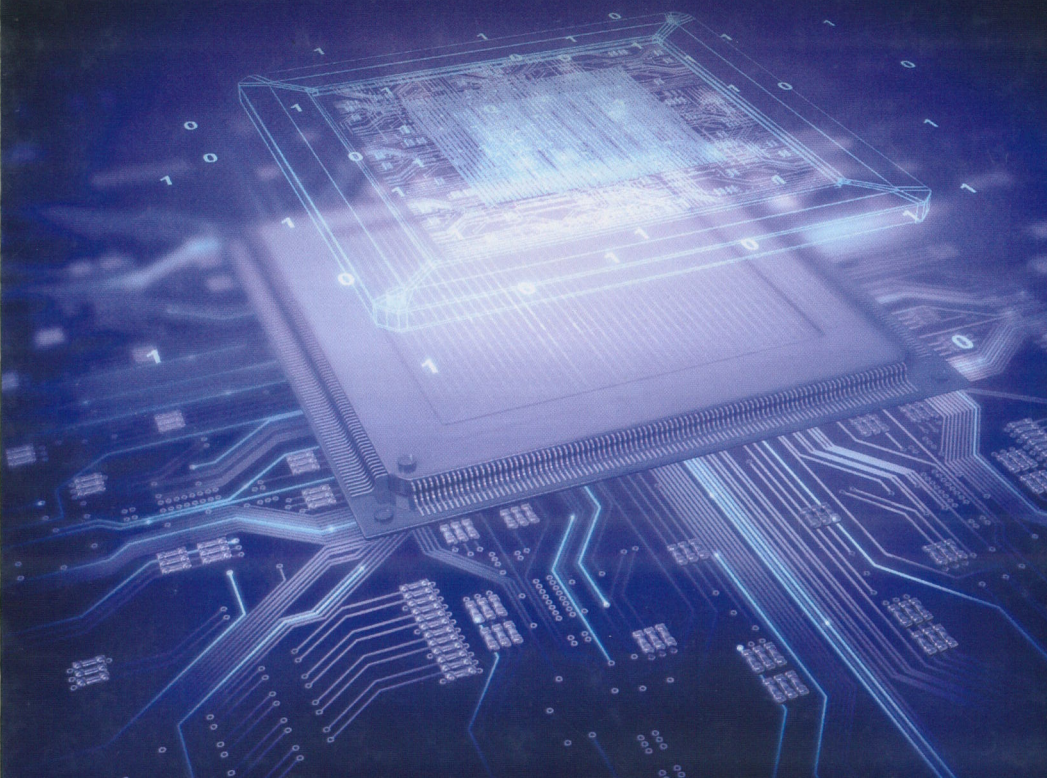




CRC Press
Taylor & Francis Group

A glowing blue microchip is shown in a 3D perspective, appearing to float above a complex printed circuit board (PCB). The chip and the board's traces are illuminated with a bright blue light, creating a futuristic, high-tech aesthetic. The background is dark, making the glowing elements stand out.

Mixed-Signal Circuits

EDITED BY

Thomas Noulis

Mixed-Signal Circuits

Mixed-Signal Circuits offers a thoroughly modern treatment of integrated circuit design in the context of mixed-signal applications. Featuring chapters authored by leading experts from industry and academia, this book:

- Demonstrates advanced design techniques that enable digital circuits and sensitive analog circuits to coexist without any compromise
- Describes the process technology needed to address the performance challenges associated with developing complex mixed-signal circuits
- Deals with modeling topics, such as reliability, variability, and crosstalk, that define pre-silicon design methodology and trends, and are the focus of companies involved in wireless applications
- Develops methods to move analog into the digital domain quickly, minimizing and eliminating common trade-offs between performance, power consumption, simulation time, verification, size, and cost
- Details approaches for very low-power performances, high-speed interfaces, phase-locked loops (PLLs), voltage-controlled oscillators (VCOs), analog-to-digital converters (ADCs), and biomedical filters
- Delineates the respective parts of a full system-on-chip (SoC), from the digital parts to the baseband blocks, radio frequency (RF) circuitries, electrostatic-discharge (ESD) structures, and built-in self-test (BIST) architectures

Mixed-Signal Circuits explores exciting opportunities in wireless communications and beyond. The book is a must for anyone involved in mixed-signal circuit design for future technologies.

K24255



CRC Press
Taylor & Francis Group
an informa business

www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

ISBN: 978-1-4822-6062-5



9 781482 260625



Mixed-Signal Circuits

EDITED BY

Thomas Noulis

Intel Corporation, Munich, Germany

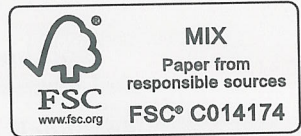


CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business



CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2016 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper
Version Date: 20150421

International Standard Book Number-13: 978-1-4822-6062-5 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Mixed-signal circuits / Thomas Noulis, editor.
pages cm. -- (Devices, circuits, and systems ; 46)
"A CRC title."
Includes bibliographical references and index.
ISBN 978-1-4822-6062-5 (alk. paper)
1. Mixed signal circuits. 2. Integrated circuits. I. Noulis, Thomas, editor.

TK7874.55.M59 2016
621.3815--dc23

2015015695

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Contents

Preface.....	xi
Editor.....	xiii
Contributors.....	xv
1. Ultra-Low Voltage Analog Filters for Biomedical Systems	1
<i>Costas Psychalinos, Farooq A. Khanday, and Georgia Tsirimokou</i>	
2. Offset Reduction Techniques in Flash A/D Converters.....	31
<i>Lampros Mountrichas and Stylianos Siskos</i>	
3. Protecting Mixed-Signal Technologies against Electrostatic Discharges: Challenges and Protection Strategies from Component to System	59
<i>Marise Bafleur, Fabrice Caignet, Nicolas Nolhier, Jean-Phillppe Laine, and Patrice Besse</i>	
4. Variability and Reliability Issues in Mixed-Signal Circuits	93
<i>Georgios D. Panagopoulos</i>	
5. Mixed-Signal Circuit Testing Using Wavelet Signatures	125
<i>Michael G. Dimopoulos, Alexios Spyronasios, and Alkis Hatzopoulos</i>	
6. Topological Investigations and Phase Noise Analyses in CMOS LC Oscillator Circuits	145
<i>Ilias Chlis, Domenico Pepe, and Domenico Zito</i>	
7. Design of an Energy-Efficient ZigBee Transceiver.....	173
<i>Antonio Ginés, Rafaella Fiorelli, Alberto Villegas, Ricardo Doldán, Manuel Barragán, Diego Vázquez, Adoración Rueda, and Eduardo Peralías</i>	
8. Simulation Techniques for Large-Scale Circuits	205
<i>Nestor Evmorfopoulos, Sotiris Bantas, and George Stamoulis</i>	
9. Mixed-Signal IC Design Addressed to Substrate Noise Immunity in Bulk Silicon toward 3D Circuits.....	233
<i>Olivier Valorge, Fengyuan Sun , Jean-Etienne Lorival, Francis Calmon, and Christian Gontrand</i>	

10. FIR Filtering Techniques for Clock and Frequency Generation	261
<i>Ni Xu, Woogeun Rhee, and Zhihua Wang</i>	
11. Design-for-Test Methods for Mixed-Signal Systems.....	279
<i>Mani Soma</i>	
12. Built-In Testing and Tuning of Mixed-Signal/RF Systems: Exploiting the Alternative Testing Paradigm.....	299
<i>Abhijit Chatterjee and Jacob Abraham</i>	
13. Spectrally Pure Clock versus Flexible Clock: Which One Is More Efficient in Driving Future Electronic Systems?.....	325
<i>Liming Xiu</i>	
14. Machine Learning-Based BIST in Analog/RF ICs	349
<i>Dzmitry Maliuk, Haralampos-G. Stratigopoulos, and Yiorgos Makris</i>	
15. Closed-Loop Spatial Audio Coding.....	379
<i>Ikhwana Elfitri</i>	
Index	401

Preface

This book addresses mixed-signal integrated circuits using advanced design techniques to enable digital circuits and sensitive analog circuits to co-exist without any compromise. Different related topics are addressed, such as the advanced process technology to address the performance challenges associated with developing these complex mixed-signal circuits, the related blocking points in the industry design flow, and the general validation of the proposed solutions and implementations. Development and implementation of innovative methodologies to move analog into the digital domain quickly, minimizing and eliminating common trade-offs between performance, power consumption, simulation time, verification, size, and cost containment are also discussed.

