Personal Sound Zone Reproduction with Room Reflections

Marek Olik

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey

Centre for Vision, Speech and Signal Processing
Faculty of Engineering and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

September 2014

© Marek Olik 2014
Summary

Loudspeaker-based sound systems, capable of a convincing reproduction of different audio streams to listeners in the same acoustic enclosure, are a convenient alternative to headphones. Such systems aim to generate “sound zones” in which target sound programmes are to be reproduced with minimum interference from any alternative programmes. This can be achieved with appropriate filtering of the source (loudspeaker) signals, so that the target sound’s energy is directed to the chosen zone while being attenuated elsewhere. The existing methods are unable to produce the required sound energy ratio (acoustic contrast) between the zones with a small number of sources when strong room reflections are present. Optimization of parameters is therefore required for systems with practical limitations to improve their performance in reflective acoustic environments. One important parameter is positioning of sources with respect to the zones and room boundaries.

The first contribution of this thesis is a comparison of the key sound zoning methods implemented on compact and distributed geometrical source arrangements. The study presents previously unpublished detailed evaluation and ranking of such arrangements for systems with a limited number of sources in a reflective acoustic environment similar to a domestic room.

Motivated by the requirement to investigate the relationship between source positioning and performance in detail, the central contribution of this thesis is a study on optimizing source arrangements when strong individual room reflections occur. Small sound zone systems are studied analytically and numerically to reveal relationships between the geometry of source arrays and performance in terms of acoustic contrast and array effort (related to system efficiency). Three novel source position optimization techniques are proposed to increase the contrast, and geometrical means of reducing the effort are determined. Contrary to previously published case studies, this work presents a systematic examination of the key problem of first order reflections and proposes general optimization techniques, thus forming an important contribution.

The remaining contribution considers evaluation and comparison of the proposed techniques with two alternative approaches to sound zone generation under reflective conditions: acoustic contrast control (ACC) combined with anechoic source optimization and sound power minimization (SPM). The study provides a ranking of the examined approaches which could serve as a guideline for method selection for rooms with strong individual reflections.

Key words: sound zones, array signal processing, sound field control, enclosures.

Email: m.olk@surrey.ac.uk
Declaration of originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.
Acknowledgements

I would like to thank Philip Jackson for supervising my work. His help in directing my research, his technical guidance, encouragement and support have been key for the progress of this work and my professional development as a researcher. I am very grateful for the time and energy he committed to assist me throughout my doctoral study.

I would like to express my gratitude to Bang & Olufsen for funding my research within the POSZ project and for their hospitality during our visits to Struer.

I greatly appreciate collaboration with Philip Coleman on the POSZ project. I have learned a great deal from him when working together on the development of the sound zone simulation software and my work has always benefited from his feedback and comments. I would also like to thank Jon Francombe and Khan Baykaner for stimulating discussions and useful feedback.

All members of the POSZ project are gratefully acknowledged for their feedback during the project meetings: Russell Mason, Chris Hummersone, and Martin Dewhirst from the University of Surrey, Martin Olsen, Martin Møller, Søren Bech, Jan Pedersen, and Adrian Celestinos from Bang & Olufsen, and Francis Rumsey. Special thanks to Martin Olsen and Martin Møller who carried out the measurements for the experimental part of this work, contributed to the development of the simulation software and provided feedback on the paper drafts and reports. The help of Jacob Dyreby, Patrick Hegarty, Mørtten Lydolf and Casper Thomsen in the experimental work is also much appreciated.

I would like to thank my colleagues from the Institute of Sound Recording for interesting discussions during the seminar days: Tim Brookes, Cleo Pike, Andy Pearce, Daisuke Koya, Tommy Ashby, Kirsten Hermes and Toby Stokes. I also appreciate the feedback on my presentations from the Machine Audition Group in CVSSP and I am grateful for the comments on my written work provided by Wenwu Wang.

Last but not least, I would like to thank the loved ones for their support during the course of my PhD. Monika’s patience, sacrifice, friendship and love have greatly supported me throughout. Natalia’s smile has always been the best reward after a hard day’s work. My great thanks go to Marianna, Zbigniew and Krzysztof for their invaluable help with looking after Natalia. Finally, thanks to Krystyna, Grzegorz, Justyna, Aleksandra and Patryk for their encouragement and reassurance.
## Contents

1 Introduction 1

1.1 Problem statement ........................................ 3

1.2 Contributions ............................................. 4

1.2.1 Comparative study of sound zone methods using a practical system 5

1.2.2 Source position optimization techniques for a reflective environment 6

1.2.3 Comparison of approaches to sound zone generation under reflective conditions ..................................... 7

1.3 Thesis structure ............................................ 8

1.4 List of publications ....................................... 9

2 Background and literature review 11

2.1 Description of a general sound zone problem .................. 11

2.2 Evaluation measures ....................................... 14

2.2.1 Acoustic contrast ...................................... 14

2.2.2 Planarity ............................................... 16

2.2.3 Array effort ........................................... 17

2.3 State of the art in sound zone reproduction ..................... 18

2.3.1 Sound energy control methods .......................... 19

2.3.2 Sound field synthesis methods .......................... 30

2.4 Possible solutions to source limitation and room reflection problems ............................................. 39

2.4.1 Limited number of sources ............................. 40

2.4.2 Room reflections ....................................... 43

2.4.3 Discussion ............................................. 47

2.5 Summary ................................................ 48
## 3 Practical implementation study

3.1 Experimental setup ............................................. 52
3.2 Sound zone reproduction procedure .......................... 54
3.3 System evaluation ........................................... 58
3.4 Discussion .................................................... 63
3.5 Summary .......................................................... 64

## 4 Optimization of source positions: analysis

4.1 Sound zone system under analysis ............................. 70
4.2 Optimal source weights ........................................ 72
4.3 Acoustic contrast expression .................................... 73
4.4 Array effort expression ......................................... 74
4.5 Minimization of sound pressure in the dark zone .......... 76
  4.5.1 Anechoic zone generation scenario ..................... 77
  4.5.2 Anechoic-reflective zone generation scenario ........ 77
  4.5.3 Reflective zone generation scenario .................. 79
4.6 Bright zone considerations .................................... 81
4.7 Summary .......................................................... 84

## 5 Optimization of source positions: simulations

5.1 Null-Split, Far-Align, and Near-Align comparison ......... 88
5.2 Numerical optimization of source positions .................. 94
  5.2.1 Compact arrays with regular spacing .................. 95
  5.2.2 Non-compact and irregular arrays ...................... 101
5.3 Summary .......................................................... 107

## 6 Evaluation of reflection control strategies

6.1 Systems and methods under evaluation ...................... 112
6.2 Contrast and effort performance ............................ 115
6.3 Robustness to implementation errors ....................... 122
6.4 Summary .......................................................... 125
7 Conclusions and further work 129
  7.1 Summary ................................................. 131
  7.2 Further work .............................................. 136
    7.2.1 Experimental work on source optimization ................. 137
    7.2.2 Advanced numerical optimization of source positions .......... 137
    7.2.3 Source optimization for complex reflective environments ........ 138
  7.3 Conclusions ................................................. 139

A Derivation of contrast and effort equations 141
  A.1 Definitions ................................................. 141
    A.1.1 Transfer functions: general definitions ...................... 141
    A.1.2 Transfer functions: a 2×2 system .......................... 142
  A.2 Derivations ................................................. 143
    A.2.1 Optimal source weights .................................. 143
    A.2.2 Sound pressure ........................................ 146
    A.2.3 Array effort .......................................... 151

B Comparison of the BC and ACC solutions 155

C Optimal source positions for compact arrays 157

D Source optimization for practical systems: pilot study 161

Bibliography 167
List of Figures

1.1 A typical personal sound reproduction scenario in a living room.  .... 2
1.2 Superposition of sound fields to create listening zones. ............. 4

2.1 Diagram of a typical sound zone system. ......................... 12
2.2 Geometry of a sound zone system based on the sound field coefficient translation theory. .................................................. 31
2.3 Geometrical relationships in the sound field coefficient translation theory. 33

3.1 Overview of the source arrays used for experimental evaluation of sound zone methods. .............................................. 53
3.2 Experimental sound zone system layout. .......................... 54
3.3 Sensor grid used for measurements in each zone. ................. 55
3.4 A normalised impulse response measured in the room with a superimposed cropping window. ........................................... 56
3.5 Assignment of LPF’s cut-off frequencies to impulse blocks in the smoothing process. ..................................................... 56
3.6 Example impulse after smoothing. ................................... 57
3.7 Sensor positions in the zones. ........................................ 57
3.8 Predicted and measured acoustic contrast. ........................ 59
3.9 Sound pressure maps in the bright zone. ............................ 60
3.10 Predicted and measured bright zone planarity. .................. 62

4.1 A 2×2 sound zone system with a reflecting rigid surface. ......... 72
4.2 Source positions for minimizing the direct sound pressure in zone B. ... 77
4.3 Source positions for minimizing the reflected sound pressure in zone B. . 79
4.4 Ranges of angles between the array axis and the monitor sensors in zone B, and the related cosine function values for configurations from Fig. 4.3. 80
4.5 (a) The normalized direct squared sound pressure component at the setup sensor in zone A. (b) Ranges of $\phi_A$ that increase the contribution of sound pressure in zone A to contrast. ................. 83

5.1 A 2\times2 system with a surface in a number of different locations. ... 89

5.2 Frequency averaged contrast produced by the 2\times2 system in the Null-Split, Far-Align, and Near-Align arrangements with a surface in different positions. ......................................................... 91

5.3 Ranges of $\theta_A$ that fulfil the requirements for increased efficiency of radiation into zone A, and arrays corresponding to the marked contrast and effort values in Figs. 5.2 and 5.4 respectively. ................. 92

5.4 Frequency averaged effort produced by the 2\times2 system in the Null-Split, Far-Align, and Near-Align arrangements with a surface in different positions. ......................................................... 94

5.5 Configuration used in the numerical search for optimal source positions. ... 95

5.6 (a) Optimized source array for a 2\times2 system under anechoic conditions. (b) Contrast produced by the optimal and alternative arrangements. ... 96

5.7 Optimized source arrays for a 2\times2 system with a single surface in different locations: (a) North, (b) East, and (c) West. ......................... 98

5.8 Results for systems with a single surface in the North location. (a) Optimized source array for a 2\times50 system. (b) Contrast produced by a 2\times50 system, using the optimal approximated Far-Align arrangement from Fig. 5.8a and the Null-Split arrangement from Fig. 5.7a. (c) Contrast produced by a 2\times2 system, using the optimal Null-Split arrangement from Fig. 5.7a and the approximated Far-Align arrangement from Fig. 5.8a. ......................................................... 99

5.9 Results for systems with a single surface in the North location. (a) Optimized source array for a 3\times50 system. (b) Contrast produced by a 3\times50 system, using the optimal Near-Align arrangement from Fig. 5.9a and the Far-Align arrangement similar to that in Fig. 5.8a. (c) Contrast produced by a 2\times50 system, using the optimal approximated Far-Align arrangement from Fig. 5.8a and the Near-Align arrangement similar to that in Fig. 5.9a. ......................................................... 100

5.10 Optimized source arrays for a 2\times2 system with up to three surfaces generating first order reflections: (a) East and North, (b) East and West, and (c) East, North and West surfaces. ......................................................... 101

5.11 Extended configuration used in the numerical search for optimal source positions. ......................................................... 102

5.12 Optimized source arrays with a single surface in the North location: (a) 2\times50, and (b) 3\times50 systems. ......................................................... 105
List of Figures

5.13 Contrast produced for the surface in the North location. (a) 3×50 system: Null-Split with compact and regular spacing, Near-Align arrangement with compact and regular spacing, and Near-Align arrangement with wide and irregular spacing. (b) 2×50 system: Null-Split arrangement with compact spacing, Near-Align arrangement with compact spacing, and Near-Align arrangement with wide spacing. 106

5.14 Optimized source arrays with a single surface in the West location: (a) 2×50, and (b) 3×50. 107

5.15 Contrast produced by Near-Align arrangements for the surface in the West location. (a) 2×50 system: compact source spacing and wide source spacing. (b) 3×50 system: compact source spacing and wide source spacing. 108

6.1 Contrast produced by ACC-based systems with a single surface in the North location: (a) 2×2, (b) 2×50, and (c) 3×50. 117

6.2 Contrast produced by systems with a single surface in the East location: (a) 2×2, (b) 2×50, and (c) 3×50. 118

6.3 Optimized ACC-based source arrays with a single surface in the West location: (a) 3×50 and (b) 2×50. 119

6.4 Effort produced by systems with a single surface in the North location: (a) 2×2, (b) 2×50, and (c) 3×50. 121

6.5 Mean contrast for 2×2 systems with numerically optimized source positions, subject to different random errors. 124

6.6 Effects of random surface position error on SPLs (1 kHz) generated by the 2×2 system optimized with the Near-Align technique (surface in the West position). 125

C.1 Optimized compact source arrays for 2×2, 2×50, and 3×50 systems based on ACC under anechoic conditions, overlaid on the sound pressure level maps at 1 kHz. 158

C.2 Optimized compact source arrays for 2×2, 2×50, and 3×50 systems based on ACC and SPM methods with a single surface, overlaid on the sound pressure level maps at 1 kHz. 159

C.3 Optimized compact source arrays for 2×2, 2×50, and 3×50 systems based on ACC with two surfaces generating first order reflections, overlaid on the sound pressure level maps at 1 kHz. 160

D.1 Configurations used to test the proposed source optimization techniques on practical systems: (a) plan view, and (b) side view. 162

D.2 Contrast produced by the ACC-based optimized array and the median contrast produced by all non-optimized configurations: (a) scenario with walls only, and (b) the complete room. 164
D.3  Sound pressure level maps at 1 kHz for the optimized array and a chosen non-optimized array in the walls-only scenario: (a) target zone A, (b) target zone B. .......................... 165
List of Symbols & Abbreviations

Greek symbols

\( \alpha \)  
Angle between the unit vector in the direction of the \( x \)-axis and the surface position vector at the same origin

\( \alpha_{d(A)} \)  
\( m \)th order coefficient of the desired sound field in zone A

\( \alpha_{\text{abs}} \)  
Absorption coefficient of a material/surface

\( \beta \)  
Lagrange multiplier (array effort constraint), eigenvalue of a matrix, regularization parameter

\( \beta_{d}^{m} \)  
\( m \)th order coefficient of the desired global sound field

\( \chi \)  
Direction of the principal energy component impinging on the bright zone \( \chi = \arg \max_{i} w_{i} \)

\( \gamma \)  
Magnitude of the surface reflection coefficient

\( \Gamma_{i} \)  
\( i \)th element of the weighting diagonal matrix for planarity control

\( \hat{\gamma} \)  
Complex surface reflection coefficient

\( \kappa \)  
Eigenvalue of a matrix

\( \Lambda \)  
Normalization factor for a source weight vector

\( \lambda \)  
Wave length

\( \mu \)  
Lagrange multiplier (squared source constraint), eigenvalue of a matrix

\( \omega \)  
Angular frequency

\( \omega'_{Anl} \)  
Reflected sound part of the transfer function between the \( n \)th setup sensor in zone A and the \( l \)th source

\( \omega'_{Bnl} \)  
Reflected sound part of the transfer function between the \( n \)th setup sensor in zone B and the \( l \)th source

\( \Omega_{Anl} \)  
Transfer function between the \( n \)th monitor sensor in zone A and the \( l \)th source
\( \omega_{Anl} \) Direct sound part of the transfer function between the \( n \)th setup sensor in zone A and the \( l \)th source

\( \Omega_{Bnl} \) Transfer function between the \( n \)th monitor sensor in zone B and the \( l \)th source

\( \omega_{Bnl} \) Direct sound part of the transfer function between the \( n \)th setup sensor in zone B and the \( l \)th source

\( \phi'_A \) Angle between the image array axis and the line between the image array centre and the setup sensor in zone A (2×2 system)

\( \phi'_B \) Angle between the image array axis and the line between the image array centre and the setup sensor in zone B (2×2 system)

\( \phi_A \) Angle between the array axis and the line between the array centre and the setup sensor in zone A (2×2 system)

\( \phi_B \) Angle between the array axis and the line between the array centre and the setup sensor in zone B (2×2 system)

\( \psi_i \) Plane wave component at the \( i \)th angle

\( \rho \) Density of air

\( \tau_l \) Time delay of the \( l \)th source

\( \theta'_{An} \) Angle between the image array axis and the line between the image array centre and the \( n \)th monitor in zone A (2×2 system)

\( \theta'_{Bn} \) Angle between the image array axis and the line between the image array centre and the \( n \)th monitor sensor in zone B (2×2 system)

\( \theta_{An} \) Angle between the array axis and the line between the array centre and the \( n \)th monitor sensor in zone A (2×2 system)

\( \theta_{Bn} \) Angle between the array axis and the line between the array centre and the \( n \)th monitor in zone B (2×2 system)

\( \zeta \) Tuning factor for the acoustic energy difference maximization method

**Roman symbols**

\( \hat{o}_{An} \) Unscaled sound pressure at the \( n \)th monitor sensor in zone A

\( \hat{o}_{Bn} \) Unscaled sound pressure at the \( n \)th monitor sensor in zone B

\( \hat{p}_A \) Unscaled sound pressure at the setup sensor in zone A (2×2 system)

\( \hat{p}_B \) Unscaled sound pressure at the setup sensor in zone B (2×2 system)

\( \hat{O}^D_{An} \) Direct unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone A, arising due to direct sound control
List of Symbols & Abbreviations

\( \hat{O}^D_{Bn} \) Direct unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone B, arising due to direct sound control

\( \hat{O}^R_{An} \) Direct unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone A, arising due to reflected sound control

\( \hat{O}^R_{Bn} \) Direct unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone B, arising due to reflected sound control

\( \hat{O}^D'_{An} \) Reflected unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone A, arising due to direct sound control

\( \hat{O}^D'_{Bn} \) Reflected unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone B, arising due to direct sound control

\( \hat{O}^R'_{An} \) Reflected unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone A, arising due to reflected sound control

\( \hat{O}^R'_{Bn} \) Reflected unscaled (complex) sound pressure component at the \( n \)th monitor sensor in zone B, arising due to reflected sound control

\( \hat{P}^D_A \) Direct unscaled (complex) sound pressure component at the setup sensor in zone A, arising due to direct sound control (2×2 system)

\( \hat{P}^R_A \) Direct unscaled (complex) sound pressure component at the setup sensor in zone A, arising due to reflected sound control (2×2 system)

\( \hat{P}^D'_{A} \) Reflected unscaled (complex) sound pressure component at the setup sensor in zone A, arising due to direct sound control (2×2 system)

\( \hat{P}^R'_{A} \) Reflected unscaled (complex) sound pressure component at the setup sensor in zone A, arising due to reflected sound control (2×2 system)

\( A \) Constraint on the sum of squared pressures in zone A

\( a \) Scaling factor

\( A_A \) Amplitude of a plane wave in zone A

\( A_B \) Amplitude of a plane wave in zone B

\( c \) Speed of sound

\( D \) Total number of source positions (features) in the optimization search

\( d \) Source spacing

\( E \) Target array effort

\( e \) Euler’s number

\( f \) Frequency

\( f_s \) Sampling frequency
List of Symbols & Abbreviations

- $f_z$: Cut-off frequency of the smoothing low-pass filter
- $g_{Al}^\prime$: Reflected sound part of the transfer function between the $m$th setup sensor in zone A and the $l$th source (2x2 notation $g_{Al}^\prime$)
- $g_{Bl}^\prime$: Reflected sound part of the transfer function between the $m$th setup sensor in zone B and the $l$th source (2x2 notation $g_{Bl}^\prime$)
- $G_{AmL}$: Transfer function between the $m$th setup sensor in zone A and the $l$th source (2x2 notation $G_{Al}$)
- $g_{AmL}$: Direct sound part of the transfer function between the $m$th setup sensor in zone A and the $l$th source (2x2 notation $g_{Al}$)
- $G_{BmL}$: Transfer function between the $m$th setup sensor in zone B and the $l$th source (2x2 notation $G_{Bl}$)
- $g_{BmL}$: Direct sound part of the transfer function between the $m$th setup sensor in zone B and the $l$th source (2x2 notation $g_{Bl}$)
- $J$: Optimization cost function
- $j$: Complex operator $\sqrt{-1}$
- $J_m$: $m$th order Bessel function
- $K$: Complex amplitude of sound pressure generated by a monopole source
- $k$: Wavenumber
- $L$: Number of sources in an array
- $l_z$: Length of the $z$th signal block used by the smoothing low-pass filter
- $M^{(0)}$: Number of modes required for accurate representation of the global sound field that includes the zones
- $M^{(A)}$: Number of modes required for accurate representation of zone A
- $M_A$: Number of setup sensors in zone A
- $M_B$: Number of setup sensors in zone B
- $N$: Number of blocks in the smoothing low-pass filter
- $N_A$: Number of monitor sensors in zone A
- $N_B$: Number of monitor sensors in zone B
- $O^{(A)}$: Origin of zone A
- $o_{An}$: Sound pressure at the $n$th monitor sensor in zone A
- $o_{Bn}$: Sound pressure at the $n$th monitor sensor in zone B
List of Symbols & Abbreviations

$p_0$ Reference magnitude of sound pressure ($2.89 \times 10^{-5} \text{ Pa}$)

$Q$ Constraint on the sum of squared source weights

$q_l$ Source weight of the $l$th source

$q_r$ Source weight of a reference source

$r_{Am}^l$ Distance between the $m$th setup sensor in zone A and the $l$th image source ($2\times2$ notation $r_{Al}^l$)

$r_A'$ Distance between the setup sensor in zone A and an image source array centre ($2\times2$ system)

$r_{Bm}^l$ Distance between the $m$th setup sensor in zone B and the $l$th image source ($2\times2$ notation $r_{Bl}^l$)

$r_B'$ Distance between the setup sensor in zone B and an image source array centre ($2\times2$ system)

$R_z^A$ Radius of zone A

$r_{Am}$ Distance between the $m$th setup sensor in zone A and the $l$th source ($2\times2$ notation $r_{Al}$)

$r_A$ Distance between the setup sensor in zone A and a source array centre ($2\times2$ system)

$r_{Bm}$ Distance between the $m$th setup sensor in zone B and the $l$th source ($2\times2$ notation $r_{Bl}$)

$r_B$ Distance between the setup sensor in zone B and a source array centre ($2\times2$ system)

$S$ Constraint on the total sum of squared sound pressures in zones A and B

$s_{Anl}^l$ Distance between the $n$th monitor sensor in zone A and the $l$th image source

$s_{An}^l$ Distance between the $n$th monitor sensor in zone A and an image source array centre ($2\times2$ system)

$s_{Bnl}^l$ Distance between the $n$th monitor sensor in zone B and the $l$th image source

$s_{Bn}^l$ Distance between the $n$th monitor sensor in zone B and an image source array centre ($2\times2$ system)

$s_{An}$ Distance between the $n$th monitor sensor in zone A and a source array centre ($2\times2$ system)
List of Symbols & Abbreviations

\( s_{Bnl} \) Distance between the \( n \)th monitor sensor in zone B and the \( l \)th source

\( s_{Bn} \) Distance between the \( n \)th monitor sensor in zone B and a source array centre (2\( \times \)2 system)

