

ETSI TS 146 020 V17.0.0 (2022-05)



**Digital cellular telecommunications system (Phase 2+) (GSM);
Half rate speech;
Half rate speech transcoding
(3GPP TS 46.020 version 17.0.0 Release 17)**

[ETSI TS 146 020 V17.0.0 \(2022-05\)](https://standards.iteh.ai/catalog/standards/sist/dbec5529-7c30-4dd8-acf9-f4c1f4699b5c/etsi-ts-146-020-v17-0-0-2022-05)

<https://standards.iteh.ai/catalog/standards/sist/dbec5529-7c30-4dd8-acf9-f4c1f4699b5c/etsi-ts-146-020-v17-0-0-2022-05>



Reference

RTS/TSGS-0446020vh00

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 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

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

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommitteeSupportStaff.aspx>

If you find a security vulnerability in the present document, please report it through our

Coordinated Vulnerability Disclosure Program:

<https://www.etsi.org/standards/coordinated-vulnerability-disclosure>

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2022.
All rights reserved.

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are 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 (<https://ipr.etsi.org/>).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, 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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

(standards.iteh.ai)

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. These shall be interpreted as being references to the corresponding ETSI deliverables.

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

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights	2
Legal Notice	2
Modal verbs terminology.....	2
Foreword.....	5
1 Scope	6
2 References	6
3 Definitions, symbols and abbreviations	6
3.1 Definitions	6
3.2 Symbols.....	8
3.3 Abbreviations	9
4 Functional description of the GSM half rate speech codec	10
4.1 GSM half rate speech encoder.....	10
4.1.1 High-pass filter	12
4.1.2 Segmentation	12
4.1.3 Fixed Point Lattice Technique (FLAT)	13
4.1.4 Spectral quantization.....	14
4.1.4.1 Autocorrelation Fixed Point Lattice Technique (AFLAT)	14
4.1.5 Frame energy calculation and quantization	16
4.1.6 Soft interpolation of the spectral parameters	16
4.1.7 Spectral noise weighting filter coefficients.....	17
4.1.8 Long Term Predictor lag determination.....	18
4.1.8.1 Open loop long term search initialization	19
4.1.8.2 Open loop lag search	20
4.1.8.3 Frame lag trajectory search (Mode $\neq$ 0)	25
4.1.8.4 Voicing mode selection	27
4.1.8.5 Closed loop lag search	27
4.1.9 Harmonic noise weighting	28
4.1.10 Code search algorithm	30
4.1.10.1 Decorrelation of filtered basis vectors.....	31
4.1.10.2 Fast search technique	32
4.1.11 Multimode gain vector quantization	33
4.1.11.1 Coding GS and P0.....	33
4.2 GSM half rate speech decoder.....	36
4.2.1 Excitation generation	37
4.2.2 Adaptive pitch prefilter	37
4.2.3 Synthesis Filter	37
4.2.4 Adaptive spectral postfilter	37
4.2.5 Updating decoder states	39
5 Homing sequences.....	39
5.1 Functional description	39
5.2 Definitions	39
5.3 Encoder homing	40
5.4 Decoder homing	40
5.5 Encoder home state	40
5.6 Decoder home state	40
Annex A (normative): Codec parameter description.....	41
A.1 Codec parameter description	41
A.1.1 MODE	41
A.1.2 R0	41
A.1.3 LPC1 - LPC3	42
A.1.4 LAG_1 - LAG_4	42

A.1.5 CODE_x_1 - CODE_x_4.....42

A.1.6 GSP0_1 - GSP0_4.....42

A.2 Basic coder parameters.....42

Annex B (normative): Order of occurrence of the codec parameters over Abis43

Annex C (informative): Bibliography44

Annex D (informative): Change history45

History46

iTeh STANDARD
PREVIEW
(standards.iteh.ai)

ETSI TS 146 020 V17.0.0 (2022-05)
<https://standards.iteh.ai/catalog/standards/sist/dbec5529-7c30-4dd8-acf9-f4c1f4699b5c/etsi-ts-146-020-v17-0-0-2022-05>

Foreword

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

The present document specifies the speech codec to be used for the GSM half rate channel for 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 specifies the speech codec to be used for the GSM half rate channel. It also specifies the test methods to be used to verify that the codec implementation complies with the present document.