Specifically, in this book, the state of the art in integrated circuit design in the context of mixed-signal applications is addressed. New, exciting opportunities in different areas like wireless communications, data networking, and simulation and verification techniques are presented. Design concepts for very low-power performance and approaches for high-speed interfaces, PLL, VCOs, ADC converters, and biomedical filters are described. Respective parts of a full system-on-chip (SoC), from the digital parts until the base-band blocks, the RF circuitries, the ESD structures and the built-in self-test architectures are provided.

Coverage includes advanced crucial topics like signal integrity, large-scale simulation, and verification and testing. Extremely hot modeling topics are also addressed such as reliability, variability, and crosstalk that define pre-silicon design methodology and trends and are the main research items for all industry leading companies involved in wireless applications.

The book is written by a mixture of top industrial experts and key academic professors and researchers. Practical enough to understand how these technologies work, but not a product manual and, at the same time, scientific enough but not pure academic theory.

This book is a must for anyone involved in mixed-signal circuit design for future technologies. The intended audience is engineers with advanced integrated circuit background working in the semiconductor industry. This book can also be used as a recommended reading and supplementary material in a graduate course curriculum and, in general, the intended audience is professionals working in the integrated circuit design field.

I hope you enjoy reading this book as much as we have enjoyed writing it!

Thomas Noulis

Editor

March 30, 2015

MATLAB® and Simulink are registered trademarks of The MathWorks, Inc. For product information, please contact:

The MathWorks, Inc.
3 Apple Hill Drive
Natick, MA 01760-2098 USA
Tel: 508 647 7000
Fax: 508-647-7001
E-mail: info@mathworks.com
Web: www.mathworks.com

Editor

Thomas Noulis is a staff RFMS engineer at Intel Corporation in the Mobile & Communications Group in Munich, Germany, specializing in circuit design, modeling–characterization, crosstalk, and SoC product active area minimization. Before joining Intel, from May 2008 to March 2012, Dr. Noulis was with HELIC Inc., initially as an analog/RF IC designer, and then as an R&D engineer specializing in substrate coupling, signal and noise integrity, and analog/RFIC design. Thomas Noulis earned a BSc in physics (2003), an MSc in electronics engineering (2005), and a PhD in the Design of Signal Processing Integrated Circuits (2009) from the Aristotle University of Thessaloniki, Greece, and in collaboration with LAAS (Laboratoire d'Analyse et d'Architectures des Systèmes), Toulouse, France. During 2004–2009, he participated as principal researcher in multiple European and national research projects related to space application and nuclear spectroscopy IC design, while between 2004 and 2010, he also collaborated as a visiting-adjunct professor with universities and technical institutes. Dr. Noulis is the author of more than 30 publications, journals, conferences, and scientific book chapters. He holds one French and World patent. His work has received more than 50 citations. He is an active reviewer of multiple international journals and has given multiple invited presentations at European research institutes and international conferences on crosstalk and radiation detection IC design. Dr. Noulis has received awards for his research at conferences and by research organizations and can be reached at t.noulis@gmail.com.

Contributors

Jacob Abraham

Department of Electrical and
Computer Engineering
The University of Texas
Austin, Texas

Marise Bafleur

Laboratoire d'Analyse et
d'Architecture des Systèmes
(LAAS)
Toulouse, France

Sotiris Bantas

Centaur Technologies
Volos, Greece

Manuel Barragán

Laboratoire TIMA
Centre National de la Recherche
Scientifique
Grenoble, France

Patrice Besse

Freescale Semiconductor Inc.
Toulouse, France

Fabrice Caignet

Laboratoire d'Analyse et
d'Architecture des Systèmes
(LAAS)
Toulouse, France

Francis Calmon

Institut des Nanotechnologies de
Lyon
Université de Lyon
Lyon, France

Abhijit Chatterjee

Electrical and Computer
Engineering
Georgia Institute of Technology
Atlanta, Georgia

Ilias Chlis

Tyndall National Institute
and
Electrical and Electronic
Engineering
School of Engineering
University College Cork
Cork, Ireland

Michael G. Dimopoulos

Laboratoire TIMA
Université Grenoble Alpes
Grenoble, France

Ricardo Doldán

ARQUIMEA DEUTSCHLAND
GmbH
Frankfurt (Oder), Germany

Ikhwana Elfitri

Department of Electrical
Engineering
Andalus University
Padang, Indonesia

Nestor Evmorfopoulos

Department of Computer
Science
University of Thessaly
Volos, Greece

Rafaella Fiorelli

Instituto de Microelectrónica de
Sevilla (IMSE-CNM-CSIC)
Universidad de Sevilla
Seville, Spain

Antonio Ginés

Instituto de Microelectrónica de
Sevilla (IMSE-CNM-CSIC)
Universidad de Sevilla
Seville, Spain

Christian Gontrand

Institut des Nanotechnologies de
Lyon
Université de Lyon
Lyon, France

Alkis Hatzopoulos

Department of Electrical and
Computer Engineering
Aristotle University of
Thessaloniki
Thessaloniki, Greece

Farooq A. Khanday

Department of Electronics
and Instrumentation
Technology
University of Kashmir
Srinagar, Jammu, and Kashmir,
India

Jean-Phillippe Laine

Freescale Semiconductor Inc.
Toulouse, France

Jean-Etienne Lorival

Institut des Nanotechnologies de
Lyon
Université de Lyon, INSA- Lyon,
CNRS-UMR
Villeurbanne, France

Yiorgos Makris

Department of Electrical
Engineering
Erik Jonsson School of Engineering
and Computer Science
University of Texas
Dallas, Texas

Dzmitry Maliuk

Quantlab Financial LLC
Houston, Texas

Lampros Mountrichas

Electronics Laboratory of the
Physics Department
Aristotle University of Thessaloniki
Thessaloniki, Greece

Nicolas Nolhier

Laboratoire d'Analyse et
d'Architecture des Systèmes
(LAAS)
Toulouse, France

Georgios D. Panagopoulos

Intel Mobile Communications
GmbH
Munich, Germany

Domenico Pepe

Tyndall National Institute
Cork, Ireland

Eduardo Peralías

Instituto de Microelectrónica de
Sevilla (IMSE-CNM-CSIC)
Universidad de Sevilla
Seville, Spain

Costas Psychalinos

Physics Department
University of Patras
Rio Patras, Greece

Woogeun Rhee

Institute of Microelectronics
Tsinghua University
Beijing, China

Adoración Rueda

Instituto de Microelectrónica de
Sevilla (IMSE-CNM-CSIC)
Universidad de Sevilla
Seville, Spain

Stylios Siskos

Electronics Laboratory of the
Physics Department
Aristotle University of
Thessaloniki
Thessaloniki, Greece

Mani Soma

Electrical Engineering
Department
University of Washington
Seattle, Washington

Alexios Spyronasios

Dialog Semiconductor GmbH
Stuttgart, Germany

George Stamoulis

Department of Computer Science
University of Thessaly
Volos, Greece

Haralampos-G. Stratigopoulos

Sorbonne Universités
Paris, France

Fengyuan Sun

Electronics Department
Northwestern Polytechnical
University
Xi'an, China

Georgia Tsirimokou

Physics Department
University of Patras
Rio Patras, Greece

Olivier Valorge

EASII-IC
Electronics Design Center
Lyon, France

Diego Vázquez

Instituto de Microelectrónica de
Sevilla (IMSE-CNM-CSIC)
Universidad de Sevilla
Seville, Spain

Alberto Villegas

Innovaciones Microelectrónicas
S.L. (Anafocus, E2V)
Seville, Spain

Zhihua Wang

Institute of Microelectronics
Tsinghua University
Beijing, China

Liming Xiu

TAF Microelectronics
Dallas, Texas

Ni Xu

Institute of Microelectronics
Tsinghua University
Beijing, China

Domenico Zito

Tyndall National Institute
and
Electrical and Electronic
Engineering
School of Engineering
University College Cork
Cork, Ireland

15

Closed-Loop Spatial Audio Coding

Ikhwana Elfitri

CONTENTS

15.1 Overview of MPEG Surround	380
15.1.1 Spatial Analysis and Synthesis	382
15.1.2 Quantization and Coding	384
15.2 Closed-Loop R-OTT Module	386
15.2.1 Analysis-by-Synthesis Framework	386
15.2.2 R-OTT Module within AbS-SAC Framework	387
15.3 Simplified AbS Algorithm	388
15.3.1 Full Search AbS Optimization	388
15.3.2 An Approach for Algorithm Simplification	389
15.3.3 Basic Scheme of the Encoder	390
15.3.4 Suboptimal Algorithm	392
15.3.5 Complexity of Sub-Optimal Algorithm	393
15.4 Results	393
15.4.1 Evaluation of Closed-Loop R-OTT Module	395
15.4.2 Evaluation of Sub-Optimal Algorithm	396
15.4.3 Objective Evaluation	396
15.4.4 Subjective Evaluation	398
15.4.5 Complexity Assessment	399
15.5 Conclusions	399
References	400

Having in mind the growing demand for the reliable delivery of high-quality multichannel audio in various multimedia applications such as home entertainment, digital audio broadcasting, computer games, music streaming services as well as teleconferencing, efficient coding techniques [1] have become paramount in the audio processing arena. The traditional approach for compressing multichannel audio has been to encode each audio channel using a mono audio coder, such as Dolby AC-3 and MPEG advanced audio coder (AAC) [2]. However, for the majority of coders adopting this method, the number of bits to be transmitted tends to increase linearly with the number of channels.