\( T_m^{(21)} \) \( m \)th order translation operator between origin 2 and 1

\( T_A \) Target spatially averaged sound pressure in zone A

\( w_i \) Energy component evaluated at the \( i \)th angle

\( X_k \) Current subset of selected source positions (features) in the optimization search

\( Y \) Superset of source positions (features) in the optimization search

**Vector and matrix symbols**

\( \hat{\mathbf{o}}_A \) Vector of unscaled sound pressures at monitor sensors in zone A

\( \hat{\mathbf{o}}_B \) Vector of unscaled sound pressures at monitor sensors in zone B

\( \hat{\mathbf{q}} \) Unscaled source weight vector

\( \mathbf{d} \) Vector of desired sound pressures at all setup sensors

\( \mathbf{d}_A \) Vector of desired sound pressures at setup sensors in zone A

\( \mathbf{d}_B \) Vector of desired sound pressures at setup sensors in zone B

\( \mathbf{e} \) Vector of errors between the desired and reproduced pressures at setup sensors

\( \mathbf{G}_q \) Matrix of transfer functions between each source and every other source

\( \mathbf{G}_A \) Matrix of transfer functions between the sources and setup sensors in zone A

\( \mathbf{G}_B \) Matrix of transfer functions between the sources and setup sensors in zone B

\( \mathbf{H}_A \) Steering matrix for monitor sensors in zone A

\( \mathbf{I} \) Identity matrix

\( \mathbf{o}_A \) Vector of sound pressures at monitor sensors in zone A

\( \mathbf{o}_B \) Vector of sound pressures at monitor sensors in zone B

\( \mathbf{p} \) Vector of sound pressures at setup sensors

\( \mathbf{p}_A \) Vector of sound pressures at setup sensors in zone A

\( \mathbf{p}_B \) Vector of sound pressures at setup sensors in zone B
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q$</td>
<td>Source weight vector</td>
</tr>
<tr>
<td>$T^d$</td>
<td>Coefficient translation matrix</td>
</tr>
<tr>
<td>$u_\phi$</td>
<td>Unit vector in the direction of the incoming plane wave</td>
</tr>
<tr>
<td>$u_{Ai}$</td>
<td>Unit vector associated with the $i$th energy component’s direction in zone A</td>
</tr>
<tr>
<td>$w_A$</td>
<td>Plane wave energy vector in zone A</td>
</tr>
<tr>
<td>$x_{Am}$</td>
<td>Position of the $m$th setup sensor in zone A</td>
</tr>
<tr>
<td>$x_{An}$</td>
<td>Position of the $n$th monitor sensor in zone A</td>
</tr>
<tr>
<td>$x_{Bm}$</td>
<td>Position of the $m$th setup sensor in zone B</td>
</tr>
<tr>
<td>$x_{Bn}$</td>
<td>Position of the $n$th monitor sensor in zone B</td>
</tr>
<tr>
<td>$y'_l$</td>
<td>Position of the image source corresponding to the $l$th source</td>
</tr>
<tr>
<td>$Y_A$</td>
<td>Steering matrix for setup sensors in zone A</td>
</tr>
<tr>
<td>$y_l$</td>
<td>Position of the $l$th source</td>
</tr>
<tr>
<td>$\alpha^d$</td>
<td>Vector of desired sound field coefficients for all zones</td>
</tr>
<tr>
<td>$\beta^d$</td>
<td>Vector of desired global sound field coefficients</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>Weighting matrix</td>
</tr>
<tr>
<td>$\Omega_A$</td>
<td>Matrix of transfer functions between the sources and monitor sensors in zone A</td>
</tr>
<tr>
<td>$\Omega_B$</td>
<td>Matrix of transfer functions between the sources and monitor sensors in zone B</td>
</tr>
</tbody>
</table>

**Coordinate notation**

- $(r, \theta)$: Position of an arbitrary observation point with respect to the global origin (in polar coordinates)
- $(R^{(A)}, \Theta^{(A)})$: Position of an observation point in zone A with respect to its origin (in polar coordinates)
- $(r^{(A0)}, \theta^{(A0)})$: Position of the origin of zone A with respect to the global origin (in polar coordinates)
- $(r_c, \theta_c)$: Position of a source with respect to the global origin (in polar coordinates)
Miscellaneous symbols

\(
\begin{align*}
* & \quad \text{Convolution operator} \\
\dagger & \quad \text{Matrix pseudo-inverse (superscript)} \\
\Re & \quad \text{Real part operator} \\
H & \quad \text{Hermitian matrix transpose (superscript)}
\end{align*}
\)
### Abbreviations and acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Acoustic contrast control</td>
</tr>
<tr>
<td>AEDM</td>
<td>Acoustic energy difference maximization</td>
</tr>
<tr>
<td>BAB</td>
<td>Branch and bound algorithm</td>
</tr>
<tr>
<td>BC</td>
<td>Brightness control</td>
</tr>
<tr>
<td>DRR</td>
<td>Direct-to-reverberant energy ratio</td>
</tr>
<tr>
<td>DS</td>
<td>Delay and sum</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic algorithm</td>
</tr>
<tr>
<td>HOA</td>
<td>Higher order ambisonics</td>
</tr>
<tr>
<td>ISM</td>
<td>Image-source model</td>
</tr>
<tr>
<td>Lasso</td>
<td>Least absolute shrinkage and selection operator</td>
</tr>
<tr>
<td>LPF</td>
<td>Low-pass filter</td>
</tr>
<tr>
<td>LS</td>
<td>Least squares</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-input, multiple-output</td>
</tr>
<tr>
<td>MLS</td>
<td>Maximum length sequence</td>
</tr>
<tr>
<td>OSD</td>
<td>Optimal source distribution</td>
</tr>
<tr>
<td>PC</td>
<td>Planarity control</td>
</tr>
<tr>
<td>PM</td>
<td>Pressure matching</td>
</tr>
<tr>
<td>PTA(l.r)</td>
<td>Plus-l Take Away-r algorithm</td>
</tr>
<tr>
<td>RT</td>
<td>Reverberation time</td>
</tr>
<tr>
<td>SFR</td>
<td>Sound field reproduction</td>
</tr>
<tr>
<td>SFS</td>
<td>Sound field synthesis</td>
</tr>
<tr>
<td>SPL</td>
<td>Sound pressure level</td>
</tr>
<tr>
<td>SPM</td>
<td>Sound power minimization</td>
</tr>
<tr>
<td>TIR</td>
<td>Target-to-interferer ratio</td>
</tr>
<tr>
<td>WDAF</td>
<td>Wave domain adaptive filtering</td>
</tr>
<tr>
<td>WFS</td>
<td>Wave field synthesis</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In a typical sound reproduction scenario in the home, the same audio content is delivered to one or more listeners in the room over loudspeakers. In some situations however, listeners may be interested in different sound programmes. For instance, parents may wish to listen to the news on the radio, while letting the children play a computer game with sound effects. For an enjoyable listening experience, the disruption from the competing audio stream should be minimized for each group of listeners.

One way to achieve this is to listen over headphones, but this has some disadvantages. First, headphones worn over longer periods may be uncomfortable, which makes them unsuitable for intensive use. Second, headphones create a physical barrier between the ears and the environment, impeding communication with others and reducing the audibility of background sounds. In contrast, a loudspeaker system operating at a moderate volume allows normal conversation and good auditory contact with the environment, thus facilitating social interaction and enhancing the listeners’ comfort. Finally, the impression of sound sources being located inside the head, inherent in conventional audio reproduction over headphones, can contribute to the listening fatigue [Bauer, 1965]. A loudspeaker-based personal audio system, capable of a convincing reproduction of different audio to listeners sharing the same acoustic space, would therefore be a useful alternative to headphones.

In order to reproduce two or more sound programmes in a single room using loudspeakers, a “sound zone” can be created for each listener, where ideally the desirable (target)
Chapter 1. Introduction

Figure 1.1: A typical personal sound reproduction scenario in a living room. Acoustically independent listening zones A and B are generated for the listeners seated in the chair and on the sofa respectively.

A programme is perceived without any interference from the competing programmes. Figure 1.1 shows a typical sound zone reproduction scenario in a living room—zones A and B are generated for the listeners seated in the chair and on the sofa respectively. Each listener occupies the acoustically “bright zone” with their own audio programme, while remaining in the “dark zone” for other programmes.

Technical problems related to delivering personal sound via loudspeakers are currently an active area of research. This thesis contributes to addressing the main unsolved challenges related to sound zone systems operating in domestic rooms. In the following, the key problems and questions are stated, the objectives and contributions of are set out, and the remainder of the thesis is outlined.
1.1 Problem statement

The main technical requirement for a sound zone system is to reproduce the target sound programme for each listener, while attenuating all competing programmes to sufficiently low levels. Adequate sound field control methods based on loudspeaker arrays are available [Coleman et al., 2014a]. Such methods generate the sound zones by appropriate filtering of the loudspeaker signals. The process is shown in Fig. 1.2 for two target sound programmes. First, the system uses appropriate signal processing to direct sound programme A to bright zone A, while attenuating it in dark zone B, as shown in Fig. 1.2a. The acoustic brightness (a high level of sound energy) is achieved by manipulating the loudspeaker signals to sum in-phase, whereas the acoustic darkness (a low level of sound energy) can be produced when the signals interfere destructively. Similar signal processing directs programme B to bright zone B, while attenuating it in dark zone A, as in Fig. 1.2b. The processing algorithm is typically based on the sound field information captured using a number of setup microphones located in each zone. Assuming the linearity of the system and acoustic transmission, individually processed programmes can be summed and reproduced over the same array of loudspeakers as in Fig. 1.2c. Provided that a sufficiently large sound energy ratio is achieved between the target and interfering programmes in each zone, each person can listen to their target audio streams without the disruption from other programmes.

Another important performance requirement is the control of the bright zone’s sound field. The system could be designed to generate a specific monophonic or stereophonic sound field, or to reproduce spatial audio. In any case, the control should ideally be consistent throughout the bright zone to allow some degree of free movement inside the zone and to ensure uniform listening experience for all listeners in the zone.

Achieving the above key performance requirements with practical sound zone systems is challenging. Such systems are subject to limitations such as a small number of available sources (loudspeakers) and sensors (microphones), and restrictions on their arrangement and placement. This reduces the system’s capacity to control the sound field. Room reflections are another factor which can significantly affect the performance. Further limitations on sound field control are placed by the requirement to keep the
Chapter 1. Introduction

Figure 1.2: Superposition of sound fields to create listening zones with a single loudspeaker array. (a) Directing programme A into zone A and attenuating it in zone B, (b) directing programme B into zone B and attenuating it in zone A, and (c) superposition of both sound fields.

This thesis aims to address the above practical challenges by scrutinizing the performance of different sound zoning methods, examining the fundamental problems related to room reflections and proposing techniques for optimizing the system’s geometrical parameters for improved performance.

1.2 Contributions

In this section, the gaps in the literature are identified, leading to formulation of the research questions, and followed by the description of the key contributions of this work.
1.2. Contributions

1.2.1 Comparative study of sound zone methods using a practical system

The existing sound zoning methods vary in their vulnerability to practical constraints such as a limited number of sources, and have different sensitivity to the influence of room reflections. Selection of the method that is best able to meet the performance requirements subject to given practical limitations and room reflections is therefore an important step in the system design process. Comparative studies of sound zone methods have been carried out, but predominantly under anechoic conditions. The methods have not been implemented and evaluated using an ensemble of metrics on different source geometries in a reflective room similar to a domestic space. As a result, the performance characteristics and capabilities of different methods and source arrangements were not fully explored.

The first contribution of this thesis aims to fill this gap in the literature by addressing the following research question:

- What is the most suitable method and source arrangement for a domestic sound zone system with practical limitations?

Chapter 3 compares the key sound zoning methods implemented on two different source arrangements—compact (linear) and distributed (circular) arrays. The systems are located in a reflective room with acoustic characteristics similar to a domestic room and the number of sources is limited to twenty-four to investigate the suitability of the methods and arrangements for practical implementations. The methods are evaluated both for the achieved acoustic separation between the zones and the properties of the bright zone sound field. The existing literature is therefore complemented with the following contribution:

- Previously unpublished detailed evaluation and ranking of key sound zone methods implemented in a reflective room on two different source geometries—compact (a line in front of the zones) and distributed (a circle surrounding the zones), using a limited number of sources.
1.2.2 Source position optimization techniques for a reflective environment

As discussed in Section 1.1, sound zone reproduction becomes increasingly difficult with deceasing number of sources and can be significantly affected by room reflections. A small system for a reflective room must therefore be carefully designed to achieve the desired performance. The key design step is optimization of parameters which ensures that the system’s full potential is exploited. An important system parameter is the position of sources with respect to the zones and surfaces. Numerical optimization studies have been published, showing that performance gains can be achieved by optimized configurations. However, the numerically-derived configurations were specific to the examined systems and environments, and may not be applicable for other conditions or system layouts. Therefore, there is a need for a systematic study on the source optimization problem to derive general solutions, applicable to a wider range of systems and rooms.

The second contribution of this thesis aims to address the above problem, focusing is on the following research question:

- How can the source positions be optimized for a sound zone system with a limited number of sources, operating in a reflective room?

In Chapters 4 and 5, the problem of optimizing source positions for small sound zone systems that are subject to room reflections is scrutinized. Particular attention is given to the key problem of strong individual reflections. A combination of analytical and numerical investigations forms a systematic study on the relationship between performance and source geometry. Techniques for optimizing source positions to mitigate the effect of strong reflections are proposed, and guidelines for increasing the system’s efficiency and robustness are formulated. The study is regarded as the most important contribution of this thesis. The specific contributions are:

- Derivation of analytic expressions for a basic system with a single reflecting surface, describing the fundamental relationship between the system’s geometrical parameters and performance.
1.2. Contributions

- Three novel source optimization techniques for mitigating the effect of individual room reflections on performance.
- Guidelines for source positioning for increasing the system’s efficiency and robustness.
- Previously unpublished numerical source position optimization results for small systems subject to strong first order room reflections.

1.2.3 Comparison of approaches to sound zone generation under reflective conditions

The existing literature presents limited examples of sound zone system evaluation under reflective room conditions. Some potentially effective approaches to sound zone generation when strong room reflections occur have not been evaluated and compared. For instance, it should be possible to reduce the impact of reflections on performance by directing the sound into the bright zone while minimizing the sound power radiated by the source array. This approach can be contrasted with a method to maximize the sound energy ratio between the bright and dark zones combined with optimization of source positions to facilitate control of the reflected sound. A comparison of these approaches could provide a useful performance ranking that could serve as a guideline for system designers.

The final contribution of this thesis is therefore centred on the following research question:

- How do different strategies for sound zone generation under reflective conditions compare?

The investigations are contained in Chapter 6, where the sound power minimization and source energy ratio maximization approaches are compared in simulations in single- and two-surface scenarios. For the latter approach, the source positions are optimized using the original techniques proposed in Chapter 4. The effect of optimizing the positions considering the direct sound only is also examined and compared to quantify
possible performance gains from including the reflections in the optimization process. The contribution of this study can be summarized as:

- Comparison of two strategies for sound zone reproduction when strong individual reflections occur: sound energy ratio maximization with optimized source positions and sound power minimization.

1.3 Thesis structure

The remainder of the thesis is organized as follows.

In Chapter 2, the literature is reviewed. Background to sound zone reproduction is given, key performance requirements are defined and the existing sound zoning methods are described. The suitability of the methods for systems with a limited number of sources, operating in reflective rooms, is discussed.

Chapter 3 presents a comparative study of sound zoning methods implemented in a reflective room using a limited number of sources. An ensemble of methods implemented on compact and distributed source arrays is evaluated to identify the combination which best satisfies the performance requirements.

In Chapter 4, a basic system with a single reflecting surface is analysed. The analytic expressions for two key performance metrics are derived. The expressions are used to propose source optimization techniques for improved performance.

In Chapter 5, the proposed techniques are examined in simulations. Guidelines for selecting the most suitable technique for a given layout of zones and surfaces are formulated. A numerical optimization of source positions is also carried out for small systems with strong individual reflections, and the relationship between the numerical and analytic solutions is examined.

In Chapter 6, the sound energy ratio maximization method with sources optimized using the proposed and alternative techniques is compared with the sound power minimization approach. The methods are evaluated on a number of small systems with up to two strongly-reflecting surfaces.
In Chapter 7, the conclusions are drawn and possible avenues of further research are discussed.

1.4 List of publications

This section lists the publications arising from the author’s PhD project work. Publications J1, C3 and C5 are closely related to the material presented in this thesis.

**Journal articles**


**Peer-reviewed conference papers**

Conference papers


Chapter 2

Sound zone reproduction: background and literature review

In this chapter, the background to the sound zone problem is provided. A two-dimensional sound zone reproduction problem is described and the key performance metrics introduced. The existing array signal processing methods, applicable to the problem of reproducing multi-zone audio in a reflective room, are described. The suitability of the existing methods to sound zone generation in a reflective acoustic environment with a limited number of sources is discussed.

2.1 Description of a general sound zone problem

A typical sound zone problem is described in this section. The majority of the research in the field has been concerned with sound zone reproduction in a 2D plane (for example [Chang et al., 2009a]). Such a problem will also be considered in this thesis, hence the description is limited to 2D here. The acoustic quantities are represented in the frequency domain, which is the most common convention in the sound zone literature (see for instance [Choi and Kim, 2002]).

Figure 2.1 shows the plan view of a typical 2D sound zone problem. Sound is reproduced
Chapter 2. Background and literature review

Figure 2.1: Diagram of a typical sound zone system. Symbols: ■ source, ● setup sensor, ○ monitor sensor, ■ surface, \(q_l\)—source weight of the \(l\)th source, \(p_{Am}\) and \(p_{Bm}\)—sound pressures at \(m\)th setup sensors in zones A and B respectively, \(o_{An}\) and \(o_{Bn}\)—sound pressures at \(n\)th monitor sensors in zones A and B respectively, \(G_{Aml}\) and \(G_{Bml}\)—transfer functions between the \(l\)th source and \(m\)th setup sensors in zones A and B respectively, \(\Omega_{Anl}\) and \(\Omega_{Bnl}\)—transfer functions between the \(l\)th source and \(n\)th monitor sensors in zones A and B respectively, \(\hat{\gamma}\)—complex reflection coefficient.

by an array of \(L\) sources, controlled by the complex source weights

\[
\mathbf{q}(f) = [q_1(f), q_2(f), \ldots, q_L(f)]^T,
\]

(2.1)

where \((f)\) denotes a function of frequency (for clarity, the frequency dependence is not indicated in the diagram). The complex source weights, also referred to as the complex source strengths, have the dimensions of the volume velocity [Kinsler et al., 2000, Chapter 7, pp. 175]. The system is located in an enclosure with reflective surfaces. The acoustic properties of each surface are characterized by the complex reflection coefficient \(\hat{\gamma}(f)\). There are two listening zones, A and B, each containing non-coincident arrays of setup and monitor sensors. The number of sensors in each zone may be equal or different. The setup sensors may be utilized to determine the source weights (setup stage), while the monitor sensors are used for evaluating system’s performance (playback stage). Using different setup and monitoring sensor locations can improve accuracy of performance predictions from simulations, as identical conditions at setup and playback, unattainable in practice, are avoided [Akeroyd et al., 2007]. It also allows minimizing bias in performance measurements. The transfer function between
2.1. Description of a general sound zone problem

the input of the $l$th source and the outputs of the $m$th setup sensor in zone A is $G_{Aml}(f)$, where $l = 1, 2, \ldots, L$ and $m = 1, 2, \ldots, M_A$. This transfer function incorporates the electro-acoustic characteristics of the source and sensor, as well as the properties of the transmission path, including the effect of reflections. Similarly, the transfer function between the $l$th source and the $n$th monitor sensor in zone A is $\Omega_{Anl}(f)$, where $n = 1, 2, \ldots, N_A$. Transfer functions between the sources and sensors in zone B can be defined similarly. Since the electro-acoustic transfer functions can be regarded in general as linear, time-invariant filters [Mourjopoulos, 1994], sound pressures at the sensors can be conveniently defined using the principle of superposition. Therefore, the vector of sound pressures at the setup sensors in zone A is

$$p_A(f) = G_A(f)q(f), \quad (2.2)$$

where $G_A(f)$ is the setup plant matrix defined as

$$G_A(f) = \begin{bmatrix} G_{A11}(f) & \cdots & G_{A1L}(f) \\ \vdots & \ddots & \vdots \\ G_{AML}(f) & \cdots & G_{AML}(f) \end{bmatrix}, \quad (2.3)$$

and the vector of sound pressures $p_A(f)$ can be written as

$$p_A(f) = [p_{A1}(f), p_{A2}(f), \ldots, p_{AM}(f)]^T. \quad (2.4)$$

By analogy, the vector of sound pressures at the monitor sensors in zone A is

$$o_A(f) = \Omega_A(f)q(f), \quad (2.5)$$

where $\Omega_A(f)$ is the monitor plant matrix written as

$$\Omega_A(f) = \begin{bmatrix} \Omega_{A11}(f) & \cdots & \Omega_{AIL}(f) \\ \vdots & \ddots & \vdots \\ \Omega_{AN1}(f) & \cdots & \Omega_{ANL}(f) \end{bmatrix}, \quad (2.6)$$

and the vector of sound pressures $o_A(f)$ is expanded as

$$o_A(f) = [o_{A1}(f), o_{A2}(f), \ldots, o_{AN}(f)]^T. \quad (2.7)$$

Sound pressures in zone B can be defined similarly.
2.2 Evaluation measures

In this section, the key evaluation measures are introduced.

2.2.1 Acoustic contrast

As discussed in Section 1.1, the main aim of a sound zone system is to reproduce the target audio programme for each listener, while attenuating all interfering programmes to adequately low levels. It is a common convention in the literature to assume that a system can reproduce all programmes with similar success (see for instance [Coleman et al., 2014a]). For the two-programme case shown in Fig. 1.2c, this means that the target to interferer sound energy ratios (TIRs) measured in zones A and B, are similar. If both target programmes are reproduced at the same levels, the TIRs will be similar to the ratio of the target energy in one zone to the interferer energy in another, also referred to as the acoustic contrast [Choi and Kim, 2002]. Thus, the system can be evaluated by considering a single programme, either as in Fig. 1.2a or 1.2b, and measuring the acoustic contrast between the bright and dark zone.

Assuming zone A is bright and zone B is dark, the contrast is formally defined as the ratio of the spatial averages of modulus squared sound pressures at the monitoring sensors in zones A and B, expressed in decibels:

\[
\text{Contrast}_{AB} = 10 \log_{10} \left( \frac{N_B \bar{\phi}_A^H \phi_A}{N_A \bar{\phi}_B^H \phi_B} \right),
\]

(2.8)

where \(N_A\) and \(N_B\) are the number of monitor sensors in zone A and B respectively, the superscript \(H\) denotes the Hermitian matrix transpose, and the frequency dependence has been omitted for convenience of notation. The assumption that zone A is bright and zone B is dark will be maintained throughout this thesis. The contrast between zone B and zone A can be similarly defined.

