

ETSI TS 126 445 V12.4.0 (2015-10)



**Universal Mobile Telecommunications System (UMTS);
LTE;
Codec for Enhanced Voice Services (EVS);
Detailed algorithmic description
(3GPP TS 26.445 version 12.4.0 Release 12)**



Reference

RTS/TSGS-0426445vc40

Keywords

LTE,UMTS

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

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 only prevailing document is the print of the Portable Document Format (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:
<https://portal.etsi.org/People/CommitteeSupportStaff.aspx>

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.

© European Telecommunications Standards Institute 2015.
All rights reserved.

DECT™, PLUGTESTS™, UMTS™ and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and LTE™ are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

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://ipr.etsi.org>).

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>.

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
Foreword.....	2
Modal verbs terminology.....	2
Foreword.....	14
1 Scope	15
2 References	15
3 Definitions, abbreviations and mathematical expressions.....	17
3.1 Definitions	17
3.2 Abbreviations	17
3.3 Mathematical Expressions.....	19
4 General description of the coder	20
4.1 Introduction	20
4.2 Input/output sampling rate.....	20
4.3 Codec delay	20
4.4 Coder overview	20
4.4.1 Encoder overview	21
4.4.1.1 Linear Prediction Based Operation	21
4.4.1.2 Frequency Domain Operation	22
4.4.1.3 Inactive Signal coding.....	22
4.4.1.4 Source Controlled VBR Coding.....	23
4.4.2 Decoder overview	23
4.4.2.1 Parametric Signal Representation Decoding (Bandwidth Extension)	23
4.4.2.2 Frame loss concealment	23
4.4.3 DTX/CNG operation.....	23
4.4.3.1 Inactive Signal coding.....	23
4.4.4 AMR-WB-interoperable option	24
4.4.5 Channel-Aware Mode.....	24
4.5 Organization of the rest of the Technical Standard	24
5 Functional description of the encoder	25
5.1 Common processing.....	25
5.1.1 High-pass Filtering	25
5.1.2 Complex low-delay filter bank analysis.....	25
5.1.2.1 Sub-band analysis	25
5.1.2.2 Sub-band energy estimation	26
5.1.3 Sample rate conversion to 12.8 kHz	27
5.1.3.1 Conversion of 16, 32 and 48 kHz signals to 12.8 kHz	27
5.1.3.2 Conversion of 8 kHz signals to 12.8 kHz.....	27
5.1.3.3 Conversion of input signals to 16, 25.6 and 32 kHz	29
5.1.4 Pre-emphasis.....	29
5.1.5 Spectral analysis	30
5.1.5.1 Windowing and DFT.....	30
5.1.5.2 Energy calculations	31
5.1.6 Bandwidth detection	32
5.1.6.1 Mean and maximum energy values per band	32
5.1.7 Bandwidth decision.....	34
5.1.8 Time-domain transient detection	37
5.1.9 Linear prediction analysis.....	38
5.1.9.1 LP analysis window	38
5.1.9.2 Autocorrelation computation.....	38
5.1.9.3 Adaptive lag windowing	39
5.1.9.4 Levinson-Durbin algorithm.....	39
5.1.9.5 Conversion of LP coefficients to LSP parameters	40
5.1.9.6 LSP interpolation	41

5.1.9.7	Conversion of LSP parameters to LP coefficients	41
5.1.9.8	LP analysis at 16kHz.....	42
5.1.10	Open-loop pitch analysis	43
5.1.10.1	Perceptual weighting	43
5.1.10.2	Correlation function computation	44
5.1.10.3	Correlation reinforcement with past pitch values.....	45
5.1.10.4	Normalized correlation computation.....	46
5.1.10.5	Correlation reinforcement with pitch lag multiples.....	46
5.1.10.6	Initial pitch lag determination and reinforcement based on pitch coherence with other half-frames.....	47
5.1.10.7	Pitch lag determination and parameter update	48
5.1.10.8	Correction of very short and stable open-loop pitch estimates	49
5.1.10.9	Fractional open-loop pitch estimate for each subframe.....	51
5.1.11	Background noise energy estimation	52
5.1.11.1	First stage of noise energy update	52
5.1.11.2	Second stage of noise energy update.....	54
5.1.11.2.1	Basic parameters for noise energy update	54
5.1.11.2.2	Spectral diversity	55
5.1.11.2.3	Complementary non-stationarity	55
5.1.11.2.4	HF energy content	56
5.1.11.2.5	Tonal stability.....	56
5.1.11.2.6	High frequency dynamic range.....	60
5.1.11.2.7	Combined decision for background noise energy update	60
5.1.11.3	Energy-based parameters for noise energy update	62
5.1.11.3.1	Closeness to current background estimate.....	62
5.1.11.3.2	Features related to last correlation or harmonic event	62
5.1.11.3.3	Energy-based pause detection	63
5.1.11.3.4	Long-term linear prediction efficiency	63
5.1.11.3.5	Additional long-term parameters used for noise estimation	64
5.1.11.4	Decision logic for noise energy update	65
5.1.12	Signal activity detection.....	68
5.1.12.1	SAD1 module.....	69
5.1.12.1.1	SNR outlier filtering	70
5.1.12.2	SAD2 module.....	72
5.1.12.3	Combined decision of SAD1 and SAD2 modules for WB and SWB signals	75
5.1.12.4	Final decision of the SAD1 module for NB signals	75
5.1.12.5	Post-decision parameter update.....	76
5.1.12.6	SAD3 module.....	77
5.1.12.6.1	Sub-band FFT.....	77
5.1.12.6.2	Computation of signal features.....	78
5.1.12.6.3	Computation of SNR parameters.....	81
5.1.12.6.4	Decision of background music	82
5.1.12.6.5	Decision of background update flag.....	83
5.1.12.6.6	SAD3 Pre-decision	84
5.1.12.6.7	SAD3 Hangover	86
5.1.12.7	Final SAD decision	86
5.1.12.8	DTX hangover addition.....	88
5.1.13	Coding mode determination.....	90
5.1.13.1	Unvoiced signal classification.....	91
5.1.13.1.1	Voicing measure	92
5.1.13.1.2	Spectral tilt	92
5.1.13.1.3	Sudden energy increase from a low energy level	93
5.1.13.1.4	Total frame energy difference	94
5.1.13.1.5	Energy decrease after spike	94
5.1.13.1.6	Decision about UC mode.....	94
5.1.13.2	Stable voiced signal classification.....	96
5.1.13.3	Signal classification for FEC.....	96
5.1.13.3.1	Signal classes for FEC.....	97
5.1.13.3.2	Signal classification parameters	97
5.1.13.3.3	Classification procedure	98
5.1.13.4	Transient signal classification	99
5.1.13.5	Modification of coding mode in special cases	100

