**New version**	**Subject/Comments**
SMG#15				4.1.1	ETSI Publication
SMG#20				5.1.0	Release 1996 version
SMG#27				6.0.0	Release 1997 version
SMG#29				7.0.0	Release 1998 version
				7.0.1	Version update to 7.0.1 for Publication
SMG#31				8.0.0	Release 1999 version

Change history							
Date	**TSG #**	**TSG Doc.**	**CR**	**Rev**	**Subject/Comment**	**Old**	**New**
03-2001	11				Version for Release 4		4.0.0
06-2002	16				Version for Release 5	4.0.0	5.0.0
12-2004	26				Version for Release 6	5.0.0	6.0.0
06-2007	36				Version for Release 7	6.0.0	7.0.0
12-2008	42				Version for Release 8	7.0.0	8.0.0
12-2009	46				Version for Release 9	8.0.0	9.0.0

History

Document history		
V9.0.0	February 2010	Publication