In practice, silence in the dark zone is not required, and it is sufficient to attenuate the unwanted programmes by a certain amount for the interference to be perceptually acceptable. Specification of the thresholds of acceptability for interfering sound programmes has been the subject of psychoacoustic research [Druyvesteyn et al., 1994;
Francombe et al., 2012; Baykaner et al., 2013]. In their pioneering work on personal sound, Druyvesteyn et al. [1994] reported the average acceptable TIRs of 11 dB for the sound programmes accompanying video material, and 20 dB when the video was switched off (the TIRs were calculated as the differences between B-weighted levels of the target and interferer programmes). Druyvesteyn et al. [1994] noted that the subjects rated consistently across the three sound programmes tested (news, music and sport), but there were larger differences in the ratings across the listeners. In more recent work, Francombe et al. [2012] confirmed that the acceptability thresholds depend largely on the listener; however, the interferer programme was found to be another important factor influencing the choice of acceptable TIRs. Furthermore, Francombe et al. [2012] observed large threshold differences across different listening scenarios, such as gathering information from audio, reading/working, and leisure listening. In the last case, 95% of the experienced and inexperienced listeners indicated average acceptable TIRs of 39 dB and 31 dB respectively. Baykaner et al. [2013] carried out a similar experiment, considering just the leisure listening scenario. A strong effect of the type of target programme was reported: the mean acceptable TIRs were in the range 11–19 dB for sport commentary, pop or classical music material. These findings suggest that a sound zone system used to reproduce entertainment audio should be able to produce at least 11 dB TIR or acoustic contrast to be successful in certain scenarios. However, a versatile system must produce more than 39 dB. Although these results must be treated with caution and additional research is required to provide further validation, the established values can serve as a useful reference when comparing the performance of different systems and methods. The 11 dB and 39 dB contrast thresholds can be referred to as lower and upper minimum system requirements respectively. This terminology is adopted in Section 2.3 where the existing sound zone methods are described and compared based on the results available in the literature.

Acoustic contrast has also been linked to perceived distraction from the interfering audio programme—a series of psychoacoustic tests undertaken by Baykaner et al. [2015] indicated that increasing the contrast results in decreasing distraction. Engineering sound zoning solutions that maximize the acoustic contrast has therefore perceptual relevance.
2.2.2 Planarity

As pointed out in Section 1.1, characteristics of the bright zone sound field are an important aspect of the system’s performance. The system could be used to generate a specific type of sound field in the zone, for instance to reproduce spatial audio effects [Poletti, 2008; Wu and Abhayapala, 2011; Coleman et al., 2014c]. The fidelity of sound reproduction is then evaluated by quantifying the error between the desired and the reproduced sound fields [Poletti, 2008; Wu and Abhayapala, 2011]. In some cases, monophonic sound reproduction may be sufficient; it is then beneficial to restrict the range of directions from which the sound arrives at the bright zone by appropriate signal processing [Coleman et al., 2013b]. This increases the sound field’s planarity, i.e. the extent to which the sound field resembles a plane wave [Jackson et al., 2013]. Highly planar bright zones have been shown to be characterized with spatial uniformity of sound levels [Coleman et al., 2013b, 2014a,b], which as discussed in Section 1.1 is a desirable property. Increasing planarity has also been shown to improve the perceived quality of the target programme reproduced by a sound zone system [Baykaner et al., 2015]. Bright zone’s planarity is therefore an important performance criterion for monophonic sound zone systems, and can be quantified using the planarity metric [Jackson et al., 2013]. Since this thesis is concerned with monophonic sound zones rather than rendering spatial audio, the focus below will be on the description of the planarity metric rather than the sound field reproduction error.

The energy distribution at the monitor sensor array in the bright zone A (over incoming plane wave direction) is given by \( w_{Ai} = \frac{1}{2}|\psi_{Ai}|^2 \), where \( w_A = [w_{A1}, \ldots, w_{AI}] \) are the energy components at the \( i \)th angle, and \( \psi_{Ai} \) is the plane wave component at the \( i \)th angle. The steering matrix \( H_A \) of dimensions \( I \times N_A \), which maps between the pressures observed at the sensors and the energy components, can then be defined so that

\[
w_A = \frac{1}{2}|H_Ao_A|^2.
\] (2.9)

A super-directive beamformer or the spatial Fourier transform can be used to determine the steering matrix weights [Jackson et al., 2013]. The planarity of zone A can be defined as the ratio between the energy due to the largest plane wave component in this zone
2.2. Evaluation measures

and the total energy flux of plane wave components:

\[
\text{Planarity}_A = \frac{\sum_i w_{Ai} \mathbf{u}_{Ai} \cdot \mathbf{u}_{A\chi}}{\sum_i w_{Ai}},
\]

(2.10)

where \( \mathbf{u}_{Ai} \) is the unit vector associated with the \( i \)th component’s direction, \( \mathbf{u}_{A\chi} \) is the unit vector in the direction \( \chi = \arg \max_i w_{Ai} \), and \( \cdot \) denotes the inner product. Planarity can be defined similarly for zone B.

Jackson et al. [2013] established planarity values for some typical sound fields. The nominal planarity score for the sound field consisting of a single plane wave is 100%. In contrast, for a standing wave field or a perfectly diffuse sound field the nominal planarity is 0%. An intermediate value of 50% can be assigned to two plane waves arriving from the directions 90° apart. A point source located at a certain distance from the measurement location should result in the planarity score of approximately 90%.

2.2.3 Array effort

In a practical system, efficiency of sound radiation is an important consideration. An inefficient system would require large electrical power of the driving signals to produce the required sound pressure level (SPL) in the bright zone, which may be beyond the capacity of the electro-acoustic equipment used. Assuming that there are no significant electro-acoustic interactions between the transducers in a source array, the electrical power used to drive the array is proportional to the sum of modulus squared source weights [Elliott et al., 2010]. The efficiency can therefore be measured with the array effort, defined as

\[
\text{Effort}_A = 10 \log_{10} \left( \frac{q_A^H q_A}{|q_r|^2} \right),
\]

(2.11)

where \( q_A^H q_A \) is the sum of modulus squared source weights reproducing target programme A, and \( q_r \) is the source weight of a single reference source that produces the same SPL in the bright zone for the same programme. Effort pertaining to programme B is defined similarly.

Large array effort indicates that the array controlled by the sound zone filters consumes much more electrical power compared to a single source with no sound zone process-
Chapter 2. Background and literature review

Excessive effort can also be associated with poor robustness to numerical and implementation errors, resulting in degradation of the achieved acoustic contrast when the errors occur [Elliott and Cheer, 2011; Elliott et al., 2012; Coleman et al., 2013a, 2014a; Coleman, 2014, Chapter 5]. Monitoring the array effort is therefore important to ensure realizable and effective sound zone filters.

2.3 State of the art in sound zone reproduction

The capability of a system to produce large acoustic contrast and generate bright zones of desired characteristics within the required array effort limits depends on the control method and the acoustic environment, as well as system parameters such as the number of sources and setup sensors, their arrangement, size and relative position of the sound zones. Generally, the effectiveness of a sound zone system improves with increasing the number of sources [Jones and Elliott, 2008; Elliott et al., 2012; Wu and Too, 2012; Coleman et al., 2014a] and sensors used to sample the sound field [Betlehem and Abhayapala, 2005; Wu and Abhayapala, 2011], and when the sources and sensors can be arranged to satisfy the requirements of the control method [Wu and Abhayapala, 2011; Coleman et al., 2014a]. Similarly, it is much easier to control small zones far apart rather than large zones close together [Møller and Olsen, 2011, Chapter 5], or under anechoic rather than reflective acoustic conditions [Druyvesteyn et al., 1994; Wen et al., 2005; Jacobsen et al., 2011; Elliott et al., 2012; Cheer, 2012; Simón-Gálvez et al., 2014].

Unfortunately, practical sound zone systems are subject to limitations: the system is required to generate sound zones of a practical size (at least encompassing the human head) using a limited number of sources and sensors, with restrictions on their arrangement and placement. Normally, the systems are implemented in a reflective room. The existing sound zoning methods vary in their capacity to tackle these practical challenges.

In this section, the state of the art sound zone generation methods are reviewed. The advantages and disadvantages of the two main categories of methods, sound energy control and sound field synthesis, are discussed in the context of domestic sound zone reproduction. Particular attention is given to the methods’ capability to perform well
with a limited number of sources and in the presence of room reflections. This provides the essential background for the remainder of the thesis, including a practical sound zone implementation study presented in the next chapter.

2.3.1 Sound energy control methods

The general principle of operation of this group of methods is controlling the acoustic field, so that the sound energy is concentrated in the bright zone, while being directed away from the dark zone. While this may produce the desired acoustic contrast between the zones, the phase of the sound pressure in the zones is not directly controlled. In consequence, the methods in this group do not provide the capacity to specify and precisely reproduce a desired sound field in the zones, which significantly limits the capability to render spatial audio. Moreover, spatial artefacts may occur, for instance uneven distribution of sound energy across the zone [Jacobsen et al., 2011; Møller et al., 2012; Coleman et al., 2013b, 2014a]. On the other hand, the lack of phase constraint allows to commit all of the system’s available degrees of freedom to generating the contrast between the zones. Therefore, large contrasts can potentially be achieved, with the interfering programmes attenuated to perceptually acceptable levels (see discussion in Section 2.2.1). In the following, the sound energy control methods are described.

Delay and sum beamforming

One of the simplest array signal processing methods that can be applied to a sound zone problem is the delay and sum beamforming (DS) [Van Veen and Buckley, 1988]. DS can be used to steer the main beam of sound energy radiated from the array at the bright zone. This is achieved by applying an appropriate delay to each source feed, so that all individual signals combine in phase at the zone’s centre. The optimal source weight vector is therefore defined as

\[
q = \left[ e^{-j\omega\tau_1}, e^{-j\omega\tau_2}, \ldots, e^{-j\omega\tau_L} \right]^T,
\]

where \( j = \sqrt{-1} \) is the imaginary unit, \( \tau_1, \tau_2, \ldots, \tau_L \) are time delays given by \( \tau_l = (\max\{r_l\} l=1 - r_l)/c \) with \( r_l \) as the distance between the \( l \)th source and the bright zone centre, \( \omega = 2\pi f \) is the angular frequency, and \( c \) is the speed of sound.
While the sound energy is concentrated in the bright zone, some amount of destructive interference also occurs in other directions. Thus, acoustic contrast can be produced. However, the sound energy reduction in off-target directions is not directly controlled and is frequency-dependent. The attenuation of low frequency signals is not very effective due to limitations in the size of practical source arrays—for long waves, the phase differences between source signals are too small for effective cancellation [Ma-band and Kellerman, 2007]. At high frequencies, the control is compromised due to spatial aliasing [Naidu, 2001, Chapter 4, Sec. 4.1.2]—the grating lobes may occur in the dark zone, affecting the contrast. This indicates that the method may not be suitable for broadband sound zone reproduction with a limited number of sources, where the physical requirements for effective beamforming at low and high frequencies, i.e. large array aperture and small inter-element spacing respectively, cannot be satisfied at the same time. The directivity performance can be improved by gradient methods (e.g. Jacobi arrays) [Weston, 1986] or optimal beamforming [Cox et al., 1986; Boone et al., 2009]; however, these methods tend to be more sensitive to implementation errors.

DS relies on the time delay estimates between the sources and the focusing point, which may be subject to implementation errors. Furthermore, the delays are calculated with respect to a single point, and so the sound energy is directed at one particular point rather than an area. Thus, the formed beam may be too narrow to cover the entire bright zone at higher frequencies. The beamformer’s performance may also be compromised in a reflective room, as the reflected sound field is not controlled by signal processing. Wen et al. [2005] demonstrated this in simulations for an “acoustic-hotspot” (bright zone) generation system based on a compact array of thirteen sources. The energy focusing capability of DS was significantly affected by uncontrolled reflections from the room surfaces (a better result was achieved by the brightness control method—see the following section).

**Brightness control**

Choi and Kim [2002] were the first to apply sound field optimization to the sound zone problem. Brightness control (BC) was one of the two optimization methods proposed in
their pioneering work (the other method will be introduced in section “Acoustic contrast control”). BC maximizes the sum of modulus squared sound pressures measured at the setup sensors in the bright zone, under the constraint that the array effort is held at a certain value. This results in sound energy concentration in the bright zone, and so acoustic contrast is possible. The optimization problem can be solved using the method of Lagrange multipliers, which defines the following Lagrangian function (cost function):

$$J_{BC} = \mathbf{q}^H \mathbf{G}_A^H \mathbf{G}_A \mathbf{q} - \beta (\mathbf{q}^H \mathbf{q} - Q),$$ (2.13)

where $\beta$ is the Lagrangian multiplier and $Q = |q_r|^2 \times 10^{E/10}$ is the sum of squared source weights corresponding to the chosen value of the array effort $E$, with $q_r$ as the source weight of the reference source. By calculating the partial differentials with respect to $\mathbf{q}$ and $\beta$ and setting them to zero, the stationary points of $J_{BC}$ can be found. This yields

$$\mathbf{G}_A^H \mathbf{G}_A \mathbf{q} = \beta \mathbf{q},$$ (2.14)

$$\mathbf{q}^H \mathbf{q} = Q.$$ (2.15)

Thus, the maximization of $J_{BC}$ reduces to solving an eigenvalue problem in Eq. (2.14), while simultaneously satisfying Eq. (2.15). Therefore, the optimal source weight vector $\mathbf{q}$ is proportional to the eigenvector $\hat{\mathbf{q}}$ corresponding to $\max \{ \beta_i \}_{i=1}^I$, where $\beta_i$ denotes the $i$th eigenvalue of $\mathbf{G}_A^H \mathbf{G}_A$, and $I$ is the total number of eigenvalues. The coefficient of proportionality $a = \mathbf{q} / \hat{\mathbf{q}}$, is found from Eq. (2.15): $a = \sqrt{Q} / \hat{\mathbf{q}}^H \hat{\mathbf{q}}$. BC therefore maximizes the squared sound pressure in the bright zone for a given chosen control effort. The method lends itself well to practical applications, where an effort limit must often be imposed to prevent equipment fatigue.

The main weakness of BC is that the sound energy reduction outside the bright zone is not targeted—pure sound focusing tends to create a relatively low contrast. On the other hand, the method is characterized by low array effort and good robustness to errors. BC has been evaluated in a number of studies. Coleman et al. [2014a] examined the method in free-field simulations. A forty-eight-element array surrounding the zones produced 10–20 dB of acoustic contrast in the frequency range 100–4000 Hz. Such contrast exceeds the lower minimum requirement defined in Sec. 2.2.1 (11 dB), but does not satisfy the upper minimum requirement (39 dB). The contrast scores are
expected to be lower in a reflective room and for a system with fewer of sources. This suggests that BC may not satisfy the psychoacoustic acceptability criteria for a wide range of programmes and listeners in a domestic sound zone system with a limited number of sources.

Jones and Elliott [2008] and Wen et al. [2005] examined BC on compact arrays. Jones and Elliott evaluated the method in simulations using a pair of closely-spaced monopoles in free field. The system aimed to generate acoustic contrast between a single bright and two dark zones, all located relatively close to the sources (up to approximately 0.8 m). The reported contrast was approximately 13–18 dB in the frequency range 50–2000 Hz. As above, this is closer to the lower rather than the upper minimum requirement. Wen et al. used BC in the “acoustic-hotspot” generation system previously described in section “Delay and sum beamforming”. BC focused the sound energy at the hotspot (bright zone) more effectively than DS. At 500 Hz, attenuation of the sound energy in a vertical plane located 2.5 m away from the array was approximately 5–20 dB, about 1 m from the bright zone and further. The performance improved slightly when the evaluation plane was 0.5 m from the array. This confirms the previous conclusion that the method may not have the capacity to generate effective sound zones in a domestic room with a limited number of sources.

**Sound power minimization**

Sound power minimization (SPM) was proposed by Jones and Elliott [2008]. The method minimizes the radiated sound power while maintaining the SPL in the bright zone at a certain chosen value. This concentrates sound in the bright zone, while reducing the sound energy flow in other directions, and so the contrast can be produced in a manner similar to BC. For an array of monopole sources, the optimization problem is expressed with the following Lagrangian function:

\[
J_{SPM} = \frac{1}{2} \mathbf{q}^H \mathcal{R}\{\mathbf{G}_q\} \mathbf{q} + \mu \left( \mathbf{q}^H \mathbf{G}_A^H \mathbf{G}_A \mathbf{q} - A \right),
\]

where \( \mu \) is the Lagrangian multiplier, \( \frac{1}{2} \mathbf{q}^H \mathcal{R}\{\mathbf{G}_q\} \mathbf{q} \) is the sound power radiated from the array, \( \mathbf{G}_q \) is the \( L \times L \) matrix of transfer functions between each source and every other
2.3. State of the art in sound zone reproduction

source, $A = 10^{\frac{T_A}{10}} p_0^2$ is the sum of squared sound pressures in zone A corresponding to the target SPL in zone A, $T_A$, and $p_0$ is the reference amplitude of sound pressure (normally the threshold of hearing for pure tone at 1 kHz; $2.89 \times 10^{-5}$ Pa [Kinsler et al., 2000, Chapter 5, pp. 131]). As in the case of BC, the stationary points of the cost function with respect to $q$ and the Lagrangian multiplier must be found by taking the partial derivatives with respect to $q$ and $\mu$ and equating the results to zero, which gives

$$-\frac{1}{2} \text{tr}(G_q) \left( G_A^H G_A \right)^{-1} q = \mu q, \quad (2.17)$$

$$q^H G_A^H G_A q = A. \quad (2.18)$$

Equation (2.17) is a familiar eigenvalue problem. The source weight vector $q$ that minimizes $J_{SPM}$ is therefore proportional to the eigenvector $\hat{q}$ that corresponds to $\min_{i=1}^I \mu_i$, where $\mu_i$ denotes the $i$th eigenvalue of $\frac{1}{2} \text{tr}(G_q) \left( G_A^H G_A \right)^{-1}$, and $I$ is the total number of eigenvalues. The coefficient of proportionality $a = q / \hat{q}$ is found from Eq. (2.18): $a = \sqrt{A / q^H G_A^H G_A \hat{q}}$.

Jones and Elliott [2008] evaluated SPM in free-field simulations for the system based on a pair of monopole sources previously discussed in section “Brightness control”. The contrast between the bright zone and the two dark zones was approximately 22–26 dB in the frequency range 50–2000 Hz. This provides improvement with respect to BC, which produced 13–18 dB, and pushes the system’s performance closer to the upper minimum contrast requirement defined in Sec. 2.2.1. Furthermore, SPM could be more robust to room reflections than BC or DS. All three methods do not directly control the dark zone’s sound field, but since SPM minimizes the radiation from the array, it also reduces the reflected energy arriving at the dark zone. This is desirable, as uncontrolled reflections are likely to increase the overall level of the interfering programme, damaging the contrast. SPM could therefore be potentially effective in practical systems; however, the method is yet to be evaluated under reflective conditions.

**Acoustic contrast control**

Acoustic contrast control (ACC) was the other method introduced in the work by Choi and Kim [2002], alongside BC. In the original cost function, it was assumed that
the dark zone is a larger region that encompasses the bright zone. With such an assumption, ACC aims to maximize the sum of squared sound pressures at the setup sensors in the bright zone, subject to the constraint that the sum of squared sound pressures in both zones is equal to a certain chosen value $S$. As for BC and SPM, the optimal source weights can be found using the method of Lagrange multipliers. The Lagrangian function is

$$ J_{ACCa} = \mathbf{q}^H \mathbf{G}_A^H \mathbf{G}_A \mathbf{q} - \mu \left[ \mathbf{q}^H \left( \mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B \right) \mathbf{q} - S \right], \quad (2.19) $$

where $\mu$ is the Lagrange multiplier. The stationary points of $J_{ACCa}$ can be found by following the familiar procedure of taking the partial derivatives with respect to $\mathbf{q}$ and $\mu$ and equating the results to zero. This yields

$$ \left( \mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B \right)^{-1} \mathbf{G}_A^H \mathbf{G}_A \mathbf{q} = \mu \mathbf{q}, \quad (2.20) $$

$$ \mathbf{q}^H \left( \mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B \right) \mathbf{q} = S. \quad (2.21) $$

The optimal source weight vector $\mathbf{q}$ is therefore the eigenvector $\hat{\mathbf{q}}$ corresponding to the largest eigenvalue of $\left( \mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B \right)^{-1} \mathbf{G}_A^H \mathbf{G}_A$, scaled to satisfy Eq. (2.21). Choi and Kim [2002] noted that the optimization problem of Eq. (2.19) is equivalent to maximization of the ratio

$$ \mu = \frac{\mathbf{q}^H \mathbf{G}_A^H \mathbf{G}_A \mathbf{q}}{\mathbf{q}^H \left( \mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B \right) \mathbf{q}}. \quad (2.22) $$

This formulation makes it clear that $J_{ACCa}$ achieves the global maximum when the sum of squared sound pressures in zone B is minimized ($\mu$ is maximum when $\mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q}$ is zero). The acoustic contrast is therefore produced between the zones.

To find the solution, the inversion of the unregularized matrix $\mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B$ is required. This may result in large numerical errors if the matrix is ill-conditioned. Moreover, the array effort is not constrained, and so the optimal filters may produce large signal gains that may be beyond the capacity of the sound reproduction equipment. Both problems are circumvented by using the “indirect” formulation of the contrast control problem, proposed by Elliott et al. [2012]. The sum of squared sound pressures in zone B is to be minimized with the constraints that SPL in zone A and array effort are set to certain chosen values. The resulting Lagrangian function is

$$ J_{ACCb} = \mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q} + \mu \left( \mathbf{q}^H \mathbf{G}_A^H \mathbf{G}_A \mathbf{q} - A \right) + \beta \left( \mathbf{q}^H \mathbf{q} - Q \right), \quad (2.23) $$
2.3. State of the art in sound zone reproduction

where $\mu$ and $\beta$ are the Lagrangian multipliers, and $A$ and $Q$ are the constraint values defined previously. The stationary points of $J_{\text{ACC}b}$ are found by calculating partial differentials with respect to $q$, $\mu$ and $\beta$ and setting them to zero, which yields

$$
(G_A^H G_A)^{-1} (G_B^H G_B + \beta I) q = \mu q,
$$

(2.24)

$$
q^H G_A^H G_A q = A,
$$

(2.25)

$$
q^H q = Q,
$$

(2.26)

where $I$ is the identity matrix with the dimensions of $G_B^H G_B$. The optimal $q$ is therefore proportional to the eigenvector $\hat{q}$ corresponding to the smallest eigenvalue of $(G_B^H G_B + \beta I) (G_A^H G_A)^{-1}$, or equivalently, the largest eigenvalue of the inverse $(G_B^H G_B + \beta I)^{-1} (G_A^H G_A)$ [Elliott et al., 2012]. The effort constraint therefore regularizes the inversion of the matrix $G_B^H G_B$. The regularizing Lagrange multiplier $\beta$ must be chosen so that Eqs. (2.25) and (2.26) are both satisfied. In practice, a numerical procedure can be used to find the multiplier’s value at each frequency [Coleman et al., 2014a]. First, the constraint from Eq. 2.25 is enforced by using appropriate scaling factor $a$, without regularization ($\beta = 0$). If $q^H q < Q$, then the effort constraint is not active. Otherwise, the value of $\beta$ such that $q^H q \leq Q$ is sought iteratively using a gradient descent search, with $A$ being fixed at each step. It is also possible to set $\beta$ to a frequency independent value that is large enough to ensure validity of the numerical solution at each frequency, for instance when the number of setup sensors in zone B is lower than the number of sources (in such cases, the matrix $G_B^H G_B$ is singular). In this case, the effort constraint may not be enforced. The scaling factor $a = q / \hat{q}$ such that Eq. (2.25) is satisfied can be determined after the regularization has been applied.