The requirements are mandatory for the codec to be used either in GSM Mobile Stations (MS)s or Base Station Systems (BSS)s that utilize the half rate GSM speech traffic channel.

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 06.02: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech processing functions".
- [2] GSM 06.06: "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec".
- [3] GSM 06.07: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Test sequences for the GSM half rate speech codec".

3 Definitions, symbols and abbreviations

3.1 Definitions

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

adaptive codebook: adaptive codebook is derived from the long term filter state. The lag value can be viewed as an index into the adaptive codebook.

adaptive pitch prefilter: in the GSM half rate speech decoder, this filter is applied to the excitation signal to enhance the periodicity of the reconstructed speech. Note that this is done prior to the application of the short term filter.

adaptive spectral postfilter: in the GSM half rate speech decoder, this filter is applied to the output of the short term filter to enhance the perceptual quality of the reconstructed speech.

allowable lags: set of lag values which may be coded by the GSM half rate speech encoder and transmitted to the GSM half rate speech decoder. This set contains both integer and fractional values (see table 3).

analysis window: for each frame, the short term filter coefficients are computed using the high pass filtered speech samples within the analysis window. The analysis window is 170 samples in length, and is centered about the last 100 samples in the frame.

basis vectors: set of M, M1, or M2 vectors of length Ns used to generate the VSELP codebook vectors. These vectors are not necessarily orthogonal.

closed loop lag search: process of determining the near optimal lag value from the weighted input speech and the long term filter state.

closed loop lag trajectory: for a given frame, the sequence of near optimal lag values whose elements correspond to each of the four subframes as determined by the closed loop lag search.

codebook: set of vectors used in a vector quantizer.

Codeword (OR Code): M, M1, or M2 bit symbol indicating the vector to be selected from a VSELP codebook.

Delta (LAG) code: four bit code indicating the change in lag value for a subframe relative to the previous subframe's coded lag. For frames in which the long term predictor is enabled (MODE 1, 2, or 3), the lag for subframe 1 is independently coded using eight bits, and delta codes are used for subframes 2, 3, and 4.

direct form coefficients: one of the formats for storing the short term filter parameters. All filters which are used to modify speech samples use direct form coefficients.

fractional lags: set of lag values having sub-sample resolution. Note that not every fractional lag value considered in the GSM half rate speech encoder is an allowable lag value.

frame: time interval equal to 20 ms, or 160 samples at an 8 kHz sampling rate.

harmonic noise weighting filter: this filter exploits the noise masking properties of the spectral peaks which occur at harmonics of the pitch frequency by weighting the residual error less in regions near the pitch harmonics and more in regions away from them. Note that this filter is only used when the long term filter is enabled (MODE = 1, 2 or 3).

high pass filter: this filter is used to de-emphasize the low frequency components of the input speech signal.

integer lags: set of lag values having whole sample resolution.

interpolating filter: FIR filter used to estimate sub-sample resolution samples, given an input sampled with integer sample resolution.

lag: long term filter delay. This is typically the pitch period, or a multiple or sub-multiple of it.

long term filter: this filter is used to generate the periodic component in the excitation for the current subframe. This filter is only enabled for MODE = 1, 2 or 3.

LPC coefficients: Linear Predictive Coding (LPC) coefficients is a generic descriptive term for describing the short term filter coefficients.

open loop lag search: process of estimating the near optimal lag directly from the weighted speech input. This is done to narrow the range of lag values over which the closed loop lag search shall be performed.

open loop lag trajectory: for a given frame, the sequence of near optimal lag values whose elements correspond to the four subframes as determined by the open loop lag search.

reflection coefficients: alternative representation of the information contained in the short term filter parameters.

residual: output signal resulting from an inverse filtering operation.

short term filter: this filter introduces, into the excitation signal, short term correlation which models the impulse response of the vocal tract.

soft interpolation: process wherein a decision is made for each frame to use either interpolated or uninterpolated short term filter parameters for the four subframes in that frame.

soft interpolation bit: one bit code indicating whether or not interpolation of the short term parameters is to be used in the current frame.

spectral noise weighting filter: this filter exploits the noise masking properties of the formants (vocal tract resonances) by weighting the residual error less in regions near the formant frequencies and more in regions away from them.

subframe: time interval equal to 5 ms, or 40 samples at an 8 kHz sampling rate.

vector quantization: method of grouping several parameters into a vector and quantizing them simultaneously.