Recently, a new concept for encoding multichannel audio signals has been proposed. It comprises the extraction of the spatial cues and the downmixing

of multiple audio channels into a mono or stereo audio signal. The downmix signals are subsequently compressed by an existing audio encoder and then transmitted, accompanied by the spatial cues coded as spatial parameters. Any receiver system that cannot handle multichannel audio can simply remove this side information and just render the downmix signals. This provides the coder with backward compatibility, which is important for implementation in various legacy systems. In addition, by utilizing the spatial parameters, the downmix signals can be directly upmixed at the decoder side into a multichannel configuration that may be different from the one used at the encoder side. This technique is known as spatial audio coding (SAC).

Various SAC techniques, such as binaural cue coding (BCC) [3] and MPEG 1/2 layer 3 (MP3) Surround [4], have been proposed. Interchannel level difference (ICLD), interchannel time difference (ICTD), and interchannel coherence (ICC) are extracted as spatial parameters that are based on human spatial hearing cues. Techniques such as parametric stereo (PS) [5] and MPEG Surround (MPS) [6,7] may also utilize signal processing techniques, such as decorrelation. The great benefit of these perceptual-based coders is that they can achieve bitrates as low as 3 kb/s for transmitting spatial parameters, as in the case of MPS.

As one can observe, each of these coding techniques has its unique advantages and disadvantages. However, they can all be classified as open-loop systems, where the encoders of BCC, MP3 Surround, and MPS do not consider the decoding process during encoding. The major drawback of an open-loop system is that there is no mechanism employed to reduce the error introduced by quantizing the spatial parameters and coding the downmix signals. In this chapter, an analysis by synthesis spatial audio coding (AbS-SAC) technique is presented, which provides the advantages of a closed-loop system in order to improve the quality of multichannel audio reproduction. We believe that the AbS-SAC technique can be implemented in any of the recent SAC schemes, even though in this work the AbS-SAC is applied solely in the context of the MPS architecture.

The rest of this chapter is organized as follows. Section 15.1 provides an overview of MPS. Subsequently, the closed-loop R-OTT module and the simplified AbS algorithm are presented in Sections 15.2 and 15.3, respectively.

Q1 Experiments that proofing the proposed methods and results are given in Section 15.4 followed by conclusions in the last section.

15.1 Overview of MPEG Surround

To provide a meaningful illustration, a basic block diagram of MPS is shown in Figure 15.1. An analysis quadrature mirror filterbank (A-QMF) is used to decompose the audio signal in each channel into subband signals, while a

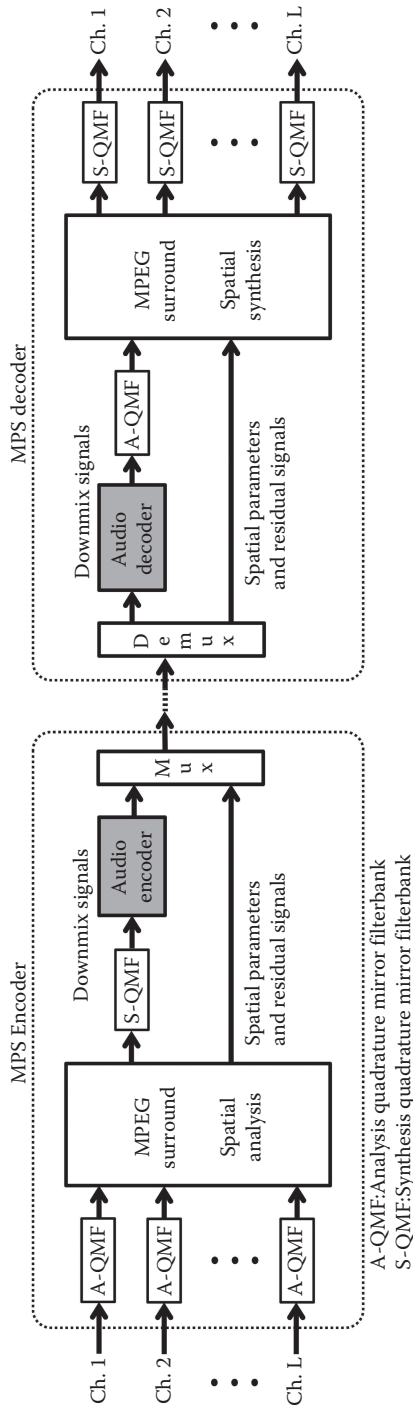


FIGURE 15.1
Block diagram of MPEG Surround.

synthesis QMF (S-QMF) is used to transform them back into a time-domain signal. Considering the importance of reducing the number of audio channels, multiple audio signals at the encoder side are typically downmixed into a mono or stereo signal, which can be further encoded by any kind of audio coder such as MP3 and high-efficiency advanced audio coder (HE-AAC). The major benefit of downmixing audio channels is a possibility to employ an existing, legacy audio coder, allowing backward compatibility. In order to be capable of creating back all audio channels at the decoder side, channel level differences (CLDs), ICCs, and channel prediction coefficients (CPCs) must be extracted as spatial parameters and transmitted as side information of the downmix signal. Furthermore, residual signal can be computed as error compensation due to downmixing process and transmitted to the decoder for enabling high-quality audio reconstruction. Interestingly, when operating at lower bitrates the residual signal can be ignored and replaced in the decoder by a synthetic signal constructed using a decorrelator.

15.1.1 Spatial Analysis and Synthesis

The MPS system comprises two pairs of elementary building blocks for channel conversion and the reverse process: one-to-two (OTT) and two-to-three (TTT) modules. The OTT module is used to convert a single channel to two channels while the TTT module is used to convert two channels to three channels. The reverse conversions are done by the reverse OTT (R-OTT) module and the reverse TTT (R-TTT) module. CLDs, ICCs, and residual signal are extracted from the R-OTT module, whereas CPCs, ICCs, and residual signal are calculated from the R-TTT module. The whole process in the encoder and decoder is built up by combining several OTT and TTT modules in a tree structure. This section simply describes the extraction of CLD, ICC as well as residual signal, as they are implemented within the proposed framework. For further details on CPC the readers can refer to ref. [10].

The schematics of the OTT and R-OTT modules are depicted in Figure 15.2. The R-OTT converts two input channels into one output channel and then extracts CLD and ICC as spatial parameters. Conversely, the OTT re-synthesizes two channels from one channel, utilizing the spatial parameters.

The first spatial parameter, CLD denoted as C , relates energies of the audio signals in the first and second channels which can be written as

$$C = \frac{e_{x_1}}{e_{x_2}} = \frac{\sum_n x_1[n]x_1^*[n]}{\sum_n x_2[n]x_2^*[n]} \quad (15.1)$$

where the $x_1^*[n]$ and $x_2^*[n]$ represent the complex conjugate of $x_1[n]$ and $x_2[n]$, respectively. For transmission, the quantized logarithmic values of the CLDs are conveyed. The second spatial parameter, ICC denoted as I , reflects

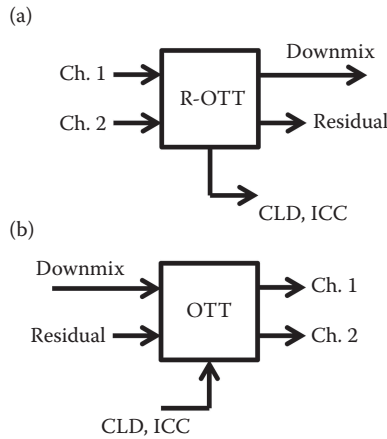


FIGURE 15.2 Block diagram of (a) the OTT module and (b) the R-OTT module as used in MPS.

the degree of correlation between both input channels which can be determined by

$$I = \text{Re} \left\{ \frac{\sum_n x_1[n]x_2^*[n]}{\sqrt{e_{x_1}e_{x_2}}} \right\} \tag{15.2}$$

The downmix signal $y[n]$ is a scaled sum of the input signals. One possible representation of the downmix signal can be written as

$$y[n] = \frac{x_1[n] + x_2[n]}{a + b} \tag{15.3}$$

where the energy constants a and b are calculated as

$$(a + b)^2 = \frac{e_{x_1} + e_{x_2} + 2I\sqrt{e_{x_1}e_{x_2}}}{e_{x_1} + e_{x_2}} \tag{15.4}$$

representing the energy preservation constraint [10].