It may be beneficial to formulate the ACC problem in the time domain for practical implementations. Such formulation was proposed by Elliott and Cheer [2011]. Cai et al. [2013, 2014b] implemented the time-domain ACC with an additional constraint, demonstrating improvements in the processed signal quality with respect to the frequency-domain implementation. Furthermore, Choi et al. [2008] proposed a modified cost function to weight the importance of different setup sensors in both zones. This provides the capacity to shape the spatial distribution of sound energy, for instance to improve the consistency of sound levels across the zones. Park et al.
[2010] tested the method’s capability to reproduce stereo sound by defining the bright zone as two regions around the listener’s ears. Chang et al. [2009b] defined the bright zone similarly and placed a rigid sphere between the two regions to model the effect of scattering from the human head. Including the scattered sound field component in the setup plant matrix reduced the system’s sensitivity to reflections from the sphere, which resulted in improved contrast.

ACC has been examined in different applications, for instance mobile phones [Elliott et al., 2010; Cheer et al., 2013a], super-directive line arrays [Choi et al., 2008; Chang et al., 2009a; Choi et al., 2010; Park et al., 2010; Simón-Gálvez et al., 2012; Simón-Gálvez and Elliott, 2013; Olivieri et al., 2013; Cai et al., 2013, 2014b], personal sound for aircraft seats [Elliott and Jones, 2006; Jones and Elliott, 2008], car cabins [Coleman et al., 2012; Cheer, 2012; Cheer et al., 2013b; Cheer and Elliott, 2013b,a], and general sound reproduction applications [Jacobsen et al., 2011; Coleman et al., 2013a, 2014a; Coleman, 2014]. Table 2.1 shows contrast scores achieved by ACC in example studies. Elliott and Jones [2006] reported up to 50 dB contrast in a small reflective room, produced by two closely spaced loudspeakers for the listeners in adjacent aircraft seats. This indicates the suitability of the method for applications when the sources can be located close to the listener. A similar result was obtained by Chang et al. [2009a] in an anechoic room, for a system producing the bright zone close in front of the compact line array and the dark zones to the left and right from the listener. Acoustic contrast up to 35 dB was reported. Jacobsen et al. [2011] achieved up to 40 dB in an acoustically “dry” room with sixteen loudspeakers surrounding the zones. These results indicate that the method has the capacity to generate large acoustic contrast in a reflective environment with a limited number of sources, and so to reach or exceed the perceptually required contrast. Therefore, it can be potentially successful in practical domestic applications.

Acoustic energy difference maximization

The acoustic energy difference maximization (AEDM) method has been proposed by Shin et al. [2010] as an alternative to ACC. The method aims to maximize the sound energy difference between zones A and B, with the sum of squared source weights being
Table 2.1: Examples of acoustic contrast scores achieved by the ACC method, reported in the literature. Contrast values are based on figure readings.

<table>
<thead>
<tr>
<th>Study</th>
<th>Source array</th>
<th>Frequency range</th>
<th>Acoustic environment</th>
<th>Acoustic contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliott and Jones [2006]</td>
<td>Two loudspeakers face to face</td>
<td>0.05–2 kHz</td>
<td>Anechoic room</td>
<td>25–50 dB</td>
</tr>
<tr>
<td>Chang et al. [2009a]</td>
<td>Nine loudspeakers (line)</td>
<td>0.8–5 kHz</td>
<td>Anechoic room</td>
<td>19–35 dB</td>
</tr>
<tr>
<td>Jacobsen et al. [2011]</td>
<td>Sixteen loudspeakers (circle)</td>
<td>0.25, 0.5 and 1 kHz</td>
<td>Acoustically “dry” room</td>
<td>up to 40 dB</td>
</tr>
</tbody>
</table>

This gives the following Lagrangian

\[ J_{AEDM} = q^H G_A^H G_A q - \zeta q^H G_B^H G_B q - \beta (q^H q - Q). \] (2.27)

After following the familiar procedure of taking the partial derivatives with respect to \( q \) and \( \beta \), and equating them to zero, it is found that \( J_{AEDM} \) is maximized by the vector \( q \) that is proportional to the eigenvector \( \hat{q} \) corresponding to the largest eigenvalue of the matrix \( q^H G_A^H G_A q - \zeta q^H G_B^H G_B q \), and that satisfies the effort constraint \( q^H q = Q \).

The matrix inversion, inherent in ACC, is therefore avoided. Thus, a numerically stable solution can be obtained without regularization. The parameter \( \zeta \) is a real-valued “tuning factor” which can be used to customize the performance characteristics of the source array. When \( \zeta = 0 \), Eq. (2.27) reduces to the cost function of BC given by Eq. (2.13). The method therefore concentrates the sound energy in zone A, but the attenuation in zone B is limited. With increasing \( \zeta \), AEDM intensifies the attenuation in zone B, in the limit achieving performance characteristics similar to ACC.

The AEDM method was compared with ACC in Shin et al.’s original paper [2010]. The 2D system consisted of a circular array of 10 loudspeakers around two closely located square zones, located in an anechoic chamber (a 3D system was also evaluated, but will not be discussed here). The contrast was below 20 dB for discrete frequencies 100, 200, and 300 Hz. These scores were higher than ACC’s contrast; however, the ACC
method was unregularized, which might have affected the measured results. Elliott et al. [2012] analysed the performance of AEDM and the regularized ACC based on a small array of free-field monopoles, showing identical contrast and effort performance of the two methods for a range of values of the tuning parameter $\zeta$. It was also indicated that above a certain value of $\zeta$ the contrast starts to drop with increasing effort, which signalled the need of careful parameter selection for optimal performance.

In more recent studies, Shin et al. [2012, 2014] evaluated the performance of AEDM implemented on a compact array of loudspeakers (sixteen transducers fitted into an enclosure back to back). Narrow directivity patterns, which can result in large contrast, were produced only when the tuning factor was selected to maximize the acoustic contrast between the zones [Shin et al., 2014]. It can therefore be concluded that AEDM has the potential to generate large contrasts using a limited number of sources in a domestic room, but it is likely that the tuning will have to be such that the method approaches the ACC’s characteristics; hence, the two methods can be regarded as similar.

Planarity control

Planarity control (PC) was proposed by Coleman et al. [2013b]. Similarly to ACC, the aim of the method is to minimize the sound pressure in the dark zone, while maintaining a chosen SPL in the bright zone. However, in this case the sound energy flow into the bright zone may be constrained to a certain range of directions. This helps in maintaining the uniformity of SPL across the zone, which otherwise may be compromised [Coleman et al., 2013b]. If the spatial window defining the range of directions is suitably chosen with respect to the zone layout, the sound energy flux can be constrained, yet effective sound attenuation can still be achieved in the dark zone. This is due to the fact that the method has the freedom to choose directions of plane wave component propagation within the spatial window, and selects those which maximize contrast. The PC cost function is

$$J_{PC} = q^H G_B^H G_B q + \mu (q^H G_A^H Y_A^H Y_A G_A q - A) + \beta (q^H q - Q),$$  \hspace{1cm} (2.28)
where $Y_A$ is a steering matrix which maps between the sound pressures at the sensors and the sound field components arriving from a set of discrete directions around the sensor array. The elements of $Y_A$ can be populated by a superdirective beamformer. If the beamformer “looks” at $I$ directions using $N_A$ sensors, $Y_A$ has the dimensions $I \times N_A$.

The term $\Gamma$ is a diagonal matrix which allows weighting of the energy components from all the observed directions (steering angles): 

$$\Gamma = \text{diag} \{ \Gamma_1, \Gamma_2, \ldots, \Gamma_I \},$$

where $0 \leq \Gamma_i \leq 1$. Stationary points of $J_{PC}$ with respect to $q$, $\mu$ and $\beta$, are:

$$- (G_A^H Y_A^H \Gamma Y_A G_A) (G_B^H G_B + \beta I) q = \mu q,$$  

$$p_A^H Y_A^H \Gamma Y_A p_A = A,$$  

$$q^H q = Q.$$ 

The optimal source weight vector $q$ is therefore proportional to the eigenvector $\hat{q}$ corresponding to the largest eigenvalue of $(G_B^H G_B + \beta I)^{-1} (G_A^H Y_A^H \Gamma Y_A G_A)$. As for ACC, the constraints in Eqs. (2.31) and (2.32) can be satisfied by selecting an appropriate value of $\beta$ and scaling.

Coleman et al. [2013b] evaluated PC in simulations, implementing the method on a forty-element circular array around the zones under the anechoic conditions. The target contrast of 76 dB was achieved over the whole frequency range of interest (50–7000 Hz), demonstrating the method’s capability to produce large contrast. Moreover, high planarity scores were achieved, indicating a uniform bright zone. In further work, Coleman [2014, Chapter 4] carried out an experimental evaluation of PC in an acoustically treated reflective room. The method was implemented on a sixty-element circular array of loudspeakers encompassing the zones. The achieved contrast was in the range 5–30 dB (50–7000 Hz), and a relatively high planarity was obtained. Furthermore, the method was shown to have the capacity for stereophonic reproduction in the zones [Coleman et al., 2014c]. PC is therefore a potentially effective method for a domestic sound zone system; however, further evaluation is required using systems with a smaller number of sources.
A uniform SPL in the bright zone could also be achieved with the sound field synthesis methods (see the following section). However, this would require specifying an appropriate (e.g. planar) sound field in advance. Selecting the most suitable directions for plane wave components at each frequency can be difficult without the help of optimization, and inappropriately chosen directions could result in significant contrast deterioration [Coleman et al., 2013b]. PC circumvents this problem by optimizing the direction of the plane wave components within the specified spatial window to achieve the largest possible contrast.

2.3.2 Sound field synthesis methods

Sound field synthesis (SFS) methods have been developed for traditional single zone sound reproduction in order to improve the fidelity of sound reproduction and facilitate generation of spatial audio. The synthesis approaches can be divided into three main groups:

- methods based on spatial harmonic expansion, such as ambisonics [Gerzon, 1973; Ahrens and Spors, 2008] or alternative mode matching techniques [Ward and Abhayapala, 2001; Poletti, 2005; Wu and Abhayapala, 2009a],

- wave field synthesis (WFS) [Berkhout et al., 1993; Spors et al., 2008],

- pressure matching techniques based on the least squares approach [Kirkeby and Nelson, 1993; Poletti, 2007].

The SFS methods facilitate accurate specification and reproduction of any given sound field. Consequently, independent listening zones can also be reproduced. The phase component of the sound field can be specified in the zones, which provides a distinct advantage over the energy control methods. This feature is particularly important in the bright zone—the listening experience may be improved if a certain sound field, e.g. a plane wave, can be specified and reproduced accurately [Coleman et al., 2014a].
2.3. State of the art in sound zone reproduction

Analytic techniques

Ambisonics and WFS belong to the category of methods, where the source weights are derived analytically. In sound zoning applications, the analytic methods require translation of the sound field coefficients defining multiple zones to the coefficients of the global sound field to be reproduced by the system. The translation theory was developed by Wu and Abhayapala [2011] for a 2D sound field [cf. Wu and Abhayapala, 2009b, 2010; Abhayapala and Wu, 2009], and is applicable to WFS and reproduction methods based on spatial harmonics expansion, such as higher order ambisonics (HOA). The translation theory is summarized below.

The diagram of the system geometry is shown in Fig. 2.2. There are two circular zones, A and B. Zone A has a radius $R_z^{(A)}$ and its centre $O^{(A)}$ is located at $(r^{(A0)}, \theta^{(A0)})$ with respect to the global origin $O$. An observation point in zone A is located at $(R^{(A)}, \Theta^{(A)})$ with respect to $O^{(A)}$, and at $(r, \theta)$ with respect to $O$. A circular array of sources surrounds the zones, with the $c$th source located at $(r_c, \theta_c)$ with respect to $O$; its source weight is $q(r_c, \theta_c, f)$. Similar quantities can be defined for zone B.

![Figure 2.2: Geometry of a sound zone system based on the sound field coefficient translation theory (modified from [Wu and Abhayapala, 2011]). There are two circular zones, A and B. Zone A has a radius $R_z^{(A)}$ and its centre $O^{(A)}$ is located at $(r^{(A0)}, \theta^{(A0)})$ with respect to the global origin $O$. An observation point in zone A is located at $(R^{(A)}, \Theta^{(A)})$ with respect to $O^{(A)}$, and at $(r, \theta)$ with respect to $O$. A circular array of sources surrounds the zones, with the $c$th source located at $(r_c, \theta_c)$ with respect to $O$; its source weight is $q(r_c, \theta_c, f)$. Similar quantities can be defined for zone B.](image-url)
\((R^{(A)}, \Theta^{(A)})\) with respect to \(O^{(A)}\), and at \((r, \theta)\) with respect to \(O\). A circular array of sources surrounds the zones, with \(i\)th source located at \((r_c, \theta_c)\) with respect to \(O\); its source weight is \(q(r_c, \theta_c, f)\). The sound pressure in zone \(A\) can be represented using the cylindrical harmonic expansion [Wu and Abhayapala, 2011]:

\[
p(R^{(A)}, \Theta^{(A)}, k) = \sum_{m=-\infty}^{\infty} \alpha_m^{(A)}(k) J_m(kR_z^{(A)}) e^{jm\Theta^{(A)}}
\]

(2.33)

where \(J_m(.)\) is the \(m\)th order Bessel function, \(\alpha_m^{(A)}(k)\) is the \(m\)th coefficient of the sound field, the superscript \(\cdot^d\) denotes the desired sound field, and \(k\) is the wavenumber. The properties of the Bessel functions allow truncating the infinite series expansion in Eq. (2.33) to a finite series representing the sound field bounded by the sources [Wu and Abhayapala, 2011]:

\[
p(R^{(A)}, \Theta^{(A)}, k) = \sum_{m=-M^{(A)}}^{M^{(A)}} \alpha_m^{(A)}(k) J_m(kR_z^{(A)}) e^{jm\Theta^{(A)}},
\]

(2.34)

where the desired sound field is mode limited to \(M^{(A)}\), i.e. the sound field is described by the total of \(2M^{(A)}+1\) modes. The minimum required number of modes depends on the wavenumber and the radius of the zone, and is defined by \([M^{(A)} = kR_z^{(A)}/2]\) [Wu and Abhayapala, 2011]. The global sound field can be expressed similarly:

\[
p(r, \theta, k) = \sum_{m=-M^{(0)}}^{M^{(0)}} \beta_m^d(k) J_m(kr) e^{jm\theta},
\]

(2.35)

where \(\beta_m^d(k)\) is the \(m\)th coefficient of the global sound field, and \(M^{(0)} = [kR_z/2]\), with \(r_c\) denoting the radius of the circle just enclosing both zones.

The geometrical translation between two arbitrary coordinate systems with origins \(O^{(1)}\) and \(O^{(2)}\), is shown in Fig. 2.3. Both systems are assumed to have the same orientation. Thus, \((r^{(12)}, \theta^{(12)})\) are the coordinates of \(O^{(2)}\) with respect to \(O^{(1)}\). The soundfield coefficients, \(\alpha_m^{(1)}\) and \(\alpha_m^{(2)}\), with respect to the coordinate systems \(O^{(1)}\) and \(O^{(2)}\), respectively, are related [Wu and Abhayapala, 2011]:

\[
\alpha_m^{(1)}(k) = \alpha_m^{(2)}(k) \ast T_m^{(21)}(r^{(12)}, \theta^{(12)}, k),
\]

(2.36)

where \(\ast\) denotes a discrete convolution, and

\[
T_m^{(21)}(r^{(12)}, \theta^{(12)}, k) \equiv J_m(kr^{(12)}) e^{-jm\theta^{(12)}}.
\]

(2.37)
2.3. State of the art in sound zone reproduction

The problem of finding the equivalent global sound field coefficients for zone A can therefore be expressed as

\[ \beta^d_m(k) * T^{(0A)}_m = \alpha^{d(A)}_m(k), \] (2.39)

where \( T^{(0A)}_m \) is the translation operator for the global coordinate system to the coordinate system of zone A. Similar definitions can be given for zone B. The convolutions of Eq. (2.39) and their equivalents for zone B can now be combined into a system of equations:

\[ \alpha^d(k) = T(k)\beta^d(k), \] (2.40)

where

\[ \beta^d(k) \equiv \begin{bmatrix} \beta^d_{-M(0)}(k) \ldots \beta^d_{M(0)}(k) \end{bmatrix}^T, \] (2.41)

\[ \alpha^d(k) \equiv \begin{bmatrix} \alpha^d_{-M(A)}(k) \ldots \alpha^d_{M(A)}(k) \alpha^d_{-M(B)}(k) \ldots \alpha^d_{M(B)}(k) \end{bmatrix}^T, \] (2.42)
Chapter 2. Background and literature review

and

\[
\mathbf{T}(k) \triangleq \begin{bmatrix}
T^{(0A)}_{-M^{(A)}+M^{(0)}} & \cdots & T^{(0A)}_{-M^{(A)}-M^{(0)}} \\
\vdots & \ddots & \vdots \\
T^{(0B)}_{M^{(B)}+M^{(0)}} & \cdots & T^{(0B)}_{M^{(B)}-M^{(0)}} \\
\vdots & \ddots & \vdots \\
T^{(0B)}_{M^{(B)}+M^{(0)}} & \cdots & T^{(0B)}_{M^{(B)}-M^{(0)}}
\end{bmatrix}.
\] (2.43)

The set of desired global coefficients are found by solving Eq. (2.40):

\[
\beta^d(k) = \mathbf{T}^\dagger(k) \alpha^d(k),
\] (2.44)

where the superscript \( \dagger \) denotes the Moore-Penrose pseudo-inverse. The set of global sound field coefficients can now be reproduced by conventional sound field synthesis method, such as HOA, alternative mode matching techniques, or WFS.

A major challenge in the application of the analytic methods to practical systems is that they assume certain characteristics of the sources and sound field which may be different in practice. Typically, it is assumed that the sources are free field monopoles, whereas conventional loudspeakers are not omnidirectional at higher frequencies and the listening room is reflective. Various approaches to circumvent this problem will be discussed in Sec. 2.4.2. Another problem is a limited number of sources in practical systems. Wu and Abhayapala [2011] indicate that 2D sound field enclosed by a circle of the radius \( r_c \) requires at least \( 2M^{(0)} + 1 \) sound field coefficients, where \( M^{(0)} = \lceil ker_c/2 \rceil \).

Thus, at least \( 2M^{(0)} + 1 \) sources are required for accurate sound field reproduction. With two non-overlapping zones \( M^{(A)} = \lceil keR^{(A)}_z \rceil \) and \( M^{(B)} = \lceil keR^{(B)}_z \rceil \) is required for zone A and B respectively. Thus, the number of modes required to reproduce the global sound field is \( M^{(0)} \geq (M^{(A)} + M^{(B)}) \). So, many sources are required for larger zones and higher frequencies. For instance, for \( R^{(A)}_z = R^{(B)}_z = 0.25 \) m and frequency 8 kHz (\( k \approx 146.5 \) m\(^{-1}\)), \( M^{(A)} = M^{(B)} = 100 \), and so at least 401 sources are needed. 3D applications will further increase these requirements. With fewer sources, the accuracy of sound reproduction will decrease. It is clear that the above requirements would be difficult to satisfy in a domestic sound zone system.

The reproduction of sound zones in 2D using the method of translating the sound field coefficients has been evaluated in simulations by Jin et al. [2013]—up to 65 dB of
contrast was reported in free field. Jacobsen et al. [2011] extended the investigations into the “2.5D” case, where a circular array of monopoles was used to reproduce a 2D sound field. The contrast was 10–40 dB in the 100–1500 Hz range under simulated free field conditions, when the source requirement was satisfied. However, this result was much lower than that achieved by ACC under the same conditions (45–180 dB). Moreover, the contrast degraded significantly when the method was implemented on sixteen loudspeakers in a highly absorbing room, and a reduction of the areas of effective reproduction was observed at higher frequencies. On the other hand, SFS reproduction was shown to be more robust to the effect of scattering from the human head and torso than the unregularized ACC method under anechoic conditions [Olsen and Møller, 2013]. However, introduction of regularization into the ACC method is likely to improve performance under such conditions.

**Least-squares pressure matching**

The pressure matching (PM) technique can be used to generate sound zones in a similar way as the analytic SFS methods—locally specified bright and dark zones can be reproduced. However, the translation of local sound field coefficients to the global equivalents is not required. Instead, a direct optimization of the sound pressure field is used to derive the source weights.

The least-squares pressure matching has been investigated extensively for conventional single-zone sound field reproduction. Kirkeby and Nelson [1993] applied the method to the problem of reproducing a plane wave over a small region sampled with a number of sensors. The optimization problem was formulated with the following cost function

$$ J_{LS} = e^H e = (d - p)^H (d - p) = (d - Gq)^H (d - Gq), \quad (2.45) $$

where $e$ is a vector of complex errors between the reproduced sound pressures $p$ and the desired sound pressures $d$. In the practical overdetermined case, i.e. when there are more sensors than sources, the solution to Eq. (2.45) is

$$ q = G^d d = (G^H G)^{-1} G^H d. \quad (2.46) $$
Chapter 2. Background and literature review

The simulation results presented by Kirkeby and Nelson [1993] showed that the sound field reproduction using the least-squares PM approach is prone to similar limitations as the analytic SFS techniques. There is no direct requirement for a specific geometry of source array, as for some analytic techniques (e.g. HOA), yet the reproduction error increases when the plane wave is to be reproduced from a direction where the sources are not present. Moreover, the reproduction is sensitive to spatial aliasing of the sensor array—a high density of sensors is required to reproduce high frequencies (Kirkeby and Nelson [1993] concluded that, as “a rule of thumb”, the spacing should be less than 1/3 of the shortest wavelength).

In their later work, Kirkeby et al. [1996, 1998b] presented a regularized solution to the least squares PM problem. The cost function from Eq. (2.45) was modified to

$$J_{LS} = e^H e = (d - Gq)^H (d - Gq) + \beta (q^H q - Q), \quad (2.47)$$

where $\beta$ is the regularization parameter. The solution is

$$q = (G^H G + \beta I)^{-1} G^H d, \quad (2.48)$$

where $I$ is the identity matrix with the dimensions of $G^H G$. The regularized formulation was examined in simulations with four sources arranged on a wide arc and three sensors arranged compactly over 2 m away from the sources [Kirkeby et al., 1996]. The target sound field was a plane wave propagating from three different directions with respect to the sensor array. It was noticed that when the plane wave direction was approaching from within the aperture of the source array, the target sound field was reproduced with good accuracy over an area extending beyond the location of the sensors. For the direction not covered by the sources, accurate target sound field reproduction was limited to the immediate proximity of the sensors. This confirms that although in principle PM could be applied to arbitrary source geometries, in practice specific source locations are required for good performance.