5.1.13.6	Speech/music classification.....	101
5.1.13.6.1	First stage of the speech/music classifier.....	101
5.1.13.6.2	Scaling of features in the first stage of the speech/music classifier.....	103
5.1.13.6.3	Log-probability and decision smoothing.....	104
5.1.13.6.4	State machine and final speech/music decision.....	105
5.1.13.6.5	Improvement of the classification for mixed and music content.....	108
5.1.13.6.6	Second stage of the speech/music classifier.....	112
5.1.13.6.7	Context-based improvement of the classification for stable tonal signals.....	114
5.1.13.6.8	Detection of sparse spectral content.....	118
5.1.13.6.9	Decision about AC mode.....	120
5.1.13.6.10	Decision about IC mode.....	120
5.1.14	Coder technology selection.....	120
5.1.14.1	ACELP/MDCT-based technology selection at 9.6kbps, 16.4 and 24.4 kbps.....	121
5.1.14.1.1	Segmental SNR estimation of the MDCT-based technology.....	121
5.1.14.1.2	Segmental SNR estimation of the ACELP technology.....	127
5.1.14.1.3	Hysteresis and final decision.....	128
5.1.14.2	TCX/HQ MDCT technology selection at 13.2 and 16.4 kbps.....	129
5.1.14.3	TCX/HQ MDCT technology selection at 24.4 and 32 kbps.....	131
5.1.14.4	TD/Multi-mode FD BWE technology selection at 13.2 kbps and 32 kbps.....	134
5.2	LP-based Coding.....	135
5.2.1	Perceptual weighting.....	135
5.2.2	LP filter coding and interpolation.....	136
5.2.2.1	LSF quantization.....	136
5.2.2.1.1	LSF weighting function.....	136
5.2.2.1.2	Bit allocation.....	139
5.2.2.1.3	Predictor allocation.....	140
5.2.2.1.4	LSF quantizer structure.....	140
5.2.2.1.5	LSFQ for voiced coding mode at 16 kHz internal sampling frequency : BC-TCVQ.....	145
5.2.2.1.6	Mid-frame LSF quantizer.....	152
5.2.3	Excitation coding.....	153
5.2.3.1	Excitation coding in the GC, VC and high rate IC/UC modes.....	153
5.2.3.1.1	Computation of the LP residual signal.....	154
5.2.3.1.2	Target signal computation.....	154
5.2.3.1.3	Impulse response computation.....	155
5.2.3.1.4	Adaptive codebook.....	155
5.2.3.1.5	Algebraic codebook.....	157
5.2.3.1.6	Combined algebraic codebook.....	167
5.2.3.1.7	Gain quantization.....	181
5.2.3.2	Excitation coding in TC mode.....	186
5.2.3.2.1	Glottal pulse codebook search.....	186
5.2.3.2.2	TC frame configurations.....	190
5.2.3.2.4	Pitch period and gain coding in the TC mode.....	192
5.2.3.2.5	Update of filter memories.....	195
5.2.3.3	Excitation coding in UC mode at low rates.....	195
5.2.3.3.1	Structure of the Gaussian codebook.....	195
5.2.3.3.2	Correction of the Gaussian codebook spectral tilt.....	196
5.2.3.3.3	Search of the Gaussian codebook.....	197
5.2.3.3.4	Quantization of the Gaussian codevector gain.....	198
5.2.3.3.5	Other parameters in UC mode.....	199
5.2.3.4	Excitation coding in IC and UC modes at 9.6 kbps.....	199
5.2.3.4.2	Gaussian noise generation.....	201
5.2.3.4.3	Gain coding.....	201
5.2.3.4.4	Memory update.....	203
5.2.3.5	Excitation coding in GSC mode.....	203
5.2.3.5.1	Determining the subframe length.....	204
5.2.3.5.2	Computing time-domain excitation contribution.....	204
5.2.3.5.3	Frequency transform of residual and time-domain excitation contribution.....	205
5.2.3.5.4	Computing energy dynamics of transformed residual and quantization of noise level.....	206
5.2.3.5.6	Find and encode the cut-off frequency.....	206
5.2.3.5.7	Band energy computation and quantization.....	208
5.2.3.5.8	PVQ Bit allocation.....	208
5.2.3.5.9	Quantization of difference signal.....	209