Furthermore, the residual signal $r[n]$ is determined from the following decomposition: Q2

$$x_1[n] = ay[n] + r[n] \tag{15.5}$$

$$x_2[n] = by[n] - r[n] \tag{15.6}$$

which produces a single residual signal for reconstructing both $x_1[n]$ and $x_2[n]$.

At the decoder side, both audio signals are recreated by estimating a and b as follows:

$$\hat{a} = X \cos(A + B) \quad (15.7)$$

$$\hat{b} = Y \cos(A - B) \quad (15.8)$$

where the X , Y , A , and B variables given as

$$X = \sqrt{\frac{\hat{C}}{1 + \hat{C}}} \quad (15.9)$$

$$Y = \sqrt{\frac{1}{1 + \hat{C}}} \quad (15.10)$$

$$A = \frac{1}{2} \arccos(\hat{I}) \quad (15.11)$$

$$B = \tan \left[- \left(\frac{X - Y}{X + Y} \right) \arctan(A) \right] \quad (15.12)$$

are determined from the quantized values of CLD, \hat{C} , and the quantized values of ICC, \hat{I} . Hence, both signals can be reconstructed as

$$\hat{x}_1[n] = \hat{a}\hat{y}[n] + \hat{r}[n] \quad (15.13)$$

$$\hat{x}_2[n] = \hat{b}\hat{y}[n] - \hat{r}[n] \quad (15.14)$$

which are similar to Equation 15.5 but use the decoded downmix and residual signals, $\hat{y}[n]$ and $\hat{r}[n]$, respectively.

15.1.2 Quantization and Coding

The logarithmic value of the extracted CLD, calculated as $10 \cdot \log_{10}(C)$, is represented using one of the following nonuniform quantization values:

$$\text{CLD} = [-150, -45, -40, -35, -30, -25, -22, -19, -16, -13, -10, -8, \\ -6, -4, -2, 0, 2, 4, 6, 8, 10, 13, 16, 19, 22, 25, 30, 35, 40, 45, 150]$$

where five bits are allocated to send the index of this quantized CLD. Additionally, the extracted ICC, I , is represented by one of the following nonuniform quantization values:

$$\text{ICC} = [-0.99, -0.589, 0, 0.36764, 0.60092, 0.84118, 0.937, 1]$$

where three bits are allocated for transmitting the index of the quantized ICC.

The residual signal is encoded in the same way as in LC-AAC. The MPS standard specifies the transformation of the residual signal from the sub-band domain to the spectral coefficients of the MDCT transform. A frame, comprised of 1024 spectral coefficients, is segmented as scale factor bands, whereas many as 49 scale factor bands are used.

For each band, a scale factor is determined and the spectral coefficients are quantized as follows:

$$ix[k] = \text{sign}(r[k]) \cdot \text{nint} \left[\left(\frac{|r[k]|}{\sqrt[4]{2^{S_F}}} \right)^{0.75} \right] \quad (15.15)$$

where $ix[k]$ is the quantized spectral coefficient with its value limited from -8191 to $+8191$, $r[k]$ is the spectral coefficient of the residual signal, and S_F is the scale factor. Consequently, 14 bits are required to represent the index of the quantized value. The quantizer may utilize a psychoacoustic model to compute the maximum allowed distortion while an analysis by synthesis procedure is carried out to select the best scale factor and quantized spectral coefficient, resulting in a minimal error.

Huffman coding is then carried out for further compression where several Huffman codebooks, designed for different sets of spectral data, are provided. The quantized spectral coefficients of one or more scale factor bands are grouped and then encoded with an appropriate codebook depending on the maximum absolute value of the quantized spectral coefficients within the group. Groups that have all of their quantized spectral coefficients at zero values are associated to codebook 0, where all of the quantized spectral coefficients within this group do not need to be transmitted. Particular attention is given to the last codebook which is provided for groups with the maximum absolute values greater than or equal to 16 when a special escape sequence is used. A set of n -tuples of quantized spectral coefficients within a group, consisting of either two or four coefficients, are then represented as Huffman codewords. To keep their size small, most codebooks are given unsigned values. Thus, the sign of each nonzero coefficient is represented as an additional bit appended to the unsigned codeword.

15.2 Closed-Loop R-OTT Module

15.2.1 Analysis-by-Synthesis Framework

Analysis-by-synthesis (AbS) technique is a generic method that has already been implemented in many areas, such as estimation and identification. Several decades ago, this concept was proposed as a framework for encoding speech signals and determining the excitation signal in a linear predictive coding (LPC)-based speech coder. Since then, many other speech coders have been proposed within this framework, such as the most popular code-excited linear prediction (CELP), which is currently specified as one of the tools in the MPEG-4 Audio standard. CELP is also currently adopted in the development of the MPEG standard, ISO/IEC 23003-3/FDIS, Unified Speech and Audio Coding (USAC). The AbS technique is currently applied in many applications, including the quantization of spectral coefficients of the MPEG-AAC audio codec.

An AbS system is able to synthesize a signal by a set of parameters where the values of these parameters are usually made variable in order to produce the best-matched synthesized signal. The difference between the observed signal and the synthesized signal, called the error signal, is utilized in an error minimization block. A set of parameters which produce minimum error signal are selected as optimal parameters and sent to the decoder.

The framework of the AbS-SAC is given in Figure 15.3. A spatial synthesis block, similar to that performed at the decoder side, is embedded within the AbS-SAC encoder as a model for reconstructing multichannel audio signals. Assuming that there is no channel error, the audio signals synthesized by the model in the encoder will be exactly the same as the reconstructed audio signals at the decoder side. The error minimization block is used to compare the input signals with the reconstructed signals based on a suitable criterion such as mean squared-error (MSE) or other perceptual relevant criterion.

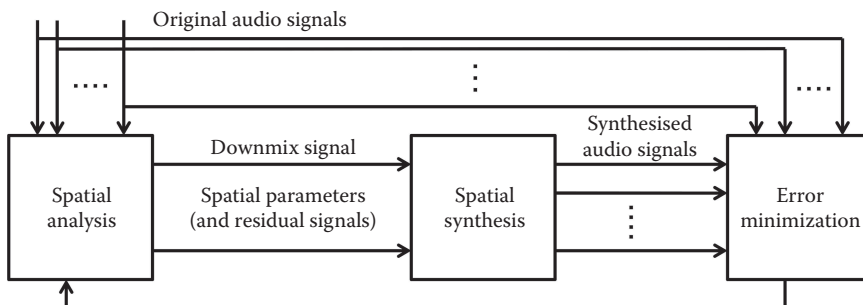


FIGURE 15.3
Framework of analysis-by-synthesis spatial audio coding (AbS-SAC).

The resultant downmix and residual signals as well as the optimal spatial parameters are then transmitted to the decoder.

Based on this framework, various implementations are possible. They are listed as follows:

- Any approach of spatial analysis and synthesis can be implemented.
- Different numbers of input and output channels can also be used.
- Various types of suitable error criterion can be utilized.
- For taking into consideration the error introduced in the communication channel, a block modeling the channel error can be inserted between the spatial analysis and synthesis block.
- The AbS-SAC approach can be implemented for only a single parameter recalculation without multiple iterations. This can be considered as a modification to the original open-loop spatial analysis block.
- The implementation can be intended to find the optimal synthesized signals by performing the trial and error procedure. Either the original blocks or the adapted version of the spatial analysis and synthesis can be applied.