GSP0 vector quantizer: process of vector quantization, its intermediate parameters (GS and P0) for the coding of the excitation gains β and γ .

VSELP codebook: Vector-Sum Excited Linear Predictive (VSELP) codebook, used in the GSM half rate speech coder, wherein each codebook vector is constructed as a linear combination of the fixed basis vectors.

zero input response: output of a filter due to all past inputs, i.e. due to the present state of the filter, given that an input of zeros is applied.

zero state response: output of a filter due to the present input, given that no past inputs have been applied, i.e. given the state information in the filter is all zeroes.

3.2 Symbols

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

$A(z)$	Short term spectral filter.
α_i	The LPC coefficients.
$b_L(n)$	The output of the long term filter state (adaptive codebook) for lag L.
β	The long term filter coefficient.
$C(z)$	Second weighting filter.
$e(n)$	Weighted error signal
$f_j(i)$	The coefficients of the j^{th} phase of the 10th order interpolating filter used to evaluate candidate fractional lag values; i ranges from 0 to P_f-1 .
$g_j(i)$	The coefficients of the j^{th} phase of the 6th order interpolating filter used to interpolate C's and G's as well as fractional lags in the harmonic noise weighting; i ranges from 0 to P_g-1 .
γ	The gain applied to the vector(s) selected from the VSELP codebook(s).
H	A M2 bit code indicating the vector to be selected from the second VSELP codebook (when operating in mode 0).
I	A M or M1 bit code indicating the vector to be selected from one of the two first VSELP codebooks.
L	The long term filter lag value.
$L_{\max}$	142 (samples), the maximum possible value for the long term filter lag.
$L_{\min}$	21 (samples), the minimum possible value for the long term filter lag.
M	9, the number of basis vectors, and the number of bits in a codeword, for the VSELP codebook used in modes 1, 2, and 3.
M1	7, the number of basis vectors, and the number of bits in a codeword, for the first VSELP codebook used in mode 0.
M2	7, the number of basis vectors, and the number of bits in a codeword, for the second VSELP codebook used in mode 0.
MODE	A two bit code indicating the mode for the current frame (see annex A).
N_A	170, the length of the analysis window. This is the number of high pass filtered speech samples used to compute the short term filter parameters for each frame.
N_F	160, the number of samples per frame (at a sampling rate of 8 kHz).
N_p	10, the short term filter order.
N_s	40, the number of samples per subframe (at a sampling rate of 8 kHz).
P1	6, the number of bits in the prequantizer for the r1 - r3 vector quantizer.
P2	5, the number of bits in the prequantizer for the r4 - r6 vector quantizer.
P3	4, the number of bits in the prequantizer for the r7 - r10 vector quantizer.
P_f	The order of one phase of an interpolating filter used to evaluate candidate fractional lag values. P_f equals 10 for $j \neq 0$ and equal to 1 for $j = 0$.
P_g	The order of one phase of an interpolating filter, $f_j(n)$, used to interpolate C's and G's as well as fractional lags in the harmonic noise weighting, P_g equals 6.
pitch	The time duration between the glottal pulses which result when the vocal chords vibrate during speech production.
Q1	11, the number of bits in the r1 - r3 reflection coefficient vector quantizer.
Q2	9, the number of bits in the r4 - r6 reflection coefficient vector quantizer.
Q3	8, the number of bits in the r7 - r10 reflection coefficient vector quantizer.
R0	A five bit code used to indicate the energy level in the current frame.
$r(n)$	The long term filter state (the history of the excitation signal); $n < 0$
$r_L(n)$	The long term filter state with the adaptive codebook output for lag L appended.

$s'(n)$	Synthesized speech.
$W(z)$	Spectral weighting filter.
λ_{hnw}	The harmonic noise weighting filter coefficient.
ξ	The adaptive pitch prefilter coefficient.
$\lceil x \rceil$	Ceiling function: the largest integer y where $y < x + 1, 0$.
$\lfloor x \rfloor$	Floor function: the largest integer y where $y \leq x$.
$\sum_{i=j}^K x(i)$	Summation: $x(j)+x(j+1)+\dots+x(K)$.
$\prod_{i=j}^K x(i)$	Product: $x(j)(x(j+1))\dots(x(K))$
$\max(x,y)$	Find the larger of two numbers x and y .
$\min(x,y)$	Find the smaller of two numbers x and y .
$\text{round}(x)$	Round the non-integer x to the closest integer y : $y = \lfloor x + 0,5 \rfloor$; $y = x + 0,5$.