5.2.3.5.10	Spectral dynamic and noise filling	209
5.2.3.5.11	Quantized gain addition, temporal and frequency contributions combination	209
5.2.3.5.12	Specifics for wideband 8kbps.....	210
5.2.3.5.13	Inverse DCT	211
5.2.3.5.14	Remove pre-echo in case of onset detection.....	211
5.2.4	Bass post-filter gain quantization.....	212
5.2.5	Source Controlled VBR Coding	212
5.2.5.1	Principles of VBR Coding	212
5.2.5.2	EVS VBR Encoder Coding Modes and Bit-Rates	213
5.2.5.3	Prototype-Pitch-Period (PPP) Encoding	213
5.2.5.3.1	PPP Algorithm.....	213
5.2.5.3.2	Amplitude Quantization	214
5.2.5.3.3	Phase Quantization	215
5.2.5.4	Noise-Excited-Linear-Prediction (NELP) Encoding.....	215
5.2.5.5	Average Data Rate (ADR) Control for the EVS VBR mode	215
5.2.6	Coding of upper band for LP-based Coding Modes	219
5.2.6.1	Bandwidth extension in time domain.....	219
5.2.6.1.1	High band target signal generation	220
5.2.6.1.2	TBE LP analysis	221
5.2.6.1.3	Quantization of linear prediction parameters.....	223
5.2.6.1.4	Interpolation of LSF coefficients.....	226
5.2.6.1.5	Target and residual energy calculation and quantization.....	228
5.2.6.1.6	Generation of the upsampled version of the lowband excitation.....	228
5.2.6.1.7	Non-Linear Excitation Generation	229
5.2.6.1.8	Spectral flip of non-linear excitation in time domain	230
5.2.6.1.9	Down-sample using all-pass filters.....	230
5.2.6.1.10	Adaptive spectral whitening	231
5.2.6.1.11	Envelope modulated noise mixing.....	231
5.2.6.1.12	Spectral shaping of the noise added excitation.....	233
5.2.6.1.13	Post processing of the shaped excitation	233
5.2.6.1.14	Estimation of temporal gain shape parameters	235
5.2.6.1.15	Estimation of frame gain parameters.....	238
5.2.6.1.16	Estimation of TEC/TFA envelope parameters.....	240
5.2.6.1.17	Estimation of full-band frame energy parameters	242
5.2.6.2	Multi-mode FD Bandwidth Extension Coding	244
5.2.6.2.1	SWB/FB Multi-mode FD Bandwidth Extension.....	245
5.2.6.2.2	WB Multi-mode FD Bandwidth Extension	256
5.2.6.3	Coding of upper band at 64 kb/s	260
5.2.6.3.1	Coding in normal mode	261
5.2.6.3.2	Coding in transient mode.....	265
5.3	MDCT Coding Mode	268
5.3.1	General description	268
5.3.2	Time-to-frequency transformations	268
5.3.2.1	Transform sizes and MDCT configurations.....	268
5.3.2.2	Long block transformation (ALDO window).....	268
5.3.2.2.1	Folding and on-the-fly window decimation.....	270
5.3.2.2.2	eDCT	272
5.3.2.3	Transient location dependent overlap and transform length	274
5.3.2.4	Short block transformation.....	275
5.3.2.4.2	Short window transform for MDCT based TCX.....	279
5.3.2.5	Special window transitions	280
5.3.2.5.1	ALDO to short transition.....	280
5.3.2.5.2	Short to ALDO transition	280
5.3.2.6	Modified Discrete Sine Transform.....	280
5.3.3	MDCT based TCX.....	281
5.3.3.1	General description	281
5.3.3.1.1	High level overview	281
5.3.3.1.2	Rate dependent configuration.....	281
5.3.3.2	General encoding procedure.....	282
5.3.3.2.1	LPC parameter calculation	282
5.3.3.2.2	Temporal Noise Shaping	286
5.3.3.2.3	LPC shaping in MDCT domain.....	288