15.2.2 R-OTT Module within AbS-SAC Framework

The OTT and R-OTT modules, as used in MPS, can be implemented within the AbS-SAC framework, where two channels of the original audio signals are fed to an R-OTT module as the spatial analysis block. On the other hand, the OTT module is performed as the spatial synthesis block for reconstructing two channels of synthesized audio signals, so that Equation 15.8 becomes the formula of the model. In this chapter, this is referred to as a closed-loop R-OTT module where a new optimized downmix signal can be approximated as

$$y_{\text{new}}[n] = \frac{x_1[n] + x_2[n]}{\hat{a} + \hat{b}} \quad (15.16)$$

Moreover, based on the new optimized downmix signal Equation 15.5 can be used to obtain the expression for the new optimized residual signal as

$$r_{\text{new}}[n] = x_1[n] - \hat{a}y_{\text{new}}[n] = \hat{b}y_{\text{new}}[n] - x_2[n] \quad (15.17)$$

where either $x_1[n] - \hat{a}y_{\text{new}}[n]$ or $\hat{b}y_{\text{new}}[n] - x_2[n]$ can be used to determine $r_{\text{new}}[n]$. If both input signals have the exact same magnitude but opposite phases (i.e., $x_1[n] = -x_2[n]$), then the downmix signal has all-zero values, $y_{\text{new}}[n] = 0$. Consequently, the residual signal can be determined as

$r_{\text{new}}[n] = x_1[n] = -x_2[n]$, and a specific information has to be transmitted to the decoder conveying this information.

The proposed closed-loop R-OTT algorithm relies on an approximation in Equation 15.16 when recalculating the downmix signal as a new, optimized signal, $y_{\text{new}}[n]$. Therefore, the signal distortion reduction process is based on how to create the new optimized downmix signal on the encoder side, such that the synthesized downmix signal on the decoder side fulfils the desired criteria. In practice, this is achieved by ensuring that the approximation error, which is the difference between the synthesized downmix signal, $\hat{y}[n]$, and the new optimized downmix signal, $y_{\text{new}}[n]$, is minimized. In order to obtain the minimum approximation error, both the synthesized and the approximated signals should be synchronized and compared.

The closed-loop R-OTT method is ideally capable of considerably minimize the error introduced by the quantization process of the spatial parameters. This is because the new optimized downmix and residual signals, $y_{\text{new}}[n]$ and $r_{\text{new}}[n]$, are computed based on estimated energy constants, \hat{a} and \hat{b} , so that the quantization errors of CLD and ICC are now compensated for through the newly optimized signals. Consequently, the quantization errors of CLD and ICC no longer affect the overall distortion of the synthesized audio signals.

15.3 Simplified AbS Algorithm

15.3.1 Full Search AbS Optimization

Referring to the proposed AbS-SAC framework for SAC as given in Figure 15.3, a case of the simplest AbS implementation to encode two-channel audio signals can be illustrated. In order to find the optimal downmix and residual signals, as well as the optimal parameters, a full search AbS optimization should be applied. An OTT module is used as a model for reconstructing two channels of audio signals. As the full searching procedure is performed, there is no need for applying the spatial analysis block.

An AbS optimization procedure can be carried out in such a way that the inputs of the optimization procedure are the quantization values of the downmix and residual signals, as well as the spatial parameters. All of these inputs can be varied to reconstruct various forms of audio signals. The purpose of the AbS optimization procedure is to examine all possible outcomes obtained by combining all inputs in every possible way, that is, all possible combinations of every variable. For each combination, the OTT module reconstructs audio signals and the error minimization block then computes signal distortion. Any combination that obtains minimum error is chosen as the optimal one.

The quantization values of the spectral coefficient quantizer of the downmix and residual signals that are becoming the inputs of the AbS optimization procedure range from -8191 to $+8191$, meaning that there are 16,383 quantization values for each spectral coefficient of the downmix signals. On the other hand, 31 and 8 quantization values of the CLD and ICC, respectively, are also available, provided that the MPS's quantizers are used. As a result of combining all those quantization values, for each index of the spectral coefficient the number of available combinations can be computed as $16,383 \times 16,383 \times 31 \times 8 = 6.6564 \times 10^{10}$. Note that for simplifying the calculation, the spatial parameters are assumed to be calculated for every spectral coefficient. The scale factor band is ignored.

15.3.2 An Approach for Algorithm Simplification

A simplified trial and error procedure can be applied in order to find suboptimal signals and parameters as a solution for the impractical implementation requirements of the full search AbS procedure. Three steps of simplifications are applied to make the algorithm simple. First, the number of parameters and spectral coefficients involved in the searching procedure are significantly reduced. For instance, rather than finding an optimal spectral coefficient from all quantization values, a suboptimal coefficient is simply chosen from a limited number of quantization values which are assigned based on decoded spectral coefficients. Considering the trade-off between the complexity and the degree of suboptimality, the number of coefficients and parameters involved in the searching procedure can be made variable.

Second, the main AbS-SAC algorithm is performed as a sequential process in that the suboptimal signals and parameters are not selected at the same time but one after the other. The suboptimal spectral coefficients of the downmix and residual signals can be determined first. Once the suboptimal downmix and residual signals are found, the suboptimal CLDs and ICCs can be selected. Alternatively, suboptimal spatial parameters are selected first followed by choosing suboptimal spectral coefficients. Performing the searching algorithm in a sequential process will significantly reduce the number of possible combinations to be examined.

Finally, the sequential process is performed iteratively until an insignificant error reduction is achieved. The reason for performing the iteration process is that the suboptimal spatial parameters are found based on the selected suboptimal downmix and residual signals. Additionally, the suboptimal spectral coefficients of the downmix and residual signals are selected based on the chosen suboptimal spatial parameters. Consequently, it is possible to re-optimize the downmix and residual signals after determining suboptimal spatial parameters. In contrast, it is also possible to re-select new suboptimal spatial parameters once suboptimal downmix and residual signals are found. It is expected that undertaking the iteration process will gradually reduce signal distortion.

The limited number of quantization values, where the suboptimal spectral coefficient is selected from, can be determined as follows: the spectral coefficients decoded by the spectral decoder become the inputs to the algorithm. A number of quantization values defined as the candidates for suboptimal spectral coefficient named as predetermined vector, $ix_p[k]$, can then be assigned based on the decoded spectral coefficients, $ix[k]$, as

$$ix_p[k] = [ix[k] - v, ix[k] - v + 1, \dots, ix[k] - 1, ix[k], ix[k] + 1, \dots, ix[k] + v - 1, ix[k] + v] \quad (15.18)$$

where $ix[k]$ is the decoded spectral coefficient as in Equation 15.15, k is the index of spectral coefficient, and $2v + 1$ is the size of the predetermined vector, $ix_p[k]$, with v an integer number reflecting the computational complexity of the searching procedure.

A limited number of quantization values of spatial parameters are determined in a similar way. The CLDs and ICCs obtained from the quantizer are used as the initial values. A set of predetermined values of CLDs, C_p , and a set of predetermined values of inter channel coherences, I_p , are determined using

$$P_p = \begin{cases} [P(1), P(2), \dots, P(p-1), P(p), P(p+1), \dots, P(2w), P(2w+1)] \\ \text{if } P(p) \leq w+1; \\ [P(p-w), P(p-w+1), \dots, P(p-1), P(p), P(p+1), \dots, P(p+w-1), \\ P(p+w)] \\ \text{if } w+1 < P(p) \leq P_{\max} - w; \\ [P(P_m - 2w), P(P_m - 2w + 1), \dots, P(p-1), P(p), P(p+1), \dots, P(P_m - 1), \\ P(P_m)] \\ \text{if } P_m - w < P(p); \end{cases} \quad (15.19)$$

where $P(p)$ is the initial decoded parameter, P_m is the size of the codebook, P is the spatial parameter which is either C , or I , P is the index of the initial decoded parameter, and w is an integer number reflecting the complexity of the procedure. For CLD $w \leq 15$ and for ICC $w \leq 3$.