Poletti [2008] applied the regularized least-squares PM solution to the problem of 2D multi-zone reproduction. The desired sound pressure $d$, which was a plane wave propagating in the chosen direction, was specified across the measurement points in the
zones. For two zones, A and B, this gives a vector $d = [d_A, d_B]^T$ with

$$d_A = A_A e^{jkx_{Am}} u_\phi, \text{ for } m = 1, 2, \ldots, M_A,$$

$$d_B = A_B e^{jkx_{Bm}} u_\phi, \text{ for } m = 1, 2, \ldots, M_B,$$

(2.49)

where $x_{Am}$ and $x_{Bm}$ are the positions of the $m$th sensors in zones A and B respectively, $\cdot$ denotes the inner product, and $u_\phi$ is the unit vector in the specified direction of the incoming plane wave. By specifying a lower amplitude for the dark zone’s wave, acoustic contrast could be produced (a 60 dB attenuation was used). Although the reproduction of a plane wave was considered in the original work, in principle any target sound field could be specified.

Poletti [2008] evaluated the sound field reproduction error in multiple zones in simulations. The reproduction region was circular (radius $r = 2$ m) and encompassed the zones. The target sound fields in the zones were plane waves of up to 4 kHz. The number of sources, $L = 300$, satisfied the minimum requirement to reproduce the sound fields accurately within the region enclosed by the source array, which is $L = 2kr$ [Ward and Abhayapala, 2001]. It was demonstrated that the reproduction error increased considerably when both the bright and dark zones were on the propagation path of the target plane wave. A similar characteristic was also observed for narrow-band signals by Radmanesh and Burnett [2011]. Judicious virtual source placement with respect to the zone layout is therefore essential for good performance of PM.

The results discussed above are based on the sound field reproduction error which is difficult to interpret. The contrast performance of PM was studied by Coleman et al. [2014a] using simulations. The desired sound pressure in the dark zone was set to zero. In such case, Eq. (2.47) can be reformulated to read:

$$J_{PM} = q^H G_B^H G_B q + (d_A - G_A q)^H (d_A - G_A q) + \beta (q^H q - Q),$$

(2.50)

and the optimal source weights are

$$q = (G_B^H G_B + G_A^H G_A + \beta I)^{-1} G_A^H d.$$

(2.51)

The method was simulated on a circular array of forty-eight free-field monopole sources. There were two circular zones contained in a circular reproduction region with the
radius $r = 0.75$ m. According to the previously mentioned criterion, $L = 2[kr]$, the highest frequency for accurate reproduction (i.e. the spatial aliasing limit) was $1.7$ kHz. Contrast was evaluated in the range $0.1–4$ kHz, and fell between $50–70$ dB in the frequency range reaching just above the spatial aliasing limit. For higher frequencies, a rapid roll-off was observed. PM is therefore subject to similar limitations as the analytic SFS approaches—it is unable to maintain sufficient contrast above the spatial aliasing limit of the source array. A large number of sources is therefore required, or the sound zone reproduction may be limited to low frequencies.

Compared to the analytic SFS methods, PM does not require pre-filtering to compensate for electro-acoustic responses of the system operating in a reflective room. Such compensation can be handled by the sound zone filters calculated based on the measured transfer function data—this facilitates implementation. Example sound zone implementations of PM in rooms include line arrays [Olivieri et al., 2013; Simón-Gálvez and Elliott, 2013; Simón-Gálvez et al., 2014] and circular arrays [Coleman, 2014, Chapter 3]. The physical implementations were shown to be subject to similar performance limitations as discussed above [Coleman, 2014, Chapter 3].

Hybrid techniques

Apart from the analytic and least-squares approaches to SFS in sound zoning, some hybrid techniques have also been proposed. Chang and Jacobsen [2012] modified the PM cost function, allowing to trade between the accuracy of sound reproduction in the bright zone and the effectiveness of sound energy attenuation in the dark zone using a weighting factor. Cai et al. [2014a] recently proposed a method based on a similar concept, referred to as SFR-ACC (SFR—sound field reproduction), where the sound pressure is optimized in the bright zone with the constraint of achieving a given minimum acoustic contrast. SFR-ACC has been shown to perform very similarly to Chang and Jacobsen’s weighted PM [Cai et al., 2014a]; however, the former has the advantage of being more computationally efficient [Cai et al., 2014a]. Møller et al. [2012] proposed another cost function, which trades between the reproduction of the desired bright zone sound field and the sound energy difference between the zones. Although
relaxing the sound field constraints allows the above methods to outperform PM in terms of contrast [Chang and Jacobsen, 2012, 2013; Cai et al., 2014a; Møller et al., 2012], the methods are still subject to the performance limitations discussed above.

2.4 Possible solutions to source limitation and room reflection problems

The discussion and results presented above indicated that a limited number of sources and the presence of room reflections are particularly challenging practical problems in sound zone reproduction. For example, limiting the number of sources can significantly reduce the produced acoustic contrast. This is shown in Table 2.2 which contains example results from the literature where arrays with different number of sources were compared under the same conditions. It can be noticed that a reduction in the number of sources always results in a contrast decrease. For instance, Wu and Too [2012] report a drop from 18 dB to 4 dB when reducing the number of sources from thirty-two to eight. In the former case, the contrast exceeds the lower minimum system requirement (11 dB), whereas in the latter case this limit is not reached.

Similar performance deterioration is observed when the system is subject to increasingly reflective conditions. Experimental results reported in the literature show that the lower minimum contrast (11 dB) is attainable and can be exceeded by most physical systems, irrespective of acoustic conditions [Druyvesteyn and Garas, 1997; Elliott and Jones, 2006; Jones and Elliott, 2008; Chang et al., 2009a; Park et al., 2010; Shin et al., 2010; Jacobsen et al., 2011; Cheer, 2012; Simón-Gálvez et al., 2012]. However, the upper minimum contrast required for the system’s versatility (39 dB) has only been approached in anechoic or acoustically treated rooms [Shin et al., 2010; Jacobsen et al., 2011; Coleman, 2014], or with the zones located close to the sources [Elliott and Jones, 2006; Cheer, 2012] where the direct-to-reverberant energy ratio (DRR) is large. This indicates the damaging effect of room reflections on performance. The amount of contrast deterioration due to reflections has been quantified in studies that evaluated the same systems both under anechoic and reflective conditions [Elliott and Jones, 2006; Jacobsen et al., 2011; Elliott et al., 2012; Cheer, 2012]. These results are summarized
Chapter 2. Background and literature review

<table>
<thead>
<tr>
<th>Study</th>
<th>Frequency</th>
<th>Acoustic environment</th>
<th>No. of sources</th>
<th>Acoustic contrast (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones and Elliott [2008]</td>
<td>1 kHz</td>
<td>Simulated anechoic</td>
<td>2</td>
<td>42</td>
</tr>
<tr>
<td>Elliott et al. [2012]</td>
<td>1 kHz</td>
<td>Simulated anechoic</td>
<td>3</td>
<td>52</td>
</tr>
<tr>
<td>Wu and Too [2012]</td>
<td>1 kHz</td>
<td>Simulated reflective</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Coleman et al. [2014b]</td>
<td>0.1–4 kHz</td>
<td>Reflective room</td>
<td>5</td>
<td>10 (mean)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>21 (mean)</td>
</tr>
</tbody>
</table>

Table 2.2: Comparisons of sound zone systems with a different number of sources reported in the literature. Contrast values are based on figure readings. In all examples, the acoustic contrast control method was used.

in Table 2.3, showing that the lowest contrast decreased by at least 5 dB when reflections occurred. For systems just meeting the minimum requirements under anechoic conditions, such a change may lower the contrast to perceptually unacceptable levels.

The above examples show that the contrast performance deteriorates significantly as the number of sources is reduced and with decreasing DRR in a reflective room. A number of methods that can potentially reduce the contrast loss due to these factors have been discussed in the sound zone literature. An outline of these methods is presented below, considering their applicability to systems with a limited number of sources implemented in a domestic room.

2.4.1 Limited number of sources

One of the basic means of improving the performance of a system with a limited number of sources is the optimization of source positions. Elliott and Jones [2006] examined how spacing between the sources influences the acoustic contrast produced by a pair of ACC-weighted free-field monopoles, arranged to simulate a personal audio system for listeners in two adjacent aircraft seats. Compared to a widely-spaced array, compact
2.4. Possible solutions to source limitation and room reflection problems

<table>
<thead>
<tr>
<th>Study</th>
<th>Source array</th>
<th>Frequency range (kHz)</th>
<th>Acoustic environment</th>
<th>Acoustic contrast (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliott and</td>
<td>Two loudspeakers face to face</td>
<td>0.05–1</td>
<td>Anechoic room</td>
<td>&gt;28</td>
</tr>
<tr>
<td>Jones [2006]</td>
<td></td>
<td></td>
<td>Reflective room</td>
<td>&gt;22</td>
</tr>
<tr>
<td>Jacobsen et al.</td>
<td>Sixteen point sources around zones</td>
<td>0.1–1</td>
<td>Simulated anechoic</td>
<td>&gt;27</td>
</tr>
<tr>
<td>[2011]</td>
<td></td>
<td></td>
<td>Simulated reflective</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Elliott et al.</td>
<td>Three point sources in a line</td>
<td>0.1–2</td>
<td>Simulated anechoic</td>
<td>&gt;9</td>
</tr>
<tr>
<td>[2012]</td>
<td></td>
<td></td>
<td>Simulated diffuse (20% influence)</td>
<td>&gt;4</td>
</tr>
<tr>
<td>Cheer [2012]</td>
<td>Eight point sources in four compact arrays(^1)</td>
<td>0.1–0.7</td>
<td>Simulated anechoic</td>
<td>&gt;30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Simulated reflective</td>
<td>&gt;25</td>
</tr>
</tbody>
</table>

Table 2.3: Comparisons of sound zone systems under anechoic and reflective conditions reported in the literature. Contrast values are based on figure readings. In all examples, the acoustic contrast control method was used.

sources resulted in a broader directivity null that encompassed all of the dark zone sensors, thus improving contrast. Since only a single dark zone was considered, the endfire orientation of the sources with respect to that zone was suitable as it produced a single wide null. In further study, Jones and Elliott [2008] demonstrated contrast gains with non-endfire arrays whose orientation was optimized empirically to split the null between two dark zones. It was also observed that for large contrast the bright zone must be located away from a null. For a fixed zone layout, the optimization of the radiation pattern must therefore consider both the bright and the dark zone. Although these results provide a useful indication of the optimal source positioning for small systems located in the free field, similar optimization of directivity for reflective environment has not been addressed in the literature.

Coleman [2014, Chapter 6] employed numerical optimization to select the best-performing subsets of sources from a superset of sixty sources arranged on a circle around the zones.

\(^1\)Example configuration. Similar performance trends were observed for other source arrangements at low frequencies.
in an acoustically treated studio [cf. Coleman et al., 2014b]. The subsets were selected using four different objective functions: acoustic contrast, array effort, condition number of the inverted matrix and planarity in the bright zone. Three methods were examined: ACC, PC, and PM. For smaller target number of sources (up to fifteen), with acoustic contrast as the objective function, the optimal subsets formed a larger cluster close to the dark zone with a small number of isolated source singles or couples located elsewhere on the circle. This resulted in a radiation pattern that was optimized for the given zone layout. Similar results were obtained for the planarity objective function. The effort and condition number criteria resulted in more distributed source geometries. Although these results demonstrate distinct trends in the optimal source selection, the choice was restricted to circular superset of sources around the zones. Furthermore, they may be specific to the examined reflective environment. A systematic study on source optimization with room reflections is therefore required to explore a range of geometries and determine more generally applicable configurations that improve performance.

Radmanesh and Burnett [2013b] examined source optimization for the PM method specifically. They showed that the number of sources required for accurate multi-zone reproduction can be reduced by employing numerical optimization to select the sources prior to the application of the PM technique, proposing a so-called Lasso-LS method (Lasso—least absolute shrinkage and selection operator, LS—least squares). The optimization result with this method is specific to the selected positions of the virtual sources (e.g. the plane wave sources); hence, changing their location requires re-optimization and source rearrangement. Another disadvantage of the method is that it relies on numerical optimization to select the best source positions for a given system and specific acoustic conditions—it does not offer general source optimization techniques based on the analysis of the system’s behaviour in the reflective sound field. Nevertheless, the algorithm has been shown to be a useful extension to conventional PM. For instance, it reduced the reproduction error in the dark zone above the plane of the source array compared to the typical PM technique [Radmanesh and Burnett, 2013a].
2.4.2 Room reflections

In the following, the main existing strategies for mitigating the effect of room reflections on sound zone performance are outlined. The strategies fall broadly into two categories: methods designed specifically for analytic sound field synthesis techniques and more generally applicable approaches, which have been investigated primarily for the energy control and least-squares PM methods.

Analytic sound field synthesis methods

A number of methods have been proposed for the analytic sound field synthesis methods to compensate for the mismatch between the ideal (anechoic) and actual transfer responses (which normally include the room effect).

López et al. [2005] proposed a method that compensates for the non-ideal electro-acoustic responses for WFS. The method is based on the MIMO (Multiple-Input Multiple-Output) acoustic filter structures. The responses between the sources and sensors are measured and equalized accordingly via direct system inversion. The main disadvantage of this approach is that the compensation is restricted to sensor locations, and so the control may not be consistent throughout the zone, particularly at higher frequencies. Spors et al. [2007] proposed a solution to this problem by introducing wave domain transformations in the sound reproduction loop. This facilitates control over the whole sound reproduction area enclosed by an array of sensors. The method, referred to as the wave domain adaptive filtering (WDAF), decouples the compensation filters, room impulse responses and the assumed free field propagation responses. It is particularly suited for large sound reproduction systems, with many sources and setup sensors. The transformations make the method more computationally efficient compared to the MIMO methods utilising the sound pressure signals directly, and eliminate the risk of conditioning problems in the matrix inversion that arise from the coupling of the source-sensor responses. A similar technique that does not involve adaptive filters has also been discussed in earlier work [Spors et al., 2003].

Betlehem and Abhayapala [2005] proposed an SFS method that extends the sound field mode matching approach directly to reflective sound fields. The method minimizes the
error between the desired and the reproduced sound field using the least-squares approach. The sound fields are described by means of the spatial harmonics expansion, which distinguishes the method from PM that utilizes the sound pressure information directly. The method requires estimation of the sound field coefficients in the chosen sound reproduction area. The estimation can be based on the sound pressure measurements using a circular array, or a double array of sensors, enclosing the chosen area. The disadvantage is that a large array of sensors may be required—the number of array elements increases with the reproduction region and the frequency.

A different approach to the problem of room reflections was proposed by Poletti et al. [2010] who extended the spatial harmonic expansion analysis to sources with first-order (frequency-independent) directivities. It was demonstrated that the sound radiation outside the spherical source array can be reduced by using radially-oriented first-order sources. This improved the DRR with respect to the monopoles, thus reducing the impact of reflections on performance. In later work, Poletti et al. [2011] showed that, up to the exterior Nyquist frequency of the array, a spherical array with variable-directivity sources can suppress the external sound field much more effectively than the fixed-directivity array. The concept of variable-directivity sources may however be difficult to implement in practice.

**Sound energy control and least-squares pressure matching methods**

A number of studies proposed methods for mitigating the effect of reflections on performance for the energy control and least-squares PM methods.

Elliott et al. [2012] used regularization to improve the robustness to the uncertainties in transfer functions due to a diffuse sound field [cf. Elliott and Cheer, 2011]. The method was shown to improve the acoustic contrast by up to 6 dB at low frequencies for a simulated three-source array of monopoles based on ACC. The proposed regularization technique is informed by the diffuse sound field properties, and so it is only valid for such sound fields. However, a more general regularization can in principle improve performance in any reflective sound field. For instance, a frequency-dependent regularization can be used to keep the array effort below a certain allowed limit [Coleman...
et al., 2014a]. This would reduce the radiation from the array at problematic frequencies, increasing the system’s robustness to reflections. However, contrast achieved for the direct sound could deteriorate [Coleman et al., 2014a].

A reduction in radiated sound energy can also be achieved by using a frequency-dependent source spacing that is optimized for each frequency band of interest. This was investigated by Takeuchi and Nelson [2002] for a binaural crosstalk cancellation system, which is conceptually similar to a sound zone system—here little “zones” are to be created around the listener’s ears, each containing a signal from either channel of the binaural reproduction. In fact, a 2×2 system crosstalk cancellation system has a solution identical to ACC [Park et al., 2010] and PM [Simón-Gálvez et al., 2012]. Takeuchi and Nelson [2002] proposed a so-called optimal source distribution (OSD) principle, which increases the spacing between sources as the frequency decreases. This greatly reduces radiation from the array, increasing DRR and thus making it more robust to reflections. The effect is similar to regularization, however here the performance is not compromised at problematic frequencies where the regularization would be large. The disadvantage of the system is that it requires dividing the frequency spectrum into bands that are assigned to source pairs with different spacings. This complicates the system’s implementation. Similar conclusions about the relationship between source spacing and radiation from arrays used for crosstalk cancellation were reported by Kirkeby et al. [1998a] and Ward and Elko [1998].

Chang and Jacobsen [2012] considered limiting the system’s radiation by surrounding the bright zone with a double-layer circular array of sources and defining a distributed dark zone outside the array. The double-layered sources allowed controlling both the internal and external sound fields. For the ACC method in a simulated anechoic environment, a reduction of sound pressure level down to −30 dB was achieved outside the array compared to the bright zone. This could reduce the impact of reflections on performance in a reflective room. However, a distributed dark zone may be inferior to a locally defined zone in terms of the achieved contrast—normally, a larger contrast can be created for a smaller dark zone [Møller and Olsen, 2011, Chapter 5].

Simón-Gálvez et al. [2012] proposed using a line array of phase-shift sources with hyper-
caroid directivity to limit radiation to the back of the array, thus reducing the impact of back wall reflections on performance. An eight-source array, controlled with ACC or PM, was able to reduce back radiation down to approximately \(-20\) dB with respect to the main lobe in an anechoic room. The investigations were extended to a reverberant sound field for a 4\(\times\)8 (four line arrays of eight sources stacked upon each other) array of hypercarioiroid sources controlled by PM [Simón-Gálvez et al., 2014]. The source weights were calculated for the anechoic environment. Much of the control over the directivity was lost under reflective conditions, with the minima occurring in different directions than for anechoic directivities. However, in most cases a reduction of at least \(-10\) dB with respect to the peak was achieved at the back of the array. Such a reduction is likely to be attributed to the directivity of the sources, however the performance of an omnidirectional array of sources was not compared.

By considering reflections in the source weight optimization process, methods such as ACC or PM will attenuate or cancel them at the setup sensors in the dark zone. This process is subject to the same limitations as local dereverberation (cancelling or reducing any reflected signals at listening positions) in room equalization [Fielder, 2003] or active noise control [Nelson and Elliott, 1992, Chapter 11]. A common problem is a rapid drop of performance with increasing distance from the setup points [Elliott and Nelson, 1989; Radlović et al., 2000; Talantzis and Ward, 2003]. This can be alleviated by controlling both pressure and pressure gradient [Elliott and Garcia-Bonito, 1995; Asano et al., 1996], adding sensors [Elliott and Nelson, 1989; Lopez et al., 1999; Tseng et al., 2000], preconditioning room responses (smoothing) [Hatziantoniou and Mourjopoulos, 2000, 2004], or geometrical optimization to improve the match between the direct and reflected wavefronts at the setup points [Howe and Hawksford, 1991; Guo and Pan, 1998]. Of the listed methods, geometrical optimization offers the most practical benefits, as it may improve attenuation away from the setup locations without additional equipment or signal processing. Guo and Pan demonstrated this for an active noise control system based on a pair of widely spaced sources aligned with a single sensor, operating in the presence of a rigid reflecting plane. Of a number of examined system geometries, the largest dark zone was produced by configurations for which the direct and reflected wavefronts matched well close to the sensor, i.e. when the source-sensor
line was parallel to the plane and located no further than half a wavelength from that plane, or when the configuration formed the right angle with the plane. By analogy, wavefront matching through source position optimization could facilitate attenuation of reflections for small sound zone systems. The right-angle alignment is particularly relevant to the general sound zone problem, as it does not restrict the source and dark zone locations to be near the surface. It is therefore worthwhile to explore the technique’s applicability to sound zone systems, by examining its influence on acoustic contrast and extending the investigations to compact source arrays.

2.4.3 Discussion

The review of methods for improving the performance of sound zone systems presented above provides some useful conclusions for the main topic of this thesis, which is the development of methods for improving performance of sound zone systems with a limited number of sources in a reflective environment.

The first set of conclusions is related to sound energy control and least-squares PM methods. When the number of available sources is limited, performance of such methods can be improved by optimizing the geometry of the source array, both in terms of source positioning and their relative spacing. The optimization produces array directivities that are suitable for a given zone layout [Elliott and Jones, 2006; Jones and Elliott, 2008; Coleman, 2014, Chapter 6], and may reduce radiation from the array which facilitates control in a reflective room [Takeuchi and Nelson, 2002; Kirkeby et al., 1998a; Ward and Elko, 1998]. Radiation control through regularization can also improve contrast performance in a reflective field [Elliott et al., 2012]. Directive sources could help limit radiation in certain directions [Simón-Gálvez et al., 2012], potentially improving contrast with respect to omnidirectional sources when reflecting surfaces are present. The attenuation of reflections from known directions can be enhanced by appropriate alignment of the sources and dark zone sensors with the surface [Guo and Pan, 1998], which results in an enlarged cancellation area. The geometrical alignment has the advantage of avoiding additional sensors or signal processing that may be used to achieve similar effects. Geometrical techniques could be utilized to reduce the impact
of individual reflections from particularly problematic surfaces in a room.

A number of conclusions can also be drawn for the analytic sound field synthesis methods. A range of techniques have been proposed for those methods to improve performance of practical systems in reflective rooms. These include integrating a point-wise or transform-based room compensation in the sound reproduction process [López et al., 2005; Spors et al., 2007], extending the method formulation to utilize estimated reflective sound field coefficients [Betlehem and Abhayapala, 2005], and controlling radiation outside the array enclosing the reproduction region by employing sources with fixed or variable directivities [Poletti et al., 2010, 2011]. All techniques were shown to give performance improvements in a reflective field. However, techniques for reducing the number of required sources while maintaining effective broadband sound field reproduction, similar to those outlined in Sec. 2.4.1 for the energy control and least-squares PM methods, have not been proposed for the analytic SFS methods. Such methods are therefore particularly sensitive to limitations of practical sound zone systems.

2.5 Summary

In this chapter, background to the sound zone problem was presented. The principles of a 2D sound zone reproduction were described. The main evaluation measures were also introduced: acoustic contrast to quantify acoustic separation between the zones, planarity to estimate the directional distribution of the sound field components in the bright zone, and array effort to measure the system’s efficiency.

Furthermore, the existing sound zone reproduction methods were described and discussed in the context of sound zone reproduction in a reflective room with a limited number of sources. Two main groups of methods were identified: sound energy control and sound field synthesis. The methods were shown to vary in their capability to produce the acoustic contrast and generate a sound field with the desired characteristics. For the DS, BC and SPM methods in the energy control group, relatively low contrasts were reported in the literature. These methods achieve contrast by focusing the sound energy at the bright zone. Other methods in the group: ACC, AEDM, and PC, were shown to achieve larger contrasts due to a more effective reduction of sound energy in
the dark zone. ACC and AEDM were shown to generate relatively large contrasts with a small number of sources. A useful feature of PC was the capability to restrict the range of angles for the energy flux into the bright zone, which helps reducing spatial artefacts that may occur with other methods. However, the method has not been tested with small source arrays.