5.3.3.2.4	Adaptive low frequency emphasis.....	289
5.3.3.2.5	Spectrum noise measure in power spectrum.....	290
5.3.3.2.6	Low pass factor detector.....	290
5.3.3.2.7	Uniform quantizer with adaptive dead-zone.....	291
5.3.3.2.8	Arithmetic coder.....	291
5.3.3.2.9	Global gain coding.....	303
5.3.3.2.10	Noise Filling.....	304
5.3.3.2.11	Intelligent Gap Filling.....	307
5.3.3.2.12	Memory updates.....	322
5.3.3.2.13	Global Gain Adjuster.....	322
5.3.4	High Quality MDCT coder (HQ).....	323
5.3.4.1	Low-rate HQ coder.....	323
5.3.4.1.1	Tonality Estimation.....	324
5.3.4.1.2	Grouping of spectral coefficients.....	324
5.3.4.1.3	Energy Envelope coding.....	329
5.3.4.1.4	MDCT coefficients quantization.....	335
5.3.4.2	High-rate HQ coder.....	371
5.3.4.2.1	Normal Mode.....	373
5.3.4.2.2	Transient Mode.....	383
5.3.4.2.3	Generic, Harmonic and HVQ mode detector.....	384
5.3.4.2.4	Harmonic Mode.....	385
5.3.4.2.5	HVQ.....	386
5.3.4.2.6	Generic Mode.....	389
5.3.4.2.7	Pyramid Vector Quantization (PVQ) and indexing.....	397
5.4	Switching of Coding Modes.....	407
5.4.1	General description.....	407
5.4.2	MDCT coding mode to CELP coding mode.....	407
5.4.2.1	MDCT to CELP transition 1 (MC1).....	408
5.4.2.2	MDCT to CELP transition 2 (MC2).....	408
5.4.2.3	MDCT to CELP transition 3 (MC3).....	408
5.4.3	CELP coding mode to MDCT coding mode.....	408
5.4.3.1	CELP coding mode to MDCT based TCX coding mode.....	408
5.4.3.2	CELP coding mode to HQ MDCT coding mode.....	409
5.4.3.2.1	Constrained CELP coding and simplified BWE coding.....	409
5.4.3.2.2	HQ MDCT coding with a modified analysis window.....	410
5.4.4	Internal sampling rate switching.....	411
5.4.4.1	Reset of LPC memory.....	411
5.4.4.2	Conversion of LP filter between 12.8 and 16 kHz internal sampling rates.....	411
5.5.4.1.1	Modification of the Power Spectrum.....	411
5.5.4.1.2	Computation of the Power Spectrum.....	412
5.5.4.1.3	Computation of the Autocorrelation.....	413
5.4.4.3	Extrapolation of LP filter.....	414
5.4.4.4	Buffer resampling with linear interpolation.....	414
5.4.4.5	Update of CELP input signal memories.....	414
5.4.4.6	Update of MDCT-based TCX input signal memories.....	415
5.4.4.7	Update of CELP synthesis memories.....	415
5.4.5	EVS primary and AMR-WB IO.....	415
5.4.5.1	Switching from primary modes to AMR-WB IO.....	415
5.4.5.2	Switching from AMR-WB IO mode to primary modes.....	416
5.4.6	Rate switching.....	416
5.4.6.1	Rate switching along with internal sampling rate switching.....	416
5.4.6.2	Rate switching along with coding mode switching.....	416
5.5	Frame erasure concealment side information.....	417
5.5.1	Signal classification parameter.....	417
5.5.2	Energy information.....	417
5.5.3	Phase control information.....	418
5.5.4	Pitch lag information.....	419
5.5.5	Spectral envelope diffuser.....	419
5.5.6	Tonality flag information.....	420
5.6	DTX/CNG operation.....	421
5.6.1	Overview.....	421
5.6.1.1	SID update.....	422