15.3.3 Basic Scheme of the Encoder

The suboptimal AbS optimization is performed based on the proposed MDCT-based closed-loop R-OTT module. The downmix and residual signals, as well as the spatial parameters extracted from the closed-loop R-OTT module, are used as the initial input of the suboptimal AbS optimization. Practically, the AbS-SAC encoder is implemented, as shown in Figure 15.4. The audio signal in each channel is transformed to spectral coefficients by

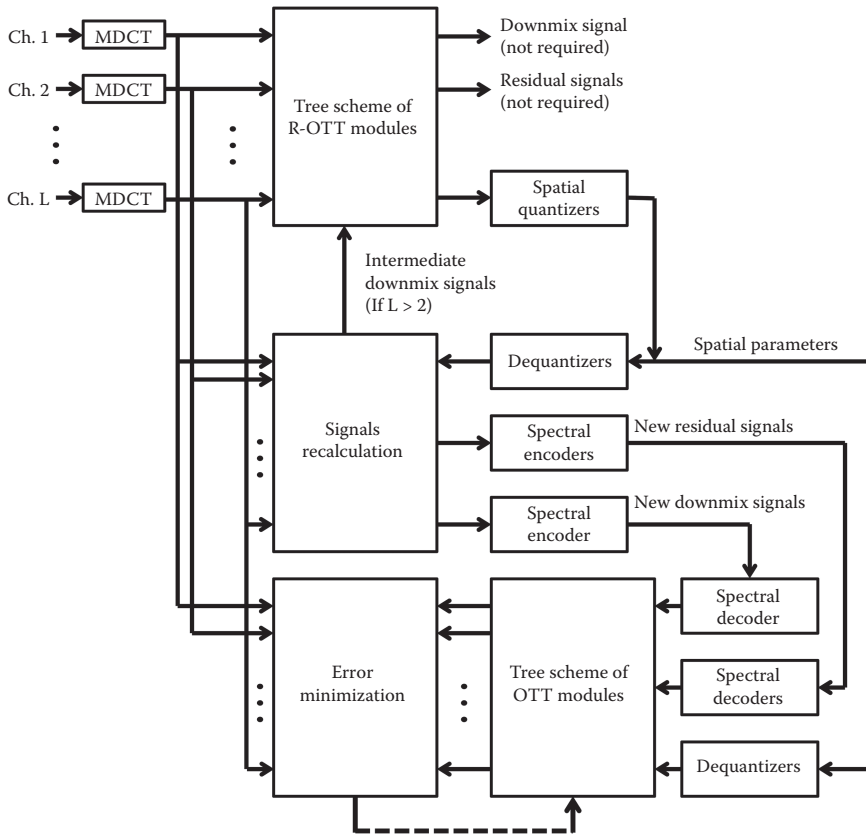


FIGURE 15.4
Block diagram of the AbS-SAC encoder.

means of MDCT. A tree scheme employing the closed-loop R-OTT modules is then performed in order to extract the spatial parameters. The new optimized downmix and residual signals are calculated based on the quantized spatial parameters. The downmix signal, as well as the residual signals, are then encoded by the spectral encoders, which actually consist of spectral coefficient quantization and the noiseless coding scheme. Prior to being supplied to the tree of OTT modules, all signals are decoded back by the spectral decoders. The decoded downmix and residual signals, as well as the quantized spatial parameters, are then used by the tree of OTT modules to upmix the audio signals. The error minimization block compares the spectral coefficients of the reproduced audio signals with those of the original signals and then computes the errors.

A closed-loop optimization procedure, utilizing the error minimization block and the tree of OTT modules, can be carried out. The inputs of the optimization loop are the decoded spectral coefficients of the downmixed

and residual signals, as well as the decoded spatial parameters. The purpose of the closed-loop approach is to examine all possible outcomes obtained by combining all quantization values of the spectral coefficient and spatial parameter quantizer in every possible way.

15.3.4 Suboptimal Algorithm

The operation of the algorithm can be explained as follows. For each index of the spectral coefficient, the candidates for suboptimal spectral coefficients of the downmix and residual signals are combined. Then, all resulting possible combinations are examined to jointly choose a suboptimal coefficient, which provides the smallest error among the other tested coefficients, for the downmix signal, as well as a suboptimal coefficient for each residual signal. In performing this task, the tree of OTT modules takes the decoded spatial parameters, given by the dequantizer, as inputs. However, the process of selecting the suboptimal spatial parameters is not performed at this stage.

Note that this task has to be carefully completed by considering the Huffman encoding process. As explained previously (see Section 15.1.2), there is a case where all spectral coefficients within a group have zero values. For such a case there is no need to transmit the magnitude of the spectral coefficients. Modifying one or more spectral coefficients within that group causes the Huffman encoding process to be associated with another codebook. As a result, the spectral coefficients within the group need to be transmitted, which causes an increase in the transmitted bitrate. However, the error reduction achieved by modifying those spectral coefficients may not provide a worthy advantage, due to an increase in the transmitted bitrate.

For this reason, the search for the suboptimal spectral coefficients, particularly downmix and residual signals, is not performed if the maximum absolute value of a group is zero. The process of choosing the suboptimal spectral coefficients is then followed by the selection of the suboptimal spatial parameters. For each parameter band, predetermined values of CLDs and ICCs are combined to form all possible combinations. A set of CLDs and ICCs is then selected as the suboptimal spatial parameters. At this stage, the downmix and the residual signal optimization is not performed.

The whole sequential process of searching for suboptimal signals and parameters is repeated as an iteration process. For the second and subsequent iterations, the chosen suboptimal signals and parameters should be used as inputs of sequential process rather than the ones from the spectral decoder and the spatial dequantizer. The iteration process is terminated when error reduction below a given threshold has been reached. The quality of the reconstructed audio is expected to improve with every iteration, however, the amount of improvement may reduce. Even if further iterations are executed the resultant values may not converge to the optimal values of both the spectral coefficients and the spatial parameters. Thus, the goal of this simplified AbS algorithm is not to provide the optimal or near-optimal

signals and parameters. Instead, it is intended to provide a solution for impractical implementation of the full search AbS procedure while minimize the signal distortion.

The whole sequential process of searching for suboptimal signals and parameters is repeated as an iteration process. For the second and subsequent iterations, the chosen suboptimal signals and parameters should be used as inputs of the sequential process rather than the ones from the spectral decoder and the spatial dequantizer. The iteration process is terminated when an error reduction below a given threshold has been reached. The quality of the reconstructed audio is expected to improve with every iteration, however, the amount of improvement may reduce. Even if further iterations are executed, the resultant values may not converge to the optimal values of both the spectral coefficients and the spatial parameters. Thus, the goal of this suboptimal algorithm is not to provide the optimal or near-optimal signals and parameters. Instead, it is intended to provide a solution for the practical implementation of the optimal searching procedure while minimizing the signal distortion.

15.3.5 Complexity of Sub-Optimal Algorithm

The proposed simplified AbS algorithm is scalable, and its algorithmic complexity depends mainly on the number of loop procedures that have to be performed in finding the suboptimal spectral coefficients and spatial parameters, as well as the number of R-OTT modules whose signals and parameters are optimized. In order to reduce the complexity of the algorithm, the number of loop procedures and the involved R-OTT modules can be decreased. Based on equations from 15.7 through 15.14, for each spectral coefficient an R-OTT module performs a number of operations, N_{ott} , consisting of 264 additions/multiplications. Hence, the number of operations, N_{op} , for each index of spectral coefficient required by the simplified AbS algorithm can be determined as

$$N_{op} = N_{ott} \times N_{loop} \quad (15.20)$$

where N_{ott} is the number of operations performed by an R-OTT module and N_{loop} is the number of loop procedures need to be performed. As an illustration, the number of loop procedures, N_{loop} , that have to be executed for encoding five-channel audio signals is given in Table 15.1.

15.4 Results

In order to evaluate the proposed system, a number of experiments designed to assess the encoding of 5 and 10 audio channels were conducted. The audio

TABLE 15.1
Complexity of Simplified AbS Algorithm

<i>v</i>	Size of Predetermined Vector	Number of Loops
0	1	0
1	3	243
2	5	3125
3	7	16,807
W	Size of Predetermined Vector	Number of Loops
0	1	0
1	3	6561
2	5	390,625

TABLE 15.2
List of Audio Excerpts for Experiments

Excerpt Name	Description
Applause	Clapping hands of hundreds of people
Drum	Drum and male vocal with guitar as background
Laughter	Sound of hundreds of people laughing
Talk	Male and female speech with music background
Vivaldi	Classical music with vocal

excerpts, sampled at 48 kHz, listed in Table 15.2 were prepared for the experiments. They were selectively chosen from a broad range of long sequence 5.1 audio signals ranging from speeches, pop, and classical music, as well as specific sounds such as clapping hands. For each audio sequence, a limited 12-s audio excerpt was selected based on the possibility of more transient events. All of the 10-channel audio signals were produced by upmixing the five-channel signals using a simple amplitude panning technique. The tree scheme of R-OTT modules for downmixing five channels into a mono downmix, as given in Figure 15.5, was used in the experiments. However,

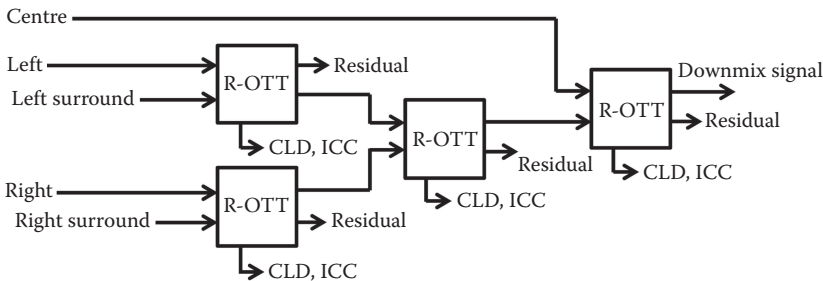


FIGURE 15.5
A tree scheme of R-OTT modules used in the experiments.

the low-frequency enhancement (LFE) channel was excluded for simplicity. Each channel of audio signals was segmented into 2048 time-domain samples with 50% overlap. For calculating CLDs and ICCs 20 parameter bands were used. The 20 parameter bands for the MDCT-based R-OTT were determined by mapping the 20 parameter bands of the CM-QMF to the 49 scale factor bands of the spectral coefficients. The downmix signal was encoded by AAC. The AAC multichannel codec, implemented as FAAC 1.28 and FAAD2 2.7, was used for benchmarking to demonstrate the usefulness of the proposed AbS-SAC approach, even though it is not the best implementation of the AAC standard.