3.3 Abbreviations

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

AFLAT	Autocorrelation Fixed point Lattice Technique
CELP	Code Excited Linear Prediction
FLAT	Fixed Point Lattice Technique
LTP	Long Term Predictor
SST	Spectral Smoothing Technique
VSELP	Vector-Sum Excited Linear Prediction

(standards.iteh.ai)

ETSI TS 146 020 V17.0.0 (2022-05)

<https://standards.iteh.ai/catalog/standards/sist/dbec5529-7c30-4dd8-acf9-f4c1f4699b5c/etsi-ts-146-020-v17-0-0-2022-05>

4 Functional description of the GSM half rate speech codec

The GSM half rate codec uses the VSELP (Vector-Sum Excited Linear Prediction) algorithm. The VSELP algorithm is an analysis-by-synthesis coding technique and belongs to the class of speech coding algorithms known as CELP (Code Excited Linear Prediction).

The GSM half rate codec's encoding process is performed on a 20 ms speech frame at a time. A speech frame of the sampled speech waveform is read and based on the current waveform and the past history of the waveform, the codec encoder derives 18 parameters that describe it. The parameters extracted are grouped into the following three general classes:

- energy parameters (R0 and GSP0);
- spectral parameters (LPC and INT_LPC);
- excitation parameters (LAG and CODE).

These parameters are quantized into 112 bits for transmission as described in annex A and their order of occurrence over Abis is given in annex B.

The GSM half rate codec is an analysis-by-synthesis codec, therefore the speech decoder is primarily a subset of the speech encoder. The quantized parameters are decoded and a synthetic excitation is generated using the energy and excitation parameters. The synthetic excitation is then filtered to provide the spectral information resulting in the generation of the synthesized speech (see figure 1).

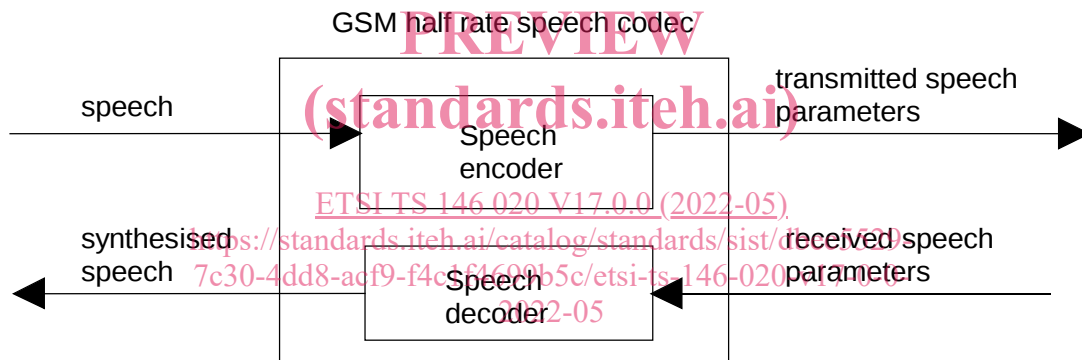


Figure 1: Block diagram of the GSM half rate speech codec

The ANSI-C code that describes the GSM half rate speech codec is given in GSM 06.06 [2] and the test sequences in GSM 06.07 [3] (see clause 5 for the codec homing test sequences).

4.1 GSM half rate speech encoder

The GSM half rate speech encoder uses an analysis by synthesis approach to determine the code to use to represent the excitation for each subframe. The codebook search procedure consists of trying each codevector as a possible excitation for the Code Excited Linear Predictive (CELP) synthesizer. The synthesized speech $s'(n)$ is compared against the input speech and a difference signal is generated. This difference signal is then filtered by a spectral weighting filter, $W(z)$, (and possibly a second weighting filter, $C(z)$) to generate a weighted error signal, $e(n)$. The power in $e(n)$ is computed. The codevector which generates the minimum weighted error power is chosen as the codevector for that subframe. The spectral weighting filter serves to weight the error spectrum based on perceptual considerations. This weighting filter is a function of the speech spectrum and can be expressed in terms of the α parameters of the short term (spectral) filter.

$$W(z) = \frac{1 - \sum_{i=1}^{N_p} \alpha_i z^{-i}}{1 - \sum_{i=1}^{N_p} \tilde{\alpha}_i z^{-i}} \quad (1)$$

The computation of the $\tilde{\alpha}_i$ coefficients is described in subclause 4.1.7.