5.6.1.2	Spectral tilt based SID transmission.....	423
5.6.1.3	CNG selector.....	424
5.6.2	Encoding for LP-CNG.....	424
5.6.2.1	LP-CNG CN parameters estimation.....	425
5.6.2.1.1	LP-CNG Hangover analysis period determination.....	425
5.6.2.1.2	LP-CNG filter parameters evaluation for low-band signal.....	425
5.6.2.1.3	LP-CNG CNG-LSF quantization for low-band signal.....	426
5.6.2.1.4	LP-CNG synthesis filter computation for local CNG synthesis.....	427
5.6.2.1.5	LP-CNG energy calculation and quantization.....	428
5.6.2.1.6	LP-CNG energy smoothing for local CNG synthesis.....	429
5.6.2.1.7	LP-CNG LF-BOOST determination and quantization.....	429
5.6.2.1.8	LP-CNG high band analysis and quantization.....	430
5.6.2.2	LP-CNG local CNG synthesis.....	432
5.6.2.3	LP-CNG CNG Memory update.....	432
5.6.3	Encoding for FD-CNG.....	432
5.6.3.1	Spectral partition energies.....	433
5.6.3.1.1	Computation of the FFT partition energies.....	433
5.6.3.1.2	Computation of the CLDFB partition energies.....	433
5.6.3.1.3	FD-CNG configurations.....	433
5.6.3.2	FD-CNG noise estimation.....	434
5.6.3.2.1	Dynamic range compression for the input energies.....	434
5.6.3.2.2	Noise tracking.....	435
5.6.3.2.3	Dynamic range expansion for the estimated noise energies.....	439
5.6.3.3	Adjusting the first SID frame in FD-CNG.....	439
5.6.3.4	FD-CNG resetting mechanism.....	439
5.6.3.5	Encoding SID frames in FD-CNG.....	439
5.6.3.6	FD-CNG local CNG synthesis.....	441
5.6.3.6.1	SID parameters interpolation.....	442
5.6.3.6.2	LPC estimation from the interpolated SID parameters.....	442
5.6.3.6.3	FD-CNG encoder comfort noise generation.....	442
5.6.3.6.4	FD-CNG encoder memory update.....	442
5.7	AMR-WB-interoperable modes.....	443
5.7.1	Pre-processing.....	443
5.7.2	Linear prediction analysis and quantization.....	443
5.7.2.1	Windowing and auto-correlation computation.....	443
5.7.2.2	Levinson-Durbin algorithm.....	443
5.7.2.3	LP to ISP conversion.....	443
5.7.2.4	ISP to LP conversion.....	444
5.7.2.5	Quantization of the ISP coefficients.....	444
5.7.2.6	Interpolation of the ISPs.....	444
5.7.3	Perceptual weighting.....	444
5.7.4	Open-loop pitch analysis.....	444
5.7.5	Impulse response computation.....	444
5.7.6	Target signal computation.....	444
5.7.7	Adaptive codebook search.....	445
5.7.8	Algebraic codebook search.....	445
5.7.9	Quantization of the adaptive and fixed codebook gains.....	445
5.7.10	Memory update.....	445
5.7.11	High-band gain generation.....	445
5.7.12	CNG coding.....	445
5.8	Channel Aware Coding.....	445
5.8.1	Introduction.....	445
5.8.2	Principles of Channel Aware Coding.....	446
5.8.3	Bit-Rate Allocation for Primary and Partial Redundant Frame Coding.....	447
5.8.3.1	Primary frame bit-rate reduction.....	447
5.8.3.2	Partial Redundant Frame Coding.....	447
5.8.3.2.1	Construction of partial redundant frame for Generic and Voiced Coding modes.....	447
5.8.3.2.2	Construction of partial redundant frame for Unvoiced Coding mode.....	447
5.8.3.2.3	Construction of partial redundant frame for TCX frame.....	448
5.8.3.2.4	RF_NO_DATA partial redundant frame type.....	448
5.8.3.3	Decoding.....	448
5.8.4	Channel aware mode encoder configurable parameters.....	448

6	Functional description of the Decoder	450
6.1	LP-based Decoding	450
6.1.1	General LP-based decoding	450
6.1.1.1	LSF decoding	450
6.1.1.1.1	General LSF decoding	450
6.1.1.1.2	LSF decoding for voiced coding mode at 16 kHz internal sampling frequency	453
6.1.1.2	Reconstruction of the excitation	455
6.1.1.2.1	Reconstruction of the excitation in GC and VC modes and high rate IC/UC modes	455
6.1.1.2.1.8	Reconstructed excitation	462
6.1.1.2.2	Reconstruction of the excitation in TC mode	462
6.1.1.2.3	Reconstruction of the excitation in UC mode at low rates	462
6.1.1.2.4	Reconstruction of the excitation in IC/UC mode at 9.6 kbps	464
6.1.1.2.5	Reconstruction of the excitation in GSC	465
6.1.1.3	Excitation post-processing	466
6.1.1.3.1	Anti-sparseness processing	466
6.1.1.3.2	Gain smoothing for noise enhancement	466
6.1.1.3.3	Pitch enhancer	467
6.1.1.3.4	Music post processing	468
6.1.2	Source Controlled VBR decoding	475
6.1.3	Synthesis	475
6.1.4	Post-processing	476
6.1.4.1.1	Long-term post-filter	476
6.1.4.1.2	Short-term post-filter	477
6.1.4.1.3	Post-filter NB parameters	477
6.1.4.1.4	Post-filter WB and SWB parameters	478
6.1.4.1.5	Tilt compensation	478
6.1.4.1.6	Adaptive gain control	479
6.1.4.2	Bass post-filter	479
6.1.5	Decoding of upper band for LP-based Coding Modes	482
6.1.5.1	Decoding Time-domain Bandwidth Extension	482
6.1.5.1.1	Generation of the upsampled version of the lowband excitation	482
6.1.5.1.2	Non-Linear Excitation Generation	483
6.1.5.1.3	De-quantization of high band parameters	484
6.1.5.1.4	LSP interpolation	484
6.1.5.1.5	Spectral flip in time domain	484
6.1.5.1.6	Down-sample using all-pass filters	484
6.1.5.1.7	Adaptive spectral whitening	485
6.1.5.1.8	Envelope modulated noise mixing	485
6.1.5.1.9	Spectral shaping of the noise added excitation	487
6.1.5.1.10	Post processing of the shaped excitation	487
6.1.5.1.11	Gain shape update	487
6.1.5.1.12	SHB synthesis	488
6.1.5.1.13	Core-switching and high-band memory updates	489
6.1.5.1.14	TEC/TFA post processing	490
6.1.5.1.15	Full-band synthesis	492
6.1.5.2	Multi-mode FD Bandwidth Extension decoding	492
6.1.5.2.1	SWB multi-mode FD BWE decoding	492
6.1.5.2.2	WB multi-mode FD BWE decoding	500
6.1.5.3	Decoding of upper band at 64 kb/s	508
6.1.5.3.1	Decoding in normal mode	508
6.1.5.3.2	Decoding in transient mode	513
6.1.5.3.3	Windowing and frequency-to-time transformation	515
6.1.5.3.4	Post-processing in temporal domain	515
6.2	MDCT Coding mode decoding	518
6.2.1	General MDCT decoding	518
6.2.2	MDCT based TCX	518
6.2.2.1	Rate dependent configurations	518
6.2.2.2	Init module parameters	518
6.2.2.2.1	TCX block configuration	518
6.2.2.2.2	LPC parameter	518
6.2.2.2.3	PLC Waveform adjustment	519
6.2.2.2.4	Global Gain	519