The analytic sound field synthesis methods, such as HOA and WFS, were shown to be applicable to the sound zone problem. A multi-zone sound field reproduction can be achieved by translating the coefficients of the sound fields in the zones to those of the global sound field enclosed by the source array. Implementation of conventional analytic sound field synthesis methods in a reflective room is not straightforward, as it requires additional precautions such as room compensation or estimation of the reflective sound field coefficients. Furthermore, the analytic SFS methods have to be applied to regular source geometries. In contrast, SFS with the least-squares PM technique has the advantage of compensating for reflections directly and being applicable to arbitrary source geometries. However, judicious placement of the virtual sources with respect to the physical sources is required to avoid performance degradation. All SFS methods provide the capacity for precise control of the sound field in the zones, as they can control both the amplitude and phase of sound pressure. Thus, they may reduce spatial artefacts compared with the energy control methods, and may be used to render spatial sound. However, SFS methods are very sensitive to the spatial aliasing limit of the source and sensor arrays: to produce contrast effectively at higher frequencies, a large number of sources may be required (several hundreds).

Source position optimization is an effective tool for alleviating the effect of the limited number of sources and the impact of room reflections on performance for the energy control and least-squares PM methods. Appropriate source positioning can optimize the array radiation pattern for a given zone layout and limit radiation in non-target directions, which facilitates control by increasing DRR in the zones. Regularization offers similar benefits, but excessive regularization may lead to performance deterioration. Using directive sources is another useful approach to limit the influence of problematic surfaces on performance. Furthermore, appropriate alignment of the sources, dark zone, and the reflective surfaces can enlarge the attenuation area without the need of
additional sensors or signal processing. Finally, although a number of methods are available to improve the performance of analytic SFS techniques in reflective rooms, methods for circumventing the requirement for a large number of sources for successful broadband sound reproduction have not yet been proposed.

The review presented above gives an indication of the performance capabilities of different sound zoning methods when they are implemented on a system with a limited number of sources in a reflective room. The presented results show that improvements are needed to achieve the acoustic contrast that is high enough to satisfy the perceptual criterion of acceptability for a broad range of programmes and for different listeners. Methods of improving performance, such as optimization of source positions, should therefore be explored and developed. This problem will be addressed in the remainder of this thesis. Before commencing with the system optimization, it is necessary to evaluate different sound zoning methods on a practical system implemented in a room to verify and extend findings from the literature and choose the most suitable method for further optimization. Such evaluation is the topic of the following chapter.
Chapter 3

Practical implementation study

The literature review presented in the previous chapter showed that sound field control methods differ in their capacity to produce acoustic contrast and generate the bright zone of the desired characteristics in a room, using a limited number of sources. In this chapter, a practical implementation study is carried out to verify and extend the findings from the literature. Representative sound zone methods are compared on a practical system with two different source arrangements to identify the combination that best satisfies the key performance requirements: large contrast, and planar and uniform bright zone. This will allow to narrow the scope of the optimization study carried out further in the thesis.

Three methods were selected for the evaluation: DS, ACC, and PM. The first two methods belong to the energy control group. DS generates contrast through sound focusing, whereas ACC supplements sound focusing with active attenuation in the dark zone. PM was chosen as the representative method from the SFS group due to its suitability for implementation in a reflective environment. Compared to the analytic SFS methods, PM does not require additional compensation for electro-acoustic responses of the system operating a reflective room—the sound zone filters can be calculated for the actual system in operation based on the measured transfer function data, rather than for the ideal analytic conditions (see section "Least-squares pressure matching" in Sec. 2.3.2).

As discussed in Sec. 1.2.1, comparative studies of sound zone methods were carried out
predominantly under anechoic conditions [Jones and Elliott, 2008; Shin et al., 2010; Møller and Olsen, 2011; Simón-Gálvez et al., 2012; Olsen and Møller, 2013; Coleman et al., 2013a, 2014a]. Examples of comparisons in reflective rooms can be found [Wen et al., 2005; Jacobsen et al., 2011; Cheer, 2012; Coleman, 2014, Chapter 3]. However, source geometries of different characteristics were not considered in the same study in an enclosure similar to a domestic room (Cheer [2012] compared different source arrangements in a car cabin). Furthermore, the properties of bright zones produced by different methods were not compared in a reflective room using a single metric in most studies. Coleman [2014, Chapter 3] evaluated bright zone’s planarity, but the study was based on a large, sixty-element circular array of sources. Therefore, this study complements the literature with the following contribution:

- Comparison of acoustic contrast and planarity performance of sound zone methods implemented in a reflective room on two different source geometries: distributed (a circle surrounding the zones) and compact (a line in front of the zones), using a limited number of sources.

3.1 Experimental setup

In this section, the experimental setup\(^1\) is detailed. The system was installed in a large room (approximate volume 320 m\(^3\)) with some \textit{ad hoc} acoustic treatment (Basotect foam panels). The room’s reverberation times (RTs) are shown in Table 3.1. The RTs have been obtained using the standard integrated impulse response method [ISO, 2009]. An overview of the source arrays is shown in Fig. 3.1, and the source and zone layout is presented in Fig. 3.2. A linear and circular array, each used independently, were chosen to represent arrangements typically used in the literature for the sound zone methods considered (line for DS, circle for PM, and both arrangements for ACC). This allowed straightforward comparisons with the existing literature. The linear and circular arrays also represent compact and spatially distributed arrangements respectively—different

\(^1\)The system installation and response measurements were carried out by Martin Møller and Martin Olsen from Bang & Olufsen. The author’s contributions to system implementation include specification of source and zone arrangements, as well as source filter design and realization.
# 3.1. Experimental setup

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT (s)</td>
<td>0.33</td>
<td>0.30</td>
<td>0.28</td>
<td>0.27</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 3.1: Reverberation times in the room used for experimental evaluation of sound zone methods: average values for octave bands (centre frequencies).

Figure 3.1: Overview of the source arrays used for experimental evaluation of sound zone methods.

types of source arrangements and their performance characteristics could therefore be compared. The line array comprised two rows of twelve loudspeakers (10 cm × 10 cm × 16 cm) fixed side to side. One row of units was placed on top of the other and shifted to obtain 5 cm spacing between the consecutive transducers. For the circular arrangement, twenty-four loudspeakers were mounted with regular spacing (45.7 cm) on a frame of 1.75 m radius. The zones were two 53 cm × 53 cm regions, located 1.51 m above the floor. A square grid of nine microphones (sensors) mounted on a rotary stand was used to sample the sound field in the zones. The grid was placed subsequently in sixteen different locations to form a larger grid of 144 sensors in each zone (3.2–5.0 cm spacing), as shown in Fig. 3.3. The sensors were used to measure system responses that were required for source weight calculation, performance prediction and measurement.
3.2 Sound zone reproduction procedure

The sound zone reproduction scenario examined here was the reproduction of a sound programme in zone A (bright zone) and its attenuation in zone B (dark zone). For the PM method, the specified sound field was a plane wave propagating in the direction of the positive y-axis in the bright zone, and a similar plane wave attenuated by 60 dB in the dark zone (as in [Poletti, 2008]).

For the DS method, the optimal source strengths were derived analytically, based on the geometry of the arrangement and the value of the speed of sound \(c = 345 \text{ m/s}\) estimated from the environmental conditions in the room (temperature 21.5 °C, static pressure 1004.7 hPa, humidity 24%), following Cramer [1993]. For the remaining cases, the responses between sources and sensors in the zone locations were measured using the maximum length sequence (MLS) technique and used to populate the transfer function matrices for zone A (described by Eqs. (2.3) and (2.6)) and their zone B equivalents.
3.2. Sound zone reproduction procedure

Prior to the calculation of filter weights, the measured responses were processed in order to reduce noise, improve robustness to mismatch between setup and playback conditions, and limit the audibility of artefacts introduced to programme material by the sound zone filters, such as pre- and post-ringing [Norcross et al., 2004, 2006]. Each impulse response was cropped 0.5 s after the direct sound by multiplying it with an asymmetrical raised-cosine window, as shown in Fig. 3.4, to retain only the part of the impulse that is above the noise floor. Subsequently, the impulse “smoothing” procedure was applied\(^2\). Each impulse was processed using a low-pass filter (LPF) with the coefficients changing dynamically, so that the filter’s cut-off frequency decreased across the impulse in steps. The chosen steps were assigned to consecutive impulse blocks with logarithmically increasing lengths, as shown in Fig. 3.5. There were \(Z\) blocks in each impulse, and the \(z\)th block length was \(l_z = f_s i / f_z\), where \(f_z\) is the cut-off frequency of the LPF, \(f_s\) is the sampling frequency, and \(i\) is the number of frequency bins before \(f_z\). To avoid discontinuities that could be caused by large changes of the filter’s response between the blocks, relatively small 1/12 octave steps were used. The filter in each case had a fourth-order (80 dB/decade slope) Butterworth response. The smoothing process filtered out high frequency noise occurring in the impulse tails, as shown for an example impulse in Fig. 3.6.

The source weights were calculated for individual frequency bins before being combined to form a filter frequency response. The ACC solutions were regularized with

\[^2\text{The impulse response smoothing algorithm was provided by Jakob Dyreby from Bang & Olufsen.}\]
Figure 3.4: A normalised impulse response measured in the room with a superimposed cropping window.

Figure 3.5: Assignment of LPF’s cut-off frequencies to impulse blocks in the smoothing process. The block lengths increase logarithmically, with the $z$th block’s length $l_z = f_s i / f_z$, where $f_z$ is the cut-off frequency of the LPF, $f_s$ is the sampling frequency, and $i$ is the number of frequency bins before $f_z$.

a frequency-independent regularization parameter $\beta = 0.3$. For PM, a frequency-dependent regularization parameter was determined using the L-curve method [Hansen, 1992]. Each filter was then band-limited to the range 300–3500 Hz. Subsequently, the negative frequencies were populated by complex conjugation. The resulting filters were transformed to the time domain for convolution with programme material. For evaluation of performance, the filters were convolved with an MLS and the total system response, which included contributions from all sources active for a given arrangement, was measured at the sensor positions.

Since independent response measurement sets were unavailable for system setup and performance predictions, the responses from half of the sensors were used to calculate the source weights (for the ACC and PM methods), as shown in Fig. 3.7. In this
3.2. Sound zone reproduction procedure

Figure 3.6: Example impulse after smoothing. The high frequency content is significantly reduced after about 50 ms. Parameters of the smoothing process: $i = 10$, $f_s = 48$ kHz, $Z = 141$, $f_1 = 8$ Hz, $f_Z = 2.376$ kHz, frequency spacing: $1/12$ oct., LPF slope 80 dB/dec.

Figure 3.7: Sensor positions in the zones. Filled circles indicate system setup locations. For performance evaluation, all positions were used.

way, the numerically independent responses (from the remaining locations) could be included in the full response set used for predicting performance, which limited the bias [Akeroyd et al., 2007]. The full set of responses was also used when evaluating the measured performance. Sensitivity of the room response inversion to geometrical mismatch between the setup and playback locations is a known problem [Radlović et al., 2000; Talantzis and Ward, 2003]; including both setup and non-setup locations in the evaluation accounted for this sensitivity in the results.
3.3 System evaluation

In this section, the systems and methods described above are evaluated in terms of acoustic contrast and bright zone planarity. Maps of sound pressure in the bright zone are also used to qualitatively assess the degree of planarity and zone uniformity at a single frequency. Finally, the computational complexity of the methods is evaluated.

Figure 3.8 shows the predicted and measured acoustic contrast plotted against frequency. The predictions and measurements show good agreement. The ACC method implemented on the line array (ACC Line) achieved the best contrast in the majority of the considered frequency range. The highest contrast is obtained in the range 1–2.5 kHz, where it fluctuates in the proximity of 20 dB. Outside this range, contrast rolls off towards the high and low ends, but remains above 10 dB in most cases. At low frequencies, this is due to the limited directivity of the array; in the higher range the method begins to lose control over the interference between the array signals as the spatial aliasing limit (3.45 kHz) is approached. A similar contrast characteristic is observed for the DS method, but in this case the low frequency drop is more pronounced due to the limited ability to maintain a focused beam (see section “Delay and sum beamforming” in Sec. 2.3.1). The ACC method based on the circular array (ACC Circle) achieved the best contrast in the low frequency range (below 500 Hz), but the contrast dropped gradually with frequency. Here, large spacing of the sources (45.7 cm) in the circular array facilitated destructive interference in the dark zone for long wavelengths, but hindered control for shorter wavelengths. For PM, the circular array imposed an upper limit of 508 Hz for reproduction of the desired sound field across the circular region of radius 1.29 m radius, concentric to the array and extending over the zones (see section “Least-squares pressure matching” in Sec. 2.3.2). Consequently, the method was able to produce relatively high contrast up to approximately 700 Hz, where the contrast started to decrease rapidly. Interestingly, a contrast peak of 10–15 dB is observed at frequencies around 1.5 kHz. This can be attributed to an energy null due to uncontrolled destructive interference occurring locally in the dark zone.

Figure 3.9 shows the measured sound pressure maps (real part and level) in the bright
3.3. System evaluation

Figure 3.8: Acoustic contrast predicted (red) and measured (blue) for all four cases: DS (top-left), ACC Line (top-right), ACC Circle (bottom-left) and PM (bottom-right). The dashed line indicates a reference 20 dB contrast.

zone, at 1.5 kHz for all four cases. Differences in the sound field characteristics can be observed. The DS and ACC Line methods tend to generate a highly planar wave propagating from the array in the direction of the bright zone, as can be observed from the wavefronts in Fig. 3.9a. Conversely, the ACC Circle method generates a standing wave pattern with pressure amplitudes rapidly changing throughout the zone. The PM method succeeds in creating a plane wave in the specified direction (along the y-axis). This shows that the method is able to control the bright zone sound field well, even above the spatial aliasing limit. The irregular SPL pattern observed for the ACC Circle method results from the multi-directional character of the sound field generated by this arrangement. The sources located at different directions with respect to the zone generate a sound field made up of plane-wave components arriving with similar amplitude from various directions. The interfering components yield a non-uniform energy distribution across the zone: SPL differences of up to 20 dB between some of the locations can be observed in Fig. 3.9b. In contrast, planar fields generated by the DS, ACC Line, and PM methods result in relatively homogeneous zones, also shown in Fig. 3.9b. In all three cases an energy beam is formed, covering the majority of the
zone, but tending to lose intensity towards the zone margins. The energy roll-off is very rapid in the upper right corner of the zone for the DS method, where the SPL drops by nearly 20 dB in relation to the beam centre. This can be attributed to a mismatch between the actual source and zone positions and those assumed when calculating the DS source weights, which shifted the beam from the centre of the zone.

Figure 3.10 shows the planarity scores in the bright zone based on predictions and performance measurements. A close match between the two sets of results can be observed. The ACC Circle method gives the lowest overall planarity result, with the highest score just above 80% and rapid variations of planarity across frequency. This indicates that
the non-planar characteristics of the sound field generated by this method, observed in Fig. 3.9 at 1.5 kHz, are maintained at other frequencies. The DS and ACC Line methods obtain very similar planarity scores and exhibit a highly planar sound field in the range 600–3000 Hz (planarity over 80% in the majority of the range). Outside this range, two major dips can be observed in the proximity of 340 Hz and 520 Hz. These can be attributed to a strong first-order reflection from the wall at the back of the array (see Fig. 3.2). At these two frequencies, the phase relationship between the direct wave and reflection encourages strong destructive interference. Thus, the influence of the planar direct part of sound field on the planarity score is limited, as the component reflected from the side wall arrives at the zone with similar strength as the direct sound. The gradual loss of planarity above 3 kHz for both arrangements can be attributed to the increasing contribution of the reflected energy due to the side lobes generated in non-target directions as the aliasing frequency is approached (3.45 kHz). Above this frequency, the contributions from individual array elements no longer combine into one dominant plane wave propagating towards the bright zone. Instead, multiple plane wave components are generated in the directions of the grating (aliased) lobes, increasing the level of reflections and causing the loss of sound field planarity in the zone.

The planarity score obtained by the PM method is generally lower than those for DS or ACC Line, with the exception of frequencies below approximately 500 Hz where a small improvement with respect to the other two methods is achieved. The planarity dips at 340 Hz and 520 Hz are less pronounced, which can be attributed to the fact that here the direct (plane) wave is directed along the side wall rather than towards it as it was the case for DS and ACC Line, thus reducing the energy of the first order reflections which would lower the planarity. A relatively low planarity score outside the lowest frequency range can be related to the previously discussed limitations of the circular array above the spatial aliasing frequency.

It is also worth considering the computational complexity of the examined methods. Table 3.2 the central processing unit (CPU) times used to calculate the sound zone filters for the examined methods and configurations. The calculations were carried out in the frequency domain (see Section 3.2) across 32768 frequency bins using the MAT-
Chapter 3. Practical implementation study

Figure 3.10: Bright zone planarity predicted (red) and measured (blue) for all four cases: DS (top-left), ACC Line (top-right), ACC Circle (bottom-left) and PM (bottom-right). The dashed line indicates a reference 80% planarity.

<table>
<thead>
<tr>
<th>Method</th>
<th>DS</th>
<th>ACC Line</th>
<th>ACC Circle</th>
<th>PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time (s)</td>
<td>11.4</td>
<td>1568.8</td>
<td>2515.9</td>
<td>6438.0</td>
</tr>
</tbody>
</table>

Table 3.2: CPU times required to calculate the sound zone filters for the examined methods.

LAB software. As expected, the DS method uses proportionally little computation time as it only involves basic arithmetic operations. The computation time is significantly longer for ACC due to increased computation cost of the matrix inversion and eigenvalue decomposition operations included in the method. The efficiency of performing these operations is data-dependent, which demonstrates itself by a longer CPU time required to derive the filters for ACC Circle compared to ACC Line. The poorest performance is observed for the PM method—the long computation time in this case can be attributed primarily to the more complex regularization method compared to ACC, which involved parameter optimization using the L-curve technique (see Section 3.2).
3.4 Discussion

The results presented above agree with the conclusions previously drawn from the review of the sound zone literature. First, DS produced lower contrast than ACC implemented on the same array, which indicates that simple sound focusing is inferior to active attenuation of sound in the dark zone in sound zone generation. Jones and Elliott [2008] and Coleman et al. [2014a] drew similar conclusions when comparing the performance of another sound focusing method—BC—with ACC. Second, ACC implemented on the circular array produced a bright zone with low planarity and large variations of SPL, which makes acoustic contrast location-dependent and could therefore affect the quality of listening experience. This problem was also observed in the system implemented by Coleman [2014, Chapter 3]. Third, PM produced the desired sound field in the bright zone, but was unable to generate large contrast above the spatial aliasing limit of the source array; this was also observed by Coleman et al. [2014a] and Coleman [2014, Chapter 3]. As a result, ACC implemented on the same array was able to produce larger contrast over a wider frequency band. A similar observation was made by Cheer [2012, Chapter 7] in a small rectangular enclosure that imitated a car cabin. The important new observations from this study concern the ACC’s performance on the line array:

- ACC implemented on a compact line array in a reflective room gives an overall better contrast than ACC implemented on a circular array surrounding the zones with the same (limited) number of sources,

- ACC implemented on a compact line array in a reflective room results in higher bright zone planarity and much improved SPL uniformity across that zone compared to the ACC implemented on a circle. These performance characteristics are comparable with DS implemented on the same line array, or PM implemented on a circle.

These observations lead to the conclusion that ACC implemented on the line array is the most suitable method for a practical system with a limited number of sources, as it is able to produce the large contrast in a reflective room, while maintaining a relatively
high planarity and avoiding large variations of SPL across the bright zone. However, the acoustic contrast achieved by this configuration was in the range of approximately 10–25 dB, which signals that further improvements through system optimization are required to satisfy the upper minimum contrast requirement defined in Sec. 2.2.1 (39 dB). Although a relatively large computational cost of ACC compared to DS may prove challenging in real-time implementations (e.g. based on adaptive filters), it is not a significant concern if the sound zone filters are calculated offline as in the study presented above and the majority of implementations discussed so far in the literature.

3.5 Summary

The study presented in this chapter compared three representative sound zone methods—DS, ACC, and PM—when implemented on linear (DS and ACC) and circular (PM and ACC) source arrangements in a reflective enclosure similar to a domestic room. The main aim of the study was to identify the method and source geometry that is most suitable for sound zone reproduction with a limited number of sources in a reflective room, in order to provide focus for the system optimization study further in the thesis.

The ACC method implemented on the line array (ACC Line) produced overall the highest contrast, exceeding 20 dB in the 1–2.5 kHz frequency range. DS was shown to be more sensitive to the limitations of the array aperture at low frequencies than ACC—the sound focusing was limited in this frequency region and the contrast was affected. ACC implemented on the circular array (ACC Circle) generated the highest contrast at low frequencies, but the contrast dropped gradually with frequency, which can be attributed to a relatively large spacing of the sources. PM was unable to generate acoustic contrast effectively above the spatial aliasing limit of the circular array for the considered sound reproduction region.

The inspection of the SPL maps in the bright zone at the chosen frequency in the mid-range (1.5 kHz) allowed initial examination of the characteristics of the sound energy distribution in that zone. The DS and ACC Line generated a relatively uniform zone, with the sound energy centred about the main beam of radiation. PM resulted in a very good zone uniformity. ACC Circle yielded highly irregular sound energy pattern
with multiple peaks and dips—up to about 20 dB differences in SPL could be observed across the zone.

ACC Circle produced the lowest overall planarity in the bright zone, with the highest score just above 80% and rapid variations of planarity with frequency. This indicated that a lack of sound energy uniformity, previously observed on the SPL map, extended to different frequencies. The DS and ACC Line produced high planarity scores, reaching over 80% in the majority of the frequency range. Two profound dips in the planarity response were attributed to the influence of strong first order reflections from the neighbouring walls. PM was able to alleviate the effect of those reflections, but produced a lower planarity than DS and ACC Line in the mid- and high-frequency range due to the limited capability of controlling the sound field above the spatial aliasing frequency.

DS and PM were shown to be most and least computationally efficient respectively, with ACC ranking between these two methods. The relatively poor efficiency of PM was attributed to a more complex regularization technique used.

ACC Line generated overall the largest contrast, while maintaining high bright zone planarity and avoiding large variations of sound energy across that zone. This indicates that ACC implemented on a compact line array is the most suitable sound zone reproduction strategy for practical systems with a limited number of sources in domestic rooms; however, further improvements through system optimization are required to satisfy the perceptual acceptability criteria. Therefore, in the following chapters of this thesis, the ACC method implemented on a compact line array with a small number of elements is thoroughly examined using analytic and numerical methods, with the aim of optimizing its performance using geometrical means.
Chapter 4

Optimization of source positions: analysis

Chapter 3 focused on the examination of representative sound zone methods from the sound energy control and SFS group. The DS, ACC, and PM methods were implemented on the linear (compact) and circular (distributed) source arrangements in a reflective room. The ACC implemented on the line array ranked best in terms of acoustic contrast and maintained high planarity of the bright zone's sound field. Confirming the findings from the literature review in Chapter 2, the method was found to require optimization to increase the produced acoustic contrast towards the upper required minimum defined in Sec. 2.2.1 (39 dB). The optimization problem for the ACC method based on a compact array is examined in this chapter.