The second weighting filter $C(z)$, if used, is a harmonic weighting filter and is used to control the amount of error in the harmonics of the speech signal. If the weighting filter(s) are moved to both input paths to the subtracter, an equivalent configuration is obtained as shown in figure 2.

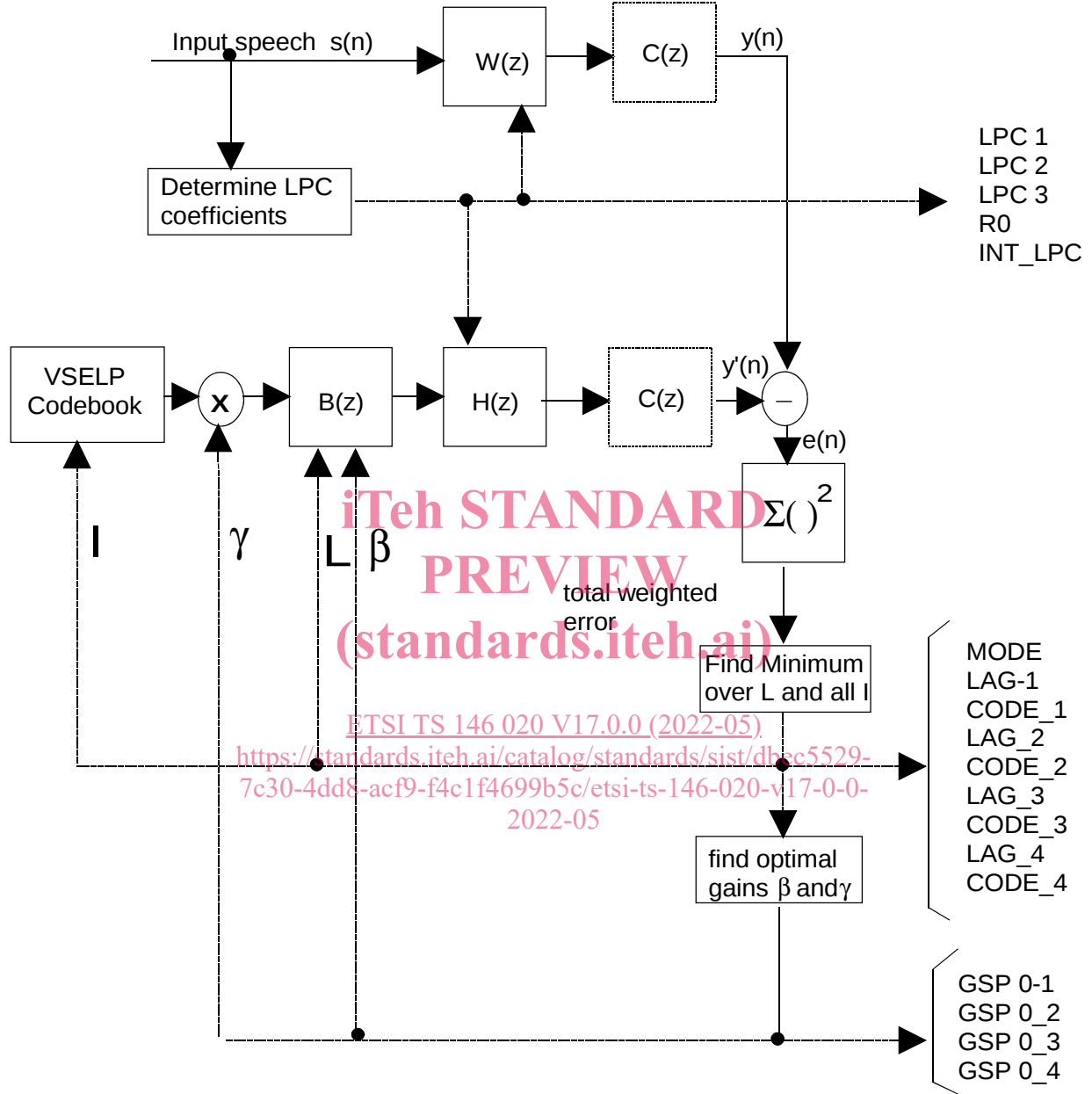


Figure 2: Block diagram of the GSM half rate speech encoder (MODE = 1,2 and 3)