6.2.2.2.5	Noise fill parameter	520
6.2.2.2.6	LTP	520
6.2.2.2.7	TNS parameter.....	520
6.2.2.2.8	Harmonic model	520
6.2.2.2.9	IGF bit stream reader.....	520
6.2.2.2.10	Spectral data	523
6.2.2.2.11	Residual bits	523
6.2.2.3	Decoding process	524
6.2.2.3.1	Arithmetic decoder	524
6.2.2.3.2	Adaptive low frequency de-emphasis.....	525
6.2.2.3.3	Global gain decoding.....	526
6.2.2.3.4	Residual bits decoding.....	526
6.2.2.3.5	TCX formant enhancement	527
6.2.2.3.6	Noise Filling	528
6.2.2.3.7	Apply global gain and LPC shaping in MDCT domain	529
6.2.2.3.8	IGF apply.....	530
6.2.2.3.9	Inverse window grouping (TCX5 separation)	536
6.2.2.3.10	Temporal Noise Shaping	536
6.2.2.3.11	IGF temporal flattening	537
6.2.3	High Quality MDCT decoder (HQ).....	537
6.2.3.1	Low-rate HQ decoder.....	537
6.2.3.1.1	Mode decoding.....	537
6.2.3.1.2	Energy Envelope decoding	537
6.2.3.1.3	Spectral coefficients decoding.....	539
6.2.3.2	High-rate HQ decoder	549
6.2.3.2.1	Normal Mode	549
6.2.3.2.2	Transient Mode.....	557
6.2.3.2.3	Harmonic Mode.....	558
6.2.3.2.4	HVQ	560
6.2.3.2.5	Generic Mode.....	561
6.2.3.2.6	PVQ decoding and de-indexing.....	566
6.2.4	Frequency-to-time transformation	568
6.2.4.1	Long block transformation (ALDO window).....	568
6.2.4.1.1	eDCT	568
6.2.4.1.2	Unfolding and windowing	568
6.2.4.1.2	Overlap-add	569
6.2.4.1.3	Pre-echo attenuation	570
6.2.4.2	Transient location dependent overlap and transform length	576
6.2.4.3	Short block transformation.....	576
6.2.4.3.1	Short window transform in TDA domain	576
6.2.4.3.2	Short window transform for MDCT based TCX.....	577
6.2.4.4	Special window transitions	577
6.2.4.4.1	ALDO to short transition	577
6.2.4.4.2	Short to ALDO transition	578
6.2.4.5	Low Rate MDCT Synthesis	578
6.3	Switching coding modes in decoding	578
6.3.1	General description	578
6.3.2	MDCT coding mode to CELP coding mode.....	579
6.3.2.1	MDCT to CELP transition 1 (MC1)	579
6.3.2.2	MDCT to CELP transition 2 (MC2)	580
6.3.2.3	MDCT to CELP transition 3 (MC3)	580
6.3.3	CELP coding mode to MDCT coding mode.....	580
6.3.3.1	CELP coding mode to MDCT based TCX coding mode	581
6.3.3.2	CELP coding mode to HQ MDCT coding mode	582
6.3.3.2.1	Constrained CELP decoding and simplified BWE decoding	583
6.3.3.2.1.1	Optimized cubic interpolation.....	583
6.3.3.2.2	HQ MDCT decoding with a modified synthesis window.....	584
6.3.3.2.3	Cross-fading	584
6.3.4	Internal sampling rate switching	584
6.3.4.1	Reset of LPC memory	584
6.3.4.2	Conversion of LP filter between 12.8 and 16 kHz internal sampling rates	585
6.3.4.3	Extrapolation of LP filter	585