In Sec. 2.4, the small number of available sources and the influence of room reflections were identified as significant factors affecting the performance of practical sound zone systems. Of the possible remedial strategies discussed, source optimization was identified as a potentially effective approach to address both problems. Other methods for improving performance under reflective conditions, such as regularization and using directive sources, can be complementary. The existing source optimization studies were identified to require extension: they were limited to anechoic conditions [Elliott and Jones, 2006; Jones and Elliott, 2008], provided results that are difficult to generalize to other rooms or candidate source arrangements [Coleman et al., 2014b; Coleman,
2014, Chapter 6], or required additional examination of the effect of source positioning on acoustic contrast and extension to compact arrays [Guo and Pan, 1998]. It can be concluded from the above discussion that there is a need for a systematic study on the effect of reflections on the performance of sound zone systems, in order to provide source optimization solutions which can be applied to a wide range of rooms.

Devising source optimization techniques that are not room-specific may not be a trivial task, since the physical and acoustic parameters of domestic rooms used for sound reproduction vary in different rooms. However, certain assumptions can be made about parameters of an “average” room. For instance, Holman and Green [2010] conducted measurements in 572 home theatre listening rooms in the USA, reporting a mean volume of approximately 96 m$^3$ and reverberation time of 0.39 s in the 1 kHz band. Such data is a valuable indicator of the physical and acoustic parameters of enclosures used for sound reproduction. A similar study carried out by Díaz and Pedrero [2005] in Spain reported an average living room volume of 46.2 m$^3$ and the mean reverberation time of 0.42 s in the 1 kHz band. Similar values were earlier reported by Burgess and Utley [1985] from the measurements of forty-seven British living rooms—the average volume of 39 m$^3$ and the reverberation time of 0.31 s in the 1 kHz band. The authors also cited figures from other measurements carried out between 1950s–1970s, noting lower reverberation values for modern living rooms at the time. This demonstrates certain trends in the parameters of domestic enclosures—they are generally relatively small and acoustically “dry” spaces, i.e. with low reverberation time. This indicates that the sound field is dominated by the direct sound and low order (early) reflections from the surfaces located relatively close to the listener. As pointed out by Walker [1998], in fact the main influence of the reflected energy in a typical room is from first order reflections; the sound forming the second order reflections normally travels far enough to be attenuated by at least 10 dB with respect to the direct sound due to the spreading loss (except for the reflections from the floor and ceiling) [cf. Walker, 2007]. Strong uncompensated first order reflections can be very damaging to the performance of sound reproduction system, as demonstrated by Sæbo [2001, Chapter 4] for a crosstalk cancellation system. Furthermore, the domestic rooms are normally shaped as parallelepiped rectangles, with the main source of absorption and scattering
being furniture and decorative elements [Díaz and Pedrero, 2005]. The regular shape and uneven distribution of absorption (e.g., a bookshelf covering one wall, with another wall exposed) make the acoustic field in rooms far from perfectly diffuse, particularly at low frequencies [Kuttruff, 2009, Chapter 5, pp. 115]. It is therefore of practical importance to investigate source optimization techniques for sound zone systems operating in non-diffuse sound fields, specifically addressing the problem of the strongest low-order reflections. Such a study is the subject of this chapter.

In the following, the source optimization problem is studied for a 2×2 system (two sources and two setup sensors) with a single surface. With the reduced complexity of the system, a fundamental relationship between performance and system geometry can be examined analytically. The expressions for the key performance metrics—acoustic contrast and array effort—are derived. The expressions provide the algebraic framework for the formulation of the source positioning optimization techniques. The study is of practical importance for the following reasons:

- examination of the performance characteristics of the 2×2 system underpins the analysis of larger systems,
- examination of a single surface scenario can provide an insight into fundamental problems related to strong low order reflections occurring in rooms.

At the beginning of this chapter, the acoustic contrast and array effort expressions are derived. First, the contrast expression is analysed to establish source positions that maximize contrast through the reduction of sound pressure in the dark zone. This results in three techniques that can be used to minimize the reflected sound pressure in that zone. Second, the geometrical requirements for increasing the efficiency of radiation into zone A are defined. Better efficiency is important in practice, as it enhances contrast through increased contribution of the bright zone’s pressure and lowers the control effort, thus reducing the risk of equipment fatigue or damage.

The specific contributions of this chapter can therefore be summarized as follows:

- Derivation of the acoustic contrast and array effort expressions for a 2×2 system
based on the ACC method, in a single surface scenario, which describes the fundamental relationship between the system’s geometry and performance.

- Proposing three source optimization techniques to increase the acoustic contrast, derived from the analysis of the 2×2 system: Null-Split, Far-Align, and Near-Align.

- Defining geometrical requirements for increasing the efficiency of radiation into the bright zone for enhanced acoustic contrast and reduced array effort.

### 4.1 Sound zone system under analysis

Fig. 4.1 shows the sound zone system under analysis. The elementary geometrical scenario of two-dimensional sound zones and reflections propagating in the same plane is considered, as discussed in Sec. 2.1. Zone A is the bright zone controlled by a single setup sensor and evaluated using N monitor sensors. The geometrical centre of the monitor sensor array coincides with the setup sensor. Zone B is the dark zone, defined similarly. There are two monopole sources whose spacing \(d\) is small compared to the distances between the sources and any of the sensors, which allows application of the well-known far-field approximation [Nelson and Elliott, 1992, Chapter 8, pp. 233]. The reflections are simulated using the image-source model (ISM) technique [Allen and Berkley, 1979]: each source is paired with an image generated by the reflecting, infinitely-long rigid surface with a magnitude of reflection coefficient, \(0 \leq \gamma \leq 1\). Therefore, the total sound field consists of the contributions from both the physical and image arrays. The distances between the source array centre and the sensors in zone B are defined as \(r_B\) and \(s_{Bn}\) for the setup and the monitor sensors respectively, where \(n = 1, 2, ..., N\). The angles between the array axis and the lines between the source array centre and the sensors in zone B are defined as \(\phi_B\) and \(\theta_{Bn}\) for the setup and monitor sensors respectively. For the image source array, the distances and angles are defined by adding the superscript ′ to the quantities defined for the physical array. Transfer functions between each source and the setup sensor in zone B, without the surface, can
4.1. Sound zone system under analysis

be written at a single frequency as

\[ G_{B1} = K \left( \frac{e^{-jk(r_B + d \cos \phi_B/2)}}{r_B} \right), \]
\[ G_{B2} = K \left( \frac{e^{-jk(r_B - d \cos \phi_B/2)}}{r_B} \right), \] (4.1)

and for the case with the surface present

\[ G_{B1} = K \left( \frac{e^{-jk(s_B + d \cos \theta_B/2)}}{s_B} + \gamma e^{-jk(s_B + d \cos \theta_B'/2)}}{s_B'} \right), \]
\[ G_{B2} = K \left( \frac{e^{-jk(s_B - d \cos \theta_B/2)}}{s_B} + \gamma e^{-jk(s_B - d \cos \theta_B'/2)}}{s_B'} \right), \] (4.2)

where \( K = j \rho c k / 4 \pi \) in which \( \rho \) is the air density, \( c \) is the speed of sound, and \( k \) is the wavenumber proportional to frequency \( (k = 2\pi f/c) \). By analogy, transfer functions between each source and the \( n \)th monitor sensor in zone B, without the surface, are

\[ \Omega_{Bn1} = K \left( \frac{e^{-jk(s_B + d \cos \theta_B/2)}}{s_B} \right), \]
\[ \Omega_{Bn2} = K \left( \frac{e^{-jk(s_B - d \cos \theta_B/2)}}{s_B} \right), \] (4.3)

and with the surface present

\[ \Omega_{Bn1} = K \left( \frac{e^{-jk(s_B + d \cos \theta_B/2)}}{s_B} + \gamma \frac{e^{-jk(s_B + d \cos \theta_B'/2)}}{s_B'} \right), \]
\[ \Omega_{Bn2} = K \left( \frac{e^{-jk(s_B - d \cos \theta_B/2)}}{s_B} + \gamma \frac{e^{-jk(s_B - d \cos \theta_B'/2)}}{s_B'} \right), \] (4.4)

It is convenient to define the sound pressure at each sensor using the principle of superposition, as discussed in Sec. 2.1. The complex sound pressure at the setup sensor in zone B (at a single frequency) can therefore be written as \( p_B = G_B q \), where \( G_B = [G_{B1}, G_{B2}] \) and \( q = [q_1, q_2]^T \) is the vector of complex source weights. For monitor sensors in zone B, observed sound pressures form the vector \( o_B = \Omega_B q \), where

\[ \Omega_B = \begin{bmatrix} \Omega_{B11} & \Omega_{B12} & \cdots & \Omega_{B1N} \\ \Omega_{B21} & \Omega_{B22} & \cdots & \Omega_{B2N} \end{bmatrix}^T. \] (4.5)

For zone A, the distances, angles, transfer functions, and sound pressures are defined similarly. The relationship between the transfer functions defined above and the general transfer functions is described in Appendix A.
4.2 Optimal source weights

The ACC problem can be solved analytically using the procedure outlined in Sec. 2.3.1. For a general 2×2 system, a regularized solution must be derived (\(G_B^H G_B\) is singular). Equation (2.24) was therefore used to find the elements of the unscaled optimal source weight vector \(\hat{q}\). The ratio of these elements formed an expression that included terms in the regularization parameter \(\beta\). These terms were neglected assuming \(\beta \to 0\), as shown in Appendix A. The resulting expression for the unscaled unit source weight vector, under anechoic conditions at the system setup stage, is

\[
\hat{q} = \frac{K}{\Lambda} \left[ \frac{e^{-jk(r_B - d \cos \phi_B/2)}}{r_B} \right],
\]

(4.6)

where

\[
\Lambda = \sqrt{2} |K|/r_B.
\]

(4.7)

Under reflective conditions at setup stage, the equivalent vector is defined as

\[
\hat{q} = \frac{K}{\Lambda} \left[ \frac{e^{-jk(r_B - d \cos \phi_B/2)}}{r_B} + \gamma e^{-jk(r_B' - d \cos \phi_B'/2)}}{r_B'} \right] + \frac{K}{\Lambda} \left[ \frac{e^{-jk(r_B + d \cos \phi_B/2)}}{r_B} - \gamma e^{-jk(r_B' + d \cos \phi_B'/2)}}{r_B'} \right],
\]

(4.8)
4.3 Acoustic contrast expression

where

$$\Lambda = \sqrt{2}|K| \left[ \frac{1}{r_B^2} + \frac{2\gamma \cos \left( kd \left( \cos \phi_B - \cos \phi_B' \right) / 2 \right) \cos \left( k(r_B - r_B') \right)}{r_B r_B'} \right]^{1/2} \left( \frac{\gamma^2}{r_B'} \right) + \frac{\gamma^2}{r_B'}. \quad (4.9)$$

Substituting Eqs. (4.2) and (4.8) into $\hat{p}_B = G_B \hat{q}$ shows that, with negligible regularization, the array simply cancels the sound pressure at the setup sensor in zone B. The ACC method can therefore be considered as comprising direct sound and reflection cancellers when reflections are considered at the setup stage. The full derivation of the optimal source weight vectors is included in Appendix A.

4.3 Acoustic contrast expression

The vectors of complex sound pressures produced by ACC in zones A and B are defined as $\hat{\mathbf{o}}_A = \Omega_A \hat{\mathbf{q}}$ and $\hat{\mathbf{o}}_B = \Omega_B \hat{\mathbf{q}}$ respectively. Substituting into Eq. (2.8) yields the contrast expression:

$$\text{Contrast}_{AB} = 10 \log_{10} \left( \frac{||\hat{\mathbf{o}}_A||^2}{||\hat{\mathbf{o}}_B||^2} \right) = 10 \log_{10} \left( \frac{\sum_{n=1}^{N} ||\hat{\mathbf{o}}_{An}||^2}{\sum_{n=1}^{N} ||\hat{\mathbf{o}}_{Bn}||^2} \right). \quad (4.10)$$

Under anechoic conditions at setup and playback (anechoic zone generation scenario), the squared sound pressure at the nth monitor sensor in zone B, contained in the contrast expression, is defined as

$$||\hat{\mathbf{o}}_{Bn}||^2 = \left( \frac{K}{\Lambda} \right)^2 \left( \hat{O}_{Bn}^D \right)^2 = \frac{4|K|}{|\Lambda|^2} \left( O_{Bn}^D \right)^2, \quad (4.11)$$

where $\Lambda$ is defined by Eq. (4.7) as in the previous case. It is also possible to set up the system under anechoic conditions (e.g. an anechoic chamber) and introduce the surface at playback (anechoic-reflective zone generation scenario). In such a scenario, the squared sound pressure is

$$||\hat{\mathbf{o}}_{Bn}||^2 = \left( \frac{K}{\Lambda} \right)^2 \left( \hat{O}_{Bn}^D + \hat{O}_{Bn}' \right)^2 = \frac{4|K|}{|\Lambda|^2} \left[ \left( O_{Bn}^D \right)^2 + \left( O_{Bn}' \right)^2 + O_{Bn}' \right], \quad (4.12)$$
with \( \Lambda \) is defined by Eq. (4.7). Finally, under reflective conditions at setup and playback (reflective zone generation scenario), the squared sound pressure expression expands to

\[
|\hat{\omega}_{Bn}|^2 = \frac{|K|^4}{|\Lambda|^2} \left| \hat{\omega}_{Bn}^D + \hat{\omega}_{Bn}^{D'} + \hat{\omega}_{Bn}^R + \hat{\omega}_{Bn}^{R'} \right|^2
\]

\[
= \frac{4|K|^4}{|\Lambda|^2} \left[ (\hat{O}_{Bn}^D)^2 + (\hat{O}_{Bn}^{D'})^2 + (\hat{O}_{Bn}^R)^2 + (\hat{O}_{Bn}^{R'})^2 \right. \\
+ \hat{O}_{Bn}^{DD'} + \hat{O}_{Bn}^{DR} + \hat{O}_{Bn}^{DR'} + \hat{O}_{Bn}^{RR'} + \hat{O}_{Bn}^{RR'} \\
\right],
\]

where \( \Lambda \) is given by Eq. (4.9). The superscripts \( D \) and \( R \) in Eqs. (4.11)–(4.13) denote the relationship of a direct pressure component with the direct sound or reflection canceller respectively, whereas the superscripts \( D' \) and \( R' \) relate reflected components with the cancellers. Therefore, the complex sound pressure components identified in Eqs. (4.11)–(4.13) are: \( \hat{O}_{Bn}^D \) and \( \hat{O}_{Bn}^{D'} \)—direct and reflected components due to direct sound canceller respectively; \( \hat{O}_{Bn}^R \) and \( \hat{O}_{Bn}^{R'} \)—direct and reflected components due to reflection canceller respectively. The corresponding magnitude components \( O_{Bn}^D, \hat{O}_{Bn}^{D'}, \hat{O}_{Bn}^R, \hat{O}_{Bn}^{R'} \), and the interaction components \( \hat{O}_{Bn}^{DD'}, \hat{O}_{Bn}^{DR}, \hat{O}_{Bn}^{DR'}, \hat{O}_{Bn}^{RR'}, \hat{O}_{Bn}^{RR'} \) are the key terms, detailed in Table 4.1. Sound pressures at the \( n \)th monitor sensor in zone A are defined similarly.

The scaling factor \( 4|K|^4 / |\Lambda|^2 \) in Eqs. (4.11)–(4.13) is independent of the monitor sensor location, and so it does not affect the contrast (it is a common factor in the contrast expression). Note that under anechoic and anechoic-reflective zone generation scenarios, the contrast expression contains only the sound pressure components arising due to the operation of the direct sound canceller. Under the reflective scenario, additional sound pressure components arise due to the activation of the reflection canceller.

The full derivation of Eqs. (4.11)–(4.13) is included in Appendix A.

### 4.4 Array effort expression

The array effort expression for the case with the anechoic conditions at setup (anechoic and anechoic-reflective zone generation scenarios) can be obtained by substituting Eq. (4.6) to Eq. (2.11), and assuming the reference source is a monopole located at the
4.4. Array effort expression

\[ O_{Bn}^D = \sin \left( kd \left( \cos \theta_{Bn} - \cos \phi_B \right) / 2 \right) / s_{Bn} r_B \]
\[ O_{Bn}^{D'} = \gamma \sin \left( kd \left( \cos \theta_{Bn} - \cos \phi_B \right) / 2 \right) / s'_{Bn} r_B \]
\[ O_{Bn}^R = \gamma \sin \left( kd \left( \cos \theta_{Bn} - \cos \phi'_B \right) / 2 \right) / s_{Bn} r'_B \]
\[ O_{Bn}^{R'} = \gamma^2 \sin \left( kd \left( \cos \theta_{Bn} - \cos \phi'_B \right) / 2 \right) / s'_{Bn} r'_B \]

**Components describing interaction**

\[ O_{Bn}^{DD'} = 2O_{Bn}^D O_{Bn}^{D'} \cos (k(s_{Bn} - s'_{Bn})) \]
\[ O_{Bn}^{DR} = 2O_{Bn}^D O_{Bn}^R \cos (k(r_B - r'_B)) \]
\[ O_{Bn}^{DR'} = 2O_{Bn}^D O_{Bn}^{R'} \cos (k(s_{Bn} - s'_{Bn} + r_B - r'_B)) \]
\[ O_{Bn}^{DR''} = 2O_{Bn}^D O_{Bn}^R \cos (k(s_{Bn} - s'_{Bn} - r_B + r'_B)) \]
\[ O_{Bn}^{DR'''} = 2O_{Bn}^D O_{Bn}^{R''} \cos (k(r_B - r'_B)) \]
\[ O_{Bn}^{RR''} = 2O_{Bn}^R O_{Bn}^{R'} \cos (k(s_{Bn} - s'_{Bn})) \]

**Table 4.1:** Components of the squared sound pressure at the nth monitor sensor in zone B, generated by a 2x2 sound pressure canceller. Superscripts D, R, and ′ indicate relationship with the direct sound canceller, the reflection canceller, and the reflected sound field respectively. Components in zone A are defined similarly.

\[
\text{Effort}^D_A = -3 + 10 \log_{10} \left( \frac{|\Lambda|^2}{2|K|^2 \left( P_A^D \right)^2 r_A^2} \right), \quad (4.14)
\]

where \( \Lambda \) is defined by Eq. 4.7. For setup under reflective conditions, the effort expression is derived by substituting Eq. (4.8) to Eq. (2.11), with the reference monopole located at array centre as previously, which gives

\[
\text{Effort}^R_A = -3 + 10 \log_{10} \left[ |\Lambda|^2 \left( \frac{1}{r_A^2} + \frac{2\gamma \cos (k(r_A - r'_A))}{r_A r'_A} + \frac{\gamma^2}{r'_A^2} \right) \right] / \left( 2|K|^2 \left( P_A^D \right)^2 + \left( P_A^{D'} \right)^2 + \left( P_A^R \right)^2 + \left( P_A^{DD'} + P_A^{DR} + P_A^{DR'} + P_A^{DR''} + P_A^{DR'''} + P_A^{RR''} \right) \right), \quad (4.15)
\]

where \( \Lambda \) is defined by Eq. (4.9). The quantities \( P_A^D, P_A^{D'}, P_A^R, P_A^{DD'}, P_A^{DR}, P_A^{DR'}, P_A^{DR''}, P_A^{RR''} \) in Eqs. (4.14) and (4.15) are the components of squared sound pressure at the setup point in zone A, arising due to the direct sound and reflection.
cancellers. The components are defined in Table 4.2. The full derivation of Eqs. (4.14) and (4.15) is included in Appendix A.

The contrast and effort expressions can now be analysed to propose techniques for optimizing source positions.

### Magnitude components

<table>
<thead>
<tr>
<th>Component</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_A^D$</td>
<td>$\sin(kd \cos \phi_A - \cos \phi_B) / r_A r_B$</td>
</tr>
<tr>
<td>$P_A^{D'}$</td>
<td>$\gamma \sin(kd \cos \phi_A' - \cos \phi_B) / r_A' r_B$</td>
</tr>
<tr>
<td>$P_A^R$</td>
<td>$\gamma \sin(kd \cos \phi_A - \cos \phi_B') / r_A r_B'$</td>
</tr>
<tr>
<td>$P_A^{R'}$</td>
<td>$\gamma^2 \sin(kd \cos \phi_A' - \cos \phi_B') / r_A' r_B'$</td>
</tr>
</tbody>
</table>

### Components describing interaction

<table>
<thead>
<tr>
<th>Component</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_A^{DD'}$</td>
<td>$2P_A^D P_A^{D'} \cos(kr_A - r_A')$</td>
</tr>
<tr>
<td>$P_A^{DR}$</td>
<td>$2P_A^D P_A^R \cos(kr_B - r_B')$</td>
</tr>
<tr>
<td>$P_A^{DR'}$</td>
<td>$2P_A^{D'} P_A^R \cos(kr_A - r_A' + r_B - r_B')$</td>
</tr>
<tr>
<td>$P_A^{DR''}$</td>
<td>$2P_A^{D''} P_A^R \cos(kr_A - r_A' - r_B + r_B')$</td>
</tr>
<tr>
<td>$P_A^{DR'''}$</td>
<td>$2P_A^{D'''} P_A^R \cos(kr_B - r_B')$</td>
</tr>
<tr>
<td>$P_A^{RR'}$</td>
<td>$2P_A^R P_A^{R'} \cos(kr_A - r_A')$</td>
</tr>
</tbody>
</table>

Table 4.2: Components of the squared sound pressure at the setup sensor in zone A, generated by a $2 \times 2$ sound pressure canceller. Superscripts $D$, $R$, and $'$ indicate relationship with the direct sound canceller, the reflection canceller, and the reflected sound field respectively.

### 4.5 Minimization of sound pressure in the dark zone

In this section, the problem of optimizing the source positions for improved contrast is examined. The solution to the source optimization problem is dominated by the poles of Eq. (2.8) when sound pressure in zone B is close to zero. Therefore, in this section, Eqs. (4.11)—(4.13) are analysed to find source positions that minimize pressure in that zone. The optimization is for three previously considered scenarios: anechoic zone generation scenario (anechoic setup and playback), anechoic-reflective scenario...
4.5. Minimization of sound pressure in the dark zone

4.5.1 Anechoic zone generation scenario

Under this scenario, the squared sound pressure in zone B is given by Eq. (4.11), and so a minimization of a single sound pressure component $O_{Dn}^B$ is required. Component $O_{Dn}^B$ can be reduced by using a source array that is endfire with respect to the setup sensor in zone B, and located at a large distance. This results in $\cos \theta_B - \cos \phi_B \approx 0$ and therefore small values of $O_{Dn}^B$. An example configuration is shown in Fig. 4.2. Elliott and Jones [2006] demonstrated such capability in simulations for a pair of closely-spaced, free-field monopoles with endfire orientation in a similar system, as discussed in Sec. 2.4.1.

4.5.2 Anechoic-reflective zone generation scenario

In this scenario, the squared sound pressure in zone B is given by Eq. (4.12), and so a minimization of sound pressure components $O_{Dn}^B$, $O_{Dn}^{D'}$, and $O_{Dn}^{DD'}$ is required. Two different techniques for achieving this are presented below.