15.4.1 Evaluation of Closed-Loop R-OTT Module

The evaluation of the closed-loop R-OTT module is aimed at demonstrating that the closed-loop approach can both improve the segSNR compared to the open-loop and perform much better in the MDCT domain. Table 15.3 shows the results of the experiment comparing segSNRs of five-channel audio encoders employing open-loop and closed-loop R-OTT modules in both the CM-QMF-based and MDCT-based scenarios. All encoders operate at a bitrate of 160 kb/s per audio channel. Here all encoders are optimized to provide maximum segSNR performance.

The results show that the closed-loop R-OTT method improves average segSNR of all the tested audio excerpts. It clearly indicates that the closed-loop R-OTT method is capable of minimizing signal distortion, resulting in segSNR improvement. Furthermore, the MDCT-based closed-loop R-OTT module outperforms, in terms of segSNR, the CM-QMF-based closed-loop R-OTT module for all tested audio excerpts other than Applause. For this Applause audio excerpt, both CM-QMF and MDCT schemes are competitive. It is an indication that the closed-loop R-OTT module generally performs better in the MDCT domain when compared with the CM-QMF domain. Moreover, the results also show that the open-loop R-OTT module has smaller segSNRs in the MDCT domain rather than in the CM-QMF domain. As expected, it indicates that the open-loop R-OTT module is basically not an appropriate

TABLE 15.3

SegSNRs (dB) of Open- and Closed-Loop R-OTT Modules

Audio	CM-QMF		MDCT	
	Open loop	Closed loop	Open loop	Closed loop
Excerpt				
Applause	24.42	30.53	21.33	30.04
Drum	23.46	32.82	21.31	39.72
Laughter	23.92	32.87	21.31	37.05
Talk	24.34	31.13	20.65	37.82
Vivaldi	21.91	27.35	19.20	44.35
Mean	23.61	30.94	20.76	37.80

method to be applied in the MDCT domain. However, employing a closed-loop approach indicates that a significant improvement is achieved, even better than the CM-QMF-based method.

15.4.2 Evaluation of Sub-Optimal Algorithm

The performance of the suboptimal algorithm is assessed by investigating the SNR of the reconstructed audio signals against the complexity of the algorithm. The suboptimal algorithm is tested for various computational complexities by assigning $v = 0, 1, 2, 3$ and $w = 0, 1, 2$. Assigning $v = 0$ and $w = 0$ means that the suboptimal algorithm is not performed. Moreover, $v = 0$ and $w \neq 0$ means that the suboptimal algorithm chooses the suboptimal spatial parameter but does not find the suboptimal spectral coefficients of the downmix and residual signals. On the other hand, defining $v \neq 0$ and $w = 0$ instructs the algorithm to select the suboptimal spectral coefficients without selecting suboptimal spatial parameters. The algorithm is terminated when the segSNR improvement is less than 10^{-4} . The average segSNRs achieved by the AbS-SAC, operating at three different bitrates, 42.8, 65.4, and 83.2 kb/s per audio channel, are given in Table 15.4.

As can be seen, the encoder employing suboptimal algorithm (i.e., $v \neq 0$ and $w \neq 0$), for various values of v and w , can improve the segSNR although the improvement is different for each complexity level. The results suggest that the best configuration to perform the proposed suboptimal algorithm, in terms of the segSNR improvement with the least complexity, is achieved by setting $v = 1$ and $w = 1$. However, to lower the complexity to a more reasonable level one can set $v = 1$ and $w = 0$ and still achieve considerable segSNR improvement.

15.4.3 Objective Evaluation

The goal of this experiments is to objectively assess perceptual quality of the proposed AbS-SAC for various operating bitrates. To our knowledge, no

TABLE 15.4
Average SegSNRs (dB) for Various Complexity Levels

Bitrate/Channel		$v = 0$	$v = 1$	$v = 2$	$v = 3$
42.8 kb/s	$w = 0$	21.82	24.31	24.45	24.50
	$w = 1$	22.38	25.12	25.16	25.18
	$w = 2$	22.41	25.19	25.19	25.22
64.4 kb/s	$w = 0$	27.82	30.41	30.45	30.46
	$w = 1$	28.33	30.77	30.79	30.79
	$w = 2$	28.34	30.79	30.81	30.82
83.2 kb/s	$w = 0$	30.66	33.51	33.54	33.72
	$w = 1$	31.12	33.72	33.74	33.74
	$w = 2$	31.13	33.73	33.75	33.75

objective perceptual test is currently available for high-quality multichannel audio signals. Thus, we have adapted the perceptual evaluation of audio quality (PEAQ), an ITU-R BS.1387-1 recommendation for assessing a mono audio signal, and currently under standardization process to include multichannel audio assessment, for evaluating multichannel audio signals. The objective difference grade (ODG), that has five grades: 0 (imperceptible), -1 (perceptible but not annoying), -2 (slightly annoying), -3 (annoying), and -4 (very annoying), was first measured for each channel of audio signal. The average values of the ODG scores over all channels are then presented as the final results for multichannel audio. A software developed by McGill University is used for calculating ODG score. Moreover, the experiments also include encoding of 10-channel audio signals. This is intended to show that, for larger channels at the given bitrate per audio channel, the performance improvement is even higher. Considering the complexity of the AbS-SAC encoder, for encoding 5 audio channels, the suboptimal algorithm is assigned with $v = 1$ and $w = 1$ while, for encoding 10 audio channels, parameters are set to $v = 1$ and $w = 0$.

The results of the experiments for encoding 5-channel and 10-channel audio signals are given in Figure 15.6. For simplicity, the results of AAC 10-channel are not shown, as they are almost identical to those achieved on AAC 5-channel. The overall ODG, averaged over all audio excerpts as shown in the lowest right plot, shows that the AbS-SAC, applied to both 5 and 10 channels, significantly outperforms, in terms of PEAQ, the tested AAC multichannel for all operating bitrates from 40 to 96 kb/s per audio

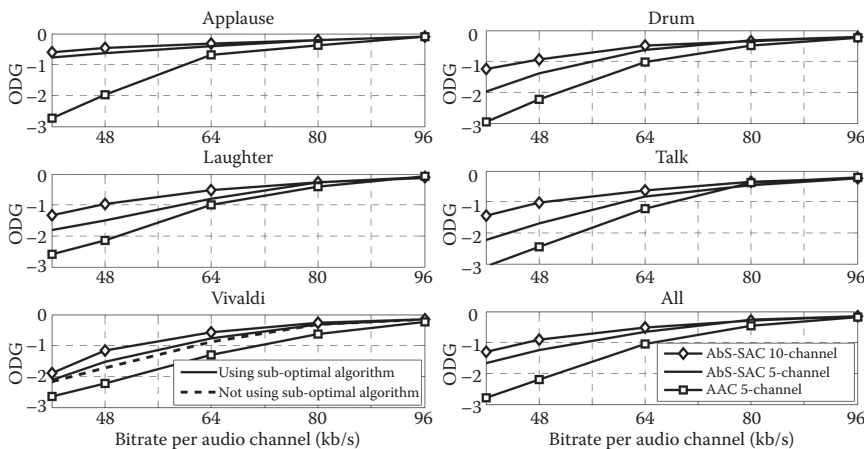


FIGURE 15.6

Objective difference grade (ODG) of the AbS-SAC for various bitrates in comparison with the tested AAC multichannel. The ODG scores of the tested AAC multichannel, for encoding 10-channel audio signals, are not plotted as they are similar to the ODG scores of the tested AAC multichannel for encoding 5-channel audio signals.

channel. However, the performance increase is greater when encoding 10 channels. It can be seen that an improvement of more than 2 points of ODG grade is achieved on the Applause audio excerpt at a bitrate of 40 kb/s per audio channel. Moreover, up to 1 point of ODG improvement is achieved on every tested audio excerpt. The results indicate that the proposed AbS-SAC technique significantly improves encoding performance for a wide range of tested audio materials.