6.3.4.4	Update of CELP synthesis memories	585
6.3.4.5	Update of CELP decoded past signal	585
6.3.4.6	Post-processing	585
6.3.4.6.1	Adaptive post-filtering	585
6.3.4.6.2	Bass post filter.....	585
6.3.4.7	CLDFB	585
6.3.5	EVS primary modes and AMR-WB IO	586
6.3.5.1	Switching from primary modes to AMR-WB IO	586
6.3.5.2	Switching from AMR-WB IO mode to primary modes	586
6.3.6	Rate switching	586
6.3.6.1	Rate switching along with internal sampling rate switching	586
6.3.6.2	Rate switching along with coding mode switching	586
6.3.6.3	Adaptive post-filter reset and smoothing	586
6.3.7	Bandwidth switching	586
6.3.7.1	Bandwidth switching detector.....	586
6.3.7.2	Super wideband switching to wideband.....	587
6.3.7.2.1	TBE mode.....	587
6.3.7.2.2	Multi-mode FD BWE mode	589
6.3.7.2.3	MDCT core.....	590
6.3.7.3	Wideband switching to super wideband.....	590
6.4	De-emphasis	590
6.5	Resampling to the output sampling frequency	590
6.6	Decoding of frame erasure concealment side information	590
6.7	Decoding in DTX/CNG operation.....	591
6.7.1	Overview	591
6.7.2	Decoding for LP-CNG.....	591
6.7.2.1	LP-CNG decoding Overview	591
6.7.2.1.1	CNG parameter updates in active periods	591
6.7.2.1.2	DTX-hangover based parameter analysis in LP-CNG mode.....	592
6.7.2.1.3	LP-CNG low-band energy decoding	594
6.7.2.1.4	LP-CNG low-band filter parameters decoding	594
6.7.2.1.5	LP-CNG low-band excitation generation	594
6.7.2.1.6	LP-CNG low-band synthesis	596
6.7.2.1.7	LP-CNG high-band decoding and synthesis.....	596
6.7.2.2	Memory update	598
6.7.3	Decoding for FD-CNG	598
6.7.3.1	Decoding SID frames in FD-CNG	598
6.7.3.1.1	SID parameters decoding.....	599
6.7.3.1.2	SID parameters interpolation.....	599
6.7.3.1.3	LPC estimation from the interpolated SID parameters.....	600
6.7.3.2	Noise tracking during active frames in FD-CNG.....	600
6.7.3.2.1	Spectral partition energies	600
6.7.3.2.2	FD-CNG noise estimation	601
6.7.3.2.3	Noise shaping in FD-CNG	601
6.7.3.3	Noise generation for SID or zero frames in FD-CNG.....	603
6.7.3.3.1	Update of the noise levels for FD-CNG	603
6.7.3.3.2	Comfort noise generation in the frequency domain.....	603
6.7.3.3.3	Comfort noise generation in the time domain	603
6.7.3.3.4	FD-CNG decoder memory update.....	605
6.8	AMR-WB-interoperable modes	606
6.8.1	Decoding and speech synthesis.....	606
6.8.1.1	Excitation decoding.....	606
6.8.1.2	Excitation post-processing	607
6.8.1.2.1	Anti-sparseness processing.....	607
6.8.1.2.2	Gain smoothing for noise enhancement	607
6.8.1.2.3	Pitch enhancer	607
6.8.1.3	Synthesis filtering	607
6.8.1.4	Music and Unvoiced/inactive Post-processing.....	607
6.8.1.4.1	Music post processing	607
6.8.1.4.2	Unvoiced and inactive post processing.....	608
6.8.1.5	Synthesis filtering and overwriting the current CELP synthesis	611
6.8.1.6	Formant post-filter	611