Here $H(z)$ is the combination of $A(z)$, the short term (spectral) filter, and $W(z)$, the spectral weighting filter. These filters are combined since the denominator of $A(z)$ is cancelled by the numerator of $W(z)$.

$$H(z) = \frac{1}{1 - \sum_{i=1}^{N_p} \tilde{\alpha}_i z^{-i}} \quad (2)$$

There are two approaches that can be used for calculating the gain, γ . The gain can be determined prior to codebook search based on residual energy. This gain would then be fixed for the codebook search. Another approach is to optimize the gain for each codevector during the codebook search. The codevector which yields the minimum weighted error would be chosen and its corresponding optimal gain would be used for γ . The latter approach generally yields better results since the gain is optimized for each codevector. This approach also implies that the gain term needs to be updated at the subframe rate. The optimal code and gain for this technique can be computed as follows.

The input speech is first filtered by a high pass filter as described in subclause 4.1.1. The short term filter parameters are computed from the filtered input speech once per frame. A fast fixed point covariance lattice technique is used. Subclauses 4.1.3 and 4.1.4 describes in detail how the short term parameters are determined and quantized. An overall frame energy is also computed and coded once per frame. Once per frame, one of the four voicing modes is selected. If $\text{MODE} \neq 0$, the long term predictor is used and the long term predictor lag, L , is updated at the subframe rate. L and a VSELP codeword are selected sequentially. Each is chosen to minimize the weighted mean square error. The long-term filter coefficient, β , and the codebook gain, γ , are optimized jointly. Subclause 4.1.8 describes the technique for selecting from among the voicing modes and, if one of voiced modes is chosen, determining the long-term filter lag. Subclause 4.1.10 describes an efficient technique for jointly optimizing β , γ and the codeword selection. Subclause 4.1.10 also includes the description of the fast VSELP codebook search technique. The β and γ parameters are transformed to equivalent parameters using the frame energy term, and are vector quantized every subframe. The coding of the frame energy and the β and γ parameters is described in subclause 4.1.11.

4.1.1 High-pass filter

The 13 bit linear Pulse Code Modulated (PCM) input speech, $x(n)$, is filtered by a fourth order pole-zero high pass filter. This filter suppresses the frequency components of the input speech which are below 120 Hz. The filter is implemented as a cascade of two second-order Infinite Impulse Response (IIR) filters. Incorporated into the filter coefficients is a gain of 0,5. The difference equation for the first filter is:

$$\tilde{y}(n) = \sum_{i=0}^2 b_{1,i} x(n-i) + \sum_{j=1}^2 a_{1,j} \tilde{y}(n-j) \quad (3)$$

where:

$$\begin{aligned} b_{10} &= 0,335052 & b_{11} &= -0,669983 & b_{12} &= 0,335052 \\ a_{11} &= 0,926117 & a_{12} &= -0,429413 \end{aligned}$$

The difference equation for the second filter is:

$$y(n) = \sum_{i=0}^2 b_{2,i} \tilde{y}(n-i) + \sum_{j=1}^2 a_{2,j} y(n-j) \quad (4)$$

where:

$$\begin{aligned} b_{20} &= 0,335052 & a_{21} &= 0,965332 \\ b_{21} &= -0,669434 & a_{22} &= -0,469513 \\ b_{22} &= 0,335052 \end{aligned}$$

4.1.2 Segmentation

A sample buffer containing the previous 195 input high pass filtered speech samples, $y(n)$, is shifted so that the oldest 160 samples are shifted out while the next 160 input samples are shifted in. The oldest 160 samples in the buffer correspond to the next frame of samples to be encoded. The analysis interval comprises the most recent 170 samples in the buffer. The samples in the buffer are labelled as $s(n)$ where $0 \leq n \leq 194$ and $s(0)$ is the first (oldest) sample.