15.4.4 Subjective Evaluation

The proposed AbS-SAC approach, for encoding five-channel audio signals, has also been evaluated using listening tests. The subjective assessment of small impairments in the audio system, as recommended in the ITU-R BS.1116-1 using the “double-blind triple stimulus with a hidden reference” method, is used. The subjective difference grade (SDG), having five grades that are similar to ODG, is used. Three codecs were taken under test: AbS-SAC, AAC multichannel, and HE-AAC multichannel. In order to reduce the difficulty, that the listeners would experience in scoring the tested audio excerpts because of too small impairment, a low but still realistic bitrate should be chosen. In the experiments, a bitrate of 51.2 kb/s per audio channel, equal to an overall five-channel bitrates of 256 kb/s, is chosen for both the AbS-SAC and AAC multichannel. Below this bitrate, the proposed AbS-SAC cannot provide a significant segSNR improvement. On the other hand, operating both coders above the chosen bitrate would increase the difficulty for the listeners in assessing the tested audio excerpts. In addition, it is still in the range of the normal operation bitrates of the AAC multichannel which is used as a benchmark. Moreover, the HE-AAC multichannel operates at its maximum typical bitrate of 32 kb/s per audio channel, which is equal to 160 kb/s for all five audio channels. Operating the HE-AAC above this bitrate is not useful in terms of coding efficiency, which means that the HE-AAC multichannel may not achieve a better performance.

A total of 20 listeners participated in this listening test. As specified in the expertise of the listeners are evaluated by averaging their SDG scores over all audio excerpts. Based on this average SDG score, a postscreening method was applied. Three listeners with an average SDG score greater than zero are assumed to be unable to correctly distinguish between the hidden reference and the tested audio object. Thus, the data from those three listeners was discarded. Only the SDG scores from the other 17 listeners were used for the results.

Figure 15.7a presents the average SDG score of each tested audio codec averaged over all audio excerpts. The error bars show the 95% confidence intervals of the mean scores. The results show that the SDG scores of all the tested codecs are competitive and very close to a grade of imperceptible. However, the proposed AbS-SAC approach achieves the highest SDG score. Furthermore, Figure 15.7b shows the SDG score of every tested audio excerpt

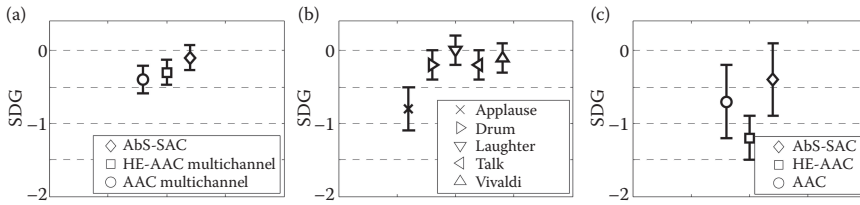


FIGURE 15.7

The results of the subjective test, to compare performance of the proposed AbS-SAC, AAC multichannel, and HE-AAC multichannel: (a) SDG scores of the tested audio codecs averaged over all audio excerpts, (b) SDG scores of all audio excerpts averaged over the tested audio codecs, (c) SDG scores of the tested audio codec for the applause audio excerpt.

averaged over all tested audio codecs, where the Applause audio excerpt has the lowest SDG score. As expected, it suggests that the Applause audio excerpt is the most critical item among the tested audio excerpts. For this Applause audio excerpt the proposed AbS-SAC approach also achieves the highest SDG score as shown in Figure 15.7c.

15.4.5 Complexity Assessment

In order to assess the performance of the suboptimal algorithm with regards to complexity, the bottom-left graph in Figure 15.6 shows the ODGs of two variants of the proposed AbS-SAC technique, with two extremely different complexity scales, for encoding five channels of the Vivaldi audio excerpt. The first one is the AbS-SAC codec using the suboptimal algorithm where $v = 1$ and $w = 0$ and all spectral coefficients at every OTT module are optimized, and the other does not use the suboptimal algorithm. The results clearly demonstrate that without the suboptimal algorithm the proposed codec is still able to achieve significant quality improvement while the suboptimal algorithm improves the performance further.

15.5 Conclusions

This chapter proposes a new AbS-SAC technique where the AbS concept is applied when choosing the suboptimal downmix signal and the spatial parameters so as to minimize the encoded signal distortion. It is demonstrated that the closed-loop R-OTT algorithm significantly reduces the error introduced by the spatial parameter quantization process resulting in significant segSNR improvement. In addition, it is shown that the frequency domain parameterization is more suitable for the AbS-SAC method instead of the sub-band domain as applied in MPS for encoding error reduction.

Additionally, the MDCT allows simplification of the coding structure by removing the transformation of the residual signals from the sub-band domain to spectral coefficients of the MDCT for the purpose of quantization. Moreover, a simplified AbS-SAC search algorithm has also demonstrated its ability to find suboptimal signals and parameters with significantly lower complexity to address the practicality of implementation of the optimal searching procedure. Subjective tests show that the AbS-SAC method outperforms, in terms of SDG score, the tested AAC multichannel, at a bitrate of 51.2 kb/s per audio channel. In addition, the AbS-SAC method has consistently higher PEAQ ODG scores than the tested AAC multichannel, for bitrates ranging from 40 to 96 kb/s per audio channel.

References

1. Brandenburg, K., Faller, C., Herre, J., Johnston, J. D., Kleijn, W. B. Perceptual coding of high-quality digital audio. *Proc. of the IEEE*, 101(9), 1905–1919, 2014.
2. Herre, J., Dietz, M. MPEG-4 high-efficiency AAC coding. *IEEE Signal Process. Mag.*, 25(3), 137–142, 2008.
3. Baumgarte, F., Faller, C. Binaural cue coding—Part I: Psychoacoustic fundamentals and design principles. *IEEE Trans. Speech, Audio, Lang. Process.*, 11(6), 509–519, 2003.
4. Moon, H. A low-complexity design for an MP3 multichannel audio decoding system. *IEEE Trans. Speech, Audio, Lang. Process.*, 20(1) 314–321, 2012.
5. Breebaart, J. van de Par, S., Kohlrausch, A. Schuijers, E. Parametric coding of stereo audio *EURASIP J. Appl. Signal Process.*, 2005, 1305–1322, 2005.
6. Hilpert, J., Disch, S. The MPEG surround audio coding standard [Standards in a nutshell]. *IEEE Signal Process. Mag.*, 26(1) 148–152, 2009.
7. Herre, J., Kjolings, K., Breebaart, J., Faller, C., Disch, S., Purnhagen, H., Koppens, J. et al. MPEG Surround—The ISO/MPEG standard for efficient and compatible multichannel audio coding. *J. Audio Eng. Soc.*, 56(11), 932–955, 2008.
8. Kondo, A. *Digital Speech: Coding for Low Bit Rate Communication Systems*. London: John Wiley Ltd, 2004.
9. Elfriti, I., Gunel, B., Kondo, A. Multichannel audio coding based on analysis by synthesis. *Proceedings of the IEEE*, 99(4), 657–670, 2011.
10. Breebaart, J., Hotho, G., Koppens, J., Schuijers, E., Oomen, W., de Par, S. V. Background, concepts, and architecture for the recent MPEG Surround standard on multichannel audio compression. *J. Audio Eng. Soc.*, 55(5), 331–351, 2007.

Q3

TO: CORRESPONDING AUTHOR

AUTHOR QUERIES - TO BE ANSWERED BY THE AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please answer these queries by marking the required corrections at the appropriate point in the text.

Query No.	Query	Response
Q1	The sense of the text “Experiments that proofing the proposed methods” is not clear. Please check?	
Q2	Equations 16.5a, 16.5b, etc. have been changed to 15.5, 15.6, etc. in sequential order as per style both in equations and cross references. Please check and confirm.	
Q3	References 8, 9 are not cited. Please cite in the text.	