6.8.1.7	Comfort noise addition.....	611
6.8.1.8	Bass post-filter	611
6.8.2	Resampling	612
6.8.3	High frequency band.....	612
6.8.3.1	Preliminary estimation steps	612
6.8.3.1.1	Estimation of tilt, figure of merit and voice factors.....	613
6.8.3.1.2	Estimation of sub-frame gains based on LP spectral envelopes	614
6.8.3.2	Generation of high-band excitation.....	617
6.8.3.2.1	DCT	617
6.8.3.2.2	High band generation	617
6.8.3.2.3	Extraction of tonal and ambiance components	619
6.8.3.2.4	Recombination.....	620
6.8.3.2.5	Filtering in DCT domain	621
6.8.3.2.6	Inverse DCT	622
6.8.3.2.7	Gain computation and scaling of excitation	622
6.8.3.3	LP filter for the high frequency band	624
6.8.3.4	High band synthesis	624
6.8.4	CNG decoding	624
6.9	Common post-processing.....	625
6.9.1	Comfort noise addition	625
6.9.1.1	Noisy speech detection.....	625
6.9.1.2	Noise estimation for CNA.....	626
6.9.1.2.1	CNA noise estimation in DTX-on mode when FD-CNG is triggered.....	626
6.9.1.2.2	CNA noise estimation in DTX-on mode when LP-CNG is triggered	626
6.9.1.2.3	CNA noise estimation in DTX-off mode.....	626
6.9.1.3	Noise generation in the FFT domain and addition in the time domain	626
6.9.1.4	Noise generation and addition in the MDCT domain	627
6.9.2	Long term prediction processing	627
6.9.2.1	Decoding LTP parameters.....	627
6.9.2.2	LTP post filtering	628
6.9.3	Complex low delay filter bank synthesis	630
6.9.4	High pass filtering.....	631
7	Description of the transmitted parameter indices.....	631
7.1	Bit allocation for the default option.....	631
7.1.1	Bit allocation at VBR 5.9, 7.2 – 9.6 kbps	631
7.1.2	Bit allocation at 13.2 kbps	632
7.1.3	Bit allocation at 16.4 and 24.4 kbps.....	632
7.1.4	Bit allocation at 32 kbps	633
7.1.5	Bit allocation at 48, 64, 96 and 128 kbps.....	634
7.2	Bit allocation for SID frames in the DTX operation	635
7.3	Bit allocation for the AMR-WB-interoperable option.....	636
7.4	Bit Allocation for the Channel-Aware Mode	636
Annex A (normative):	RTP Payload Format and SDP Parameters	638
A.0	General	638
A.1	RTP Header Usage	638
A.2	RTP Payload Format	638
A.2.1	EVS codec Compact Format	638
A.2.1.1	Compact format for EVS Primary mode.....	639
A.2.1.2	Compact format for EVS AMR-WB IO mode (except SID).....	639
A.2.1.2.1	Representation of Codec Mode Request (CMR) in Compact format for EVS AMR-WB IO mode ...	639
A.2.1.2.2	Payload structure of Compact EVS AMR-WB IO mode frame.....	640
A.2.1.3	Special case for 56 bit payload size (EVS Primary or EVS AMR-WB IO SID).....	641
A.2.2	EVS codec Header-Full format	641
A.2.2.1	EVS RTP payload structure	641
A.2.2.1.1	CMR byte.....	642
A.2.2.1.2	ToC byte.....	644
A.2.2.1.3	Speech Data.....	645
A.2.2.1.4	Zero padding	646

A.2.2.1.4.1	Zero padding for octet alignment of speech data (EVS AMR-WB IO).....	646
A.2.2.1.4.2	Zero padding for size collision avoidance	646
A.2.2.1.4.3	Additional zero padding	646
A.2.3	Header-Full/Compact format handling.....	646
A.2.3.1	Default format handling	646
A.2.3.2	Header-Full-only format handling	647
A.2.4	AMR-WB backward compatible EVS AMR-WB IO mode format	647
A.2.5	Sessions with multiple mono channels.....	647
A.2.5.1	Encoding of multiple mono channels.....	647
A.2.5.2	RTP header usage	647
A.2.5.3	Construction of the RTP payload.....	647
A.2.6	Storage Format	648
A.2.6.1	Header.....	648
A.2.6.2	Speech Frames	648
A.3	Payload Format Parameters.....	649
A.3.1	EVS Media Type Registration.....	649
A.3.2	Mapping Media Type Parameters into SDP	651
A.3.3	Detailed Description of Usage of SDP Parameters	652
A.3.3.1	Offer-Answer Model Considerations.....	652
A.3.3.2	Examples	653
Annex B (informative):	Change history	655
History		656

Foreword

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

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 is a detailed description of the signal processing algorithms of the Enhanced Voice Services coder.

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] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [2] 3GPP TS 26.441: "Codec for Enhanced Voice Services (EVS); General Overview".
- [3] 3GPP TS 26.442: "Codec for Enhanced Voice Services (EVS); ANSI C code (fixed-point)".
- [4] 3GPP TS 26.444: "Codec for Enhanced Voice Services (EVS); Test Sequences".
- [5] 3GPP TS 26.446: "Codec for Enhanced Voice Services (EVS); AMR-WB Backward Compatible Functions".
- [6] 3GPP TS 26.447: "Codec for Enhanced Voice Services (EVS); Error Concealment of Lost Packets".
- [7] 3GPP TS 26.448: "Codec for Enhanced Voice Services (EVS); Jitter Buffer Management".
- [8] 3GPP TS 26.173: "ANSI-C code for the Adaptive Multi Rate - Wideband (AMR-WB) speech codec".
- [9] 3GPP TS 26.190: "Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions".
- [10] 3GPP TS 26.103: "Speech codec list for GSM and UMTS".
- [11] 3GPP TS 26.290: "Extended AMR Wideband codec; Transcoding functions".
- [12] 3GPP TS 26.401: "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description".
- [13] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction".
- [14] 3GPP TS 22.105: "Services and Service capabilities".
- [15] IETF RFC 4867 (2007): "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", J. Sjoberg, M. Westerlund, A. Lakaniemi and Q. Xie.
- [16] IEEE Transactions on Acoustic, Speech and Signals Processing, Vol. 2, April, pp. 9-12 (1993): "Immittance spectral pairs (ISP) for speech encoding", Bistriz, Y., and Peller, S.
- [17] IEEE Journal on Selected Areas in Communications, Vol. 6, No. 2, February, pp. 314-323 (1988): "Transform Coding of Audio Signals Using Perceptual Noise Criteria", Johnston, J.D.