Null-Split technique

Component $O_{Dn}^{D'}$ can be reduced directly by appropriate orientation of the source array with respect to zone B. The orientation must be such that the difference $\cos \theta_B' - \cos \phi_B$, 

![Figure 4.2: Source positions for minimizing the direct sound pressure in zone B. Symbols: ■ source, — array axis, ● setup sensor, □ monitor sensor area.](image-url)
Chapter 4. Optimization of source positions: analysis

contained in $O_{B_n}'\$ is minimized. Figure 4.3a shows an example configuration to achieve this aim. Ensuring that $\pm \phi_B = \mp \phi_B'$, where $\phi_B \neq 0^\circ$, steers the directivity nulls of the physical and image arrays symmetrically at the setup sensor located centrally in zone B. In other words, the null generated by the physical array is split between the direct and reflected paths to the zone centre; hence, the technique is referred to as the Null-Split. The principle is similar to null-splitting between two dark zones, discussed by Jones and Elliott [2008]. Thus, the principle of operation of the technique is optimization of the array’s directivity that considers the reflection propagation path into the dark zone. When the condition $\pm \phi_B = \mp \phi_B'$ is fulfilled, the value of $\cos \phi_B$ falls within the range of values of $\cos \theta_B'$ that is small for large distances of the image array from the zone. This is shown in Fig. 4.4a (top) for the configuration from Fig. 4.3a. The values of the cosine difference are therefore relatively low: $-0.057 \leq (\cos \theta_B' - \cos \phi_B) \leq 0.053$. Since the array is no longer endfire as in the anechoic zone generation scenario, the difference $\cos \theta_B' - \cos \phi_B$ increases compared to that case—the component $O_{B_n}'D'$ is thus also larger: the example Null-Split array shown in Fig. 4.4a (bottom) yields $-0.159 \leq (\cos \theta_B' - \cos \phi_B) \leq 0.110$. The configuration is therefore suboptimal for direct sound attenuation. Since $O_{B_n}'D'$ contains $O_{B_n}'$ as a factor, the former reduces with the latter.

**Far-Align technique**

Components $O_{B_n}'$ and $O_{B_n}'D'$ can also be reduced by aligning the source array with the setup sensor in zone B and the surface, as shown in Fig. 4.3b. The array is located further from the surface than the zone; therefore, the technique is referred to as the Far-Align. As in the Null-Split case, Far-Align is based on the optimization of the array’s directivity that considers the reflection propagation path into the dark zone. When the condition $\phi_B = \phi_B'$ is met, the values of $\cos \phi_B$ are included in the range of the values of $\cos \theta_B'$, as shown in Fig. 4.4b (top) for the configuration from Fig. 4.3b. In this case, the endfire orientation of the array significantly limits the range of $\cos \theta_B'$, leading to $-0.002 \leq (\cos \theta_B' - \cos \phi_B) \leq 0$. The technique therefore allows for larger reductions of the cosine difference values compared to the Null-Split. This is also the case for with the component $O_{B_n}'$. Source array is endfire with respect to the setup sensor in zone B, and located at a large distance. As demonstrated in Figs. 4.4b for
4.5. Minimization of sound pressure in the dark zone

\[ \phi'_{B} = 34^\circ \]
\[ B \phi_{B} = -34^\circ \]

(a) Null-Split:
\[ \phi_{B} \neq 0^\circ, \pm \phi_{B} = \mp \phi'_{B} \]

(b) Far-Align:
\[ \phi_{B} = \phi'_{B} \]

(c) Near-Align:
\[ \phi_{B} \in \{0^\circ, 180^\circ\}, \phi'_{B} = \phi_{B} - 180^\circ \]

Figure 4.3: Source positions for minimizing the reflected sound pressure in zone B. Symbols: ■ source, □ image source, — array axis, ● setup sensor, □ monitor sensor area, ■ surface.

the configuration from Figs. 4.3b, this results in \(-0.030 \leq (\cos \theta_{Bn} - \cos \phi_{B}) \leq 0\) and therefore small values of \(O_{Bn}^{D} \). In consequence, the attenuation of the direct sound and reflection is effective over a larger region. Non-endfire orientations of Far-Align are also possible, although they will attenuate over a smaller area.

4.5.3 Reflective zone generation scenario

In this scenario, the squared sound pressure in zone B is given by Eq. (4.13). Therefore, a minimization of sound pressure components \(O_{Bn}^{D}, O_{Bn}^{D'}, O_{Bn}^{R}, O_{Bn}^{R'}, O_{Bn}^{DD'}, O_{Bn}^{DR}, O_{Bn}^{DR'}, O_{Bn}^{R'R'}, O_{Bn}^{D'R'},\) and \(O_{Bn}^{R'R'}\) is required to increase contrast.

The Null-Split and Far-Align techniques are both applicable in this case. In the former case, the condition \(\pm \phi_{B} = \mp \phi'_{B}\) reduces Eq. (4.8) to Eq. (4.6); thus, the reflection canceller is deactivated. Similarly, the condition \(\phi_{B} = \phi'_{B}\) deactivated the reflection canceller in the Far-Align case. The components \(O_{Bn}^{R}, O_{Bn}^{R'}, O_{Bn}^{DR}, O_{Bn}^{DR'}, O_{Bn}^{D'R'},\) and \(O_{Bn}^{RR'}\)
Figure 4.4: Ranges of angles between the array axis and the monitor sensors in zone B, and the related cosine function values for configurations from Fig. 4.3 (indicated by the filled regions).

are therefore not produced, and the components $O_{Bn}^{D}, O_{Bn}^{D'}, O_{Bn}^{D'D'}$ are reduced as in the anechoic-reflective scenario.

The reflective scenario allows for an additional optimization technique, which is presented below.

Near-Align technique

This technique, in contrast to Null-Split and Far-Align, employs the reflection canceller. The technique optimizes the canceller’s operation by spatially matching the sources and their images with respect to the dark zone. This can be achieved by aligning the source array with the zone and surface, while ensuring that $\phi_B \in \{0^\circ, 180^\circ\}$ and
4.6 Bright zone considerations

\( \phi_B' = \phi_B - 180^\circ \), as shown in Fig. 4.3c. Since the sources are located nearer to the surface than the zone, the method is referred to as the Near-Align. The alignment, the endfire orientation, and the large distance of the physical and image arrays from the zone result in \( \cos \theta_{Bn} - \cos \phi_B \approx -(\cos \theta_{Bn} - \cos \phi_B') \). This is demonstrated in Fig. 4.4c for the configuration from Fig. 4.3c. In this case, \( 1.986 \leq (\cos \theta_{Bn} - \cos \phi_B) \leq 2 \) and \( -2 \leq (\cos \theta_{Bn} - \cos \phi_B') \leq -1.970 \) is observed. The large distance of the sources and images from the zone also results in \( r_B \approx s_{Bn} \) and \( r_B' \approx s_{Bn}' \), which in combination with the cosine relationship yields \( O_{Bn}^{D'} \approx -O_{Bn}^{R} \). This approximation is fundamental to active attenuation of the reflection, since it gives \( (O_{Bn}^{D'})^2 + (O_{Bn}^{R})^2 + O_{Bn}^{RR} \approx 0 \), as well as \( O_{Bn}^{D'} + O_{Bn}^{RR} \approx 0 \) and \( O_{Bn}^{D'} + O_{Bn}^{RR} \approx 0 \), which according to Eq. (4.13) reduces the influence of \( O_{Bn}^{D'} \) and \( O_{Bn}^{R} \), and the associated components on the squared pressure in zone B. The Near-Align also yields \( \cos \theta_{Bn} - \cos \phi_B \approx 0 \), which reduces the remaining reflection-related components \( O_{Bn}^{R'} \) and \( O_{Bn}^{RR} \). The configuration from Fig. 4.3c achieves \( -0.014 \leq (\cos \theta_{Bn} - \cos \phi_B) \leq 0 \), as shown in Fig. 4.4c (top). Moreover, Near-Align remains optimal for direct sound attenuation. Configuration in Fig. 4.3c produces \( 0 \leq (\cos \theta_{Bn} - \cos \phi_B) \leq 0.030 \), as shown in Fig. 4.4c. The component \( O_{Bn}^{D} \) is therefore significantly reduced.

The properties of Near-Align indicate that an extended reflection attenuation region is produced. This is in agreement with Guo and Pan’s result for a pair of widely spaced sources aligned with the sensor, and forming the right angle with the surface [Guo and Pan, 1998], discussed in Sec. 2.4.2.

4.6 Bright zone considerations

While attenuating sound in zone B is of primary importance for producing large acoustic contrast, the influence of the above techniques on the sound pressure in zone A must also be examined. Source weights \( \hat{\mathbf{q}} \) will produce one or more peaks of SPL which should coincide with zone A to achieve the largest contrast [Jones and Elliott, 2008] and reduced control effort [Coleman et al., 2014a]. With a single surface, the peaks maintain a high DRR, and so collocating a maximum of the direct sound radiation with zone A is a valid means of enhancing contrast and reducing effort under all scenarios. Jones
Elliott [2008] observed that for a pair of free-field monopoles with the ACC weights, the position of directivity peaks changed with the array orientation with respect to the dark zone. Although these results indicate suitable zone A locations for the examined system, it is desirable to define precise geometrical requirements that must be met to increase this zone’s pressure contribution to contrast and decrease effort. To satisfy such requirements, it may be necessary to position zone A in a particular location with respect to zone B. This may not always be feasible—in certain scenarios the zone positions may be fixed in a room. However, it will be demonstrated in this section that allowing some flexibility in relative zone positioning can lead to enhanced contrast and lower effort.

When determining the required position for zone A, it is sensible to consider the geometrical centre of the monitor sensor array, which coincides with the setup sensor. This simplifies the contrast expression in Eq. (4.10) to the following form:

\[
\text{Contrast}_{AB} = 10 \log_{10} \left( \frac{|\hat{p}_A|^2}{\sum_{n=1}^{N} |\hat{p}_{Bn}|^2} \right). \tag{4.16}
\]

The direct component of the squared sound pressure \(|\hat{p}_A|^2\) at the setup sensor in zone A, \(P_D^A\), is defined in Table 4.2. Since the numerator of Eq. (4.16) contains \((P_D^A)^2\), an increase in the value of this component will result in larger contrast. Likewise, Eqs. (4.14) and (4.15) show that larger \((P_D^A)^2\) gives lower effort (the component is contained in the denominators of the logarithmic expressions). Fig. 4.5a shows \((P_D^A)^2\), normalized and plotted on a logarithmic scale against \(\cos \phi_A - \cos \phi_B\) and \(kd\), that characterize the component’s dependence on the angular distance from zone B and on frequency respectively. The range of \(\cos \phi_A - \cos \phi_B\) covers all angular distances, and \(kd\) extends up to the spatial aliasing limit of the source array. The pressure decreases by no more than 6 dB with respect to the maximum value in at least \(2/3\) of the frequency range for \(\cos \phi_A - \cos \phi_B \leq -1\) or \(\cos \phi_A - \cos \phi_B \geq 1\), i.e. outside of the greyed-out region of Fig. 4.5a. Similar observations can be made for other arbitrarily chosen thresholds. For instance, the pressure decreases by no more than 12 dB in at least \(5/6\) of the frequency range outside the greyed-out region, whereas a 18 dB decrease covers \(11/12\) of the frequency range. The key observation here is that for any chosen threshold value, the frequency range related to that threshold shrinks rapidly when the cosine difference
4.6. Bright zone considerations

Figure 4.5: (a) The normalized direct squared sound pressure component at the setup sensor in zone A, \((P_D^A)^2\) (decibel scale with a −60 dB limit). Outside the greyed-out region, \(10 \log_{10}((P_D^A)^2) \geq -6\ dB\) in at least \(2/3\) of the frequency range. (b) Ranges of \(\phi_A\) that increase the contribution of sound pressure in zone A to contrast (white regions).

falls within the greyed-out region of Fig. 4.5a. Those values of the cosine differences should therefore be avoided.

For \(-180^\circ \leq \phi_B \leq -90^\circ\) or \(90^\circ \leq \phi_B \leq 180^\circ\), the inequalities \(\cos \phi_A - \cos \phi_B \leq -1\) or \(\cos \phi_A - \cos \phi_B \geq 1\) hold if \(-\arccos\left(1 + \cos \phi_B\right) \leq \phi_A \leq \arccos\left(1 + \cos \phi_B\right)\), whereas for \(-90^\circ \leq \phi_B \leq 90^\circ\) we must have \(-180^\circ \leq \phi_A \leq -\arccos\left(-1 + \cos \phi_B\right)\) or \(\arccos\left(-1 + \cos \phi_B\right) \leq \phi_A \leq 180^\circ\). Such ranges of \(\phi_A\) are indicated by the white regions in Fig. 4.5b. Large pressures in zone A are therefore be achieved if the zone’s centre is located on the other side of the axis normal than zone B’s setup sensor, and within the required range of \(\phi_A\) that is the largest for \(\phi_B \in \{0^\circ, \pm 180^\circ\}\) (the endfire orientation with respect to the dark zone) and decreases as \(\phi_B \to \pm 90^\circ\); in the limit, the most suitable locations for zone A centre are perpendicular to the axis normal (in either direction). Note that except for this special case, for any given value of \(\phi_B\) the required range of \(\phi_A\) consists of pairs of identical values with opposite signs. This means that the centre of zone A can be suitably located on either side of the array axis.


4.7 Summary

A 2×2 system was analysed to describe the fundamental relationship between its geometry and performance. The system consisted of two sources and one setup sensor in each zone, which were surrounded by a number of monitor sensors used to evaluate the sound pressures in the zones. The system was assumed to operate in the presence of an infinitely long rigid surface generating image sources. It was also assumed that the spacing of the sources in the physical and image arrays was much smaller than the arrays’ distances to any of the sensors, which allowed applying the well-known far-field approximation in the transfer function definitions.

The regularized ACC problem was solved analytically to find the optimal source weights. Two cases were considered: anechoic and reflective (with surface) system setup. It was shown that with negligible regularization, ACC cancels the sound pressure at the setup sensor in zone B. With the reflective conditions at the system setup stage, the method therefore comprises the direct sound and the reflection cancellers.

The optimal source weights were used to derive the expression for acoustic contrast and array effort produced by the system. Three sound zone generation scenarios were considered: anechoic (with anechoic setup and playback), anechoic-reflective (with anechoic setup and reflective playback), and reflective (with reflective setup and playback).

In the first two cases, the contrast was shown to contain the direct and reflected sound pressure components related to the direct sound canceller, whereas in the last case components related to both the direct and reflection cancellers were observed. It was demonstrated that the effort depends heavily on the sound pressures produced by the cancellers at the setup sensor in zone A. The contrast and effort expressions provided a framework for the analysis that aimed to propose source optimization techniques for the considered 2×2 system.

First, methods of increasing the acoustic contrast through minimization of the sound pressure in zone B were examined. Under the anechoic zone generation scenario, a compact array located far from zone B and with the endfire orientation was found to minimize the direct sound pressure in that zone. The anechoic-reflective scenario was subsequently examined. This resulted in two distinct techniques: Null-Split (pointing
the directivity null at the surface and exploiting the array symmetry) and Far-Align (null sharing between the dark zone and surface). Both techniques were based on optimizing array directivity. Finally, the reflective scenario was investigated. It was shown that both Null-Split and Far-Align apply in this case. Moreover, an additional technique—Near-Align—was proposed. The technique optimizes the operation of the reflective canceller by taking advantage of the spatial match between the sources and their images with respect to the dark zone. The endfire orientation of the array inherent in the Far- and Near-Align techniques allowed better control of the direct sound component, making these two techniques superior to Null-Split.

Investigations into enhancing contrast and reducing effort by increasing the bright zone sound pressure generated by the cancellers were also carried out. It was found that for the largest contrast and lowest effort, the bright zone centre must be located on the other side of the array axis normal than the dark zone, within a certain specified range of angles. This range was found to be the largest when the array is oriented endfire with respect to the dark zone.

In the following, the source optimization techniques and the bright zone positioning requirements for the 2×2 system will be validated using simulations. The analytic techniques will be compared in a broad range of geometrical scenarios in order to provide guidelines for selecting the most suitable technique for a given positioning of the surface with respect to the zones. Numerical search will also be used to test the applicability of the techniques to larger systems with up to three surfaces.
Chapter 5

Optimization of source positions: simulations

In Chapter 4, source optimization problem was studied analytically for a $2 \times 2$ system with a single surface. The Null-Split, Far-Align, and Near-Align techniques were proposed to minimize the reflected sound pressure in the dark zone, thus increasing the contrast. The Far- and Near-Align techniques, due to their inherent endfire orientation at the dark zone, were shown to be optimal for the direct sound control in that zone. Moreover, it was demonstrated that further enhancement of contrast and reduction of array effort can be achieved by appropriate location of the bright zone with respect to the optimized source array.

In this chapter, the Null-Split, Far-Align, and Near-Align techniques are compared for a $2 \times 2$ system with a single surface in various configurations. The produced contrasts and efforts are examined. The analysis leads to formulating guidelines for choosing the most suitable technique for a given position and orientation of the surface with respect to the zones. Furthermore, a numerical search is employed to optimize source positions for the $2 \times 2$ and larger systems with up to three surfaces. The search is constrained to compact arrays with regular source spacing. This provides independent validation of the techniques for the $2 \times 2$ system, and tests their applicability to larger systems and conditions when two or three strongly reflective surfaces are present.
Numerical optimization is also applied to an extended source candidate set to find the best positions when the source array compactness and regular spacing constraints are lifted. This provides information about possible performance improvements with respect to regularly and closely spaced sources.

The specific contributions of this chapter can be summarized as follows:

- Analysis of performance of the proposed source optimization techniques for different surface positions and orientations; formulating guidelines for selecting the most suitable technique for a given surface positioning.

- Numerical optimization results for compact and regular source arrays, for systems with up to fifty sensors, three sources, and three surfaces; analysis of the relationship of the results with the analytically-derived techniques.

- Numerical optimization results for an extended candidate set; analysis of the relationship of the results with the constrained candidate set solutions.

### 5.1 Null-Split, Far-Align, and Near-Align comparison

In this section, simulations are carried out to compare the proposed Null-Split, Far-Align, and Near-Align techniques in a number of geometrical scenarios. Contrast and effort achieved by different configurations is analysed, which results in formulating guidelines for selecting the most appropriate technique for a given positioning of the surface with respect to the zones.

Figure 5.1 shows the considered 2×2 system with a surface in various positions and orientations. The surfaces were defined as lines tangent to a semi-circle with the radius of 4.25 m and centred at (0.75 m, −1.25 m), which were spaced by 5°. This gives a range of geometrical scenarios, $0^\circ \leq \alpha \leq 180^\circ$, from the “East” position ($\alpha = 0^\circ$), through “North” ($\alpha = 90^\circ$), and further to the “West” location ($\alpha = 180^\circ$). The range $180^\circ < \alpha < 360^\circ$ is not considered, as the surface orientations with respect to the zones are similar to those included in the $0^\circ \leq \alpha \leq 180^\circ$ range. In each case, the frequency independent magnitude of the reflection coefficient was set to $\gamma = 1$. The two square
5.1. Null-Split, Far-Align, and Near-Align comparison

Figure 5.1: A 2×2 system with a surface in a number of different locations. The surface is tangent to the semi-circle at each angular position indicated by the arrowhead ▲. Remaining symbols: □ zone, × source array centre: Far- and Near-Align, ⊗ source array centre: Null-Split, — example array axis, ■ example surface; in close-up: ● setup sensor, ○ monitor sensor.

zones, A and B, each contained one setup and thirty-six monitor sensors with 5 cm spacing. For each surface, Null-Split, Far-Align, and Near-Align were evaluated. For all three techniques to be applicable in each case, only the reflective zone generation scenario was considered. An array of two sources with \( d = 5 \text{ cm} \) spacing was moved around a circle of 2 m radius, centred at the zone B’s setup sensor, to form the Far- and Near-Align arrangements. The centres of these arrays are shown as the × symbols in Fig. 5.1; note that for the example surfaces shown, larger symbols are used and the array axes are indicated. The Null-Split was produced by rotating the array located halfway between the zones. In this location, the requirements for increased efficiency of radiation into zone A, defined in Sec. 4.6, were satisfied for a large number of surface positions. The centre of this array is shown as the ⊗ symbol in Fig. 5.1, and the array axes are also indicated for the example surfaces shown.

Setting the reflection coefficient to unity \( (\gamma = 1) \) across all frequency bands provided the worst-case perspective on the effect of reflections. Table 5.1 shows reflections coefficients of typical surfaces and materials (values are taken as \( \gamma = 1 - \alpha_{\text{abs}} \), where \( \alpha_{\text{abs}} \) are the
Chapter 5. Optimization of source positions: simulations

<table>
<thead>
<tr>
<th>Material</th>
<th>Centre frequency of octave band (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>125</td>
</tr>
<tr>
<td>Hard surfaces (brick walls, plaster, hard floors, etc.)</td>
<td>0.98</td>
</tr>
<tr>
<td>Slightly vibrating walls (suspended ceilings, etc.)</td>
<td>0.90</td>
</tr>
<tr>
<td>Strongly vibrating surfaces (wooden panelling over air space, etc.)</td>
<td>0.60</td>
</tr>
<tr>
<td>Carpet, 5 mm thick, on hard floor</td>
<td>0.98</td>
</tr>
<tr>
<td>Plush curtain, flow resistance 450 Ns/m³, deeply folded, in front of solid wall</td>
<td>0.85</td>
</tr>
<tr>
<td>Polyurethane foam, 27 kg/m³, 15 mm thick on solid wall</td>
<td>0.92</td>
</tr>
<tr>
<td>Acoustic plaster, 10 mm thick, sprayed on solid wall</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 5.1: Reflection coefficients $\gamma$ of typical surfaces and materials (based on absorption coefficient values in [Kuttruff, 2009, Chapter 9, pp. 307]).

Absorption coefficients given by Kuttruff [2009, Chapter 9, pp. 307]). It can be noticed that hard, exposed and acoustically untreated surfaces such as plaster walls or hard floors have reflection coefficients approaching unity in all frequency bands—this worst case was represented in this simulation study. Other materials are characterized by reduced reflectivity. For instance, lower reflectivity at low frequencies is characteristic of vibrating structures, such as suspended ceilings. Significant reduction of reflectivity at higher frequencies can be achieved by porous materials such as carpets and curtains, as well as by polyurethane foam and acoustic plaster.

Source weights were determined using the regularized ACC method described in section
5.1. Null-Split, Far-Align, and Near-Align comparison

Figure 5.2: Frequency averaged contrast produced by the 2×2 system in the Null-Split, Far-Align, and Near-Align arrangements with a surface in different positions, as shown in Fig. 5.1. The thick portions of the curves correspond to the arrays that fulfil the requirements for increased efficiency of radiation into zone A, specified in Sec. 4.6. The arrays producing the marked contrast values are indicated in Fig. 5.3.

"Acoustic contrast control" included in Sec. 2.3.1. A frequency independent regularization with $\beta = 10^{-6}$ was applied; this value was found to be the smallest that avoided singularity in the numerical solutions. This type of regularization parallels the approach from the analysis in Chapter 4, where $\beta$ was assumed to be negligible or infinitesimal. The squared pressures in each zone were calculated for forty-four frequency bins with one-twelfth octave band spacing in the range 250–3175 Hz. To evaluate the system, the frequency-averaged squared pressure was calculated at each monitor sensor. The motivation behind the logarithmic spacing was to avoid bias from regular variations of pressure over frequency when calculating the averages. The upper frequency limit is selected to be two bands below the spatial aliasing limit of the array (3430 Hz) to reduce the influence of the aliasing-related effects, such as beam splitting, on the results. When calculating the averages, values at each frequency were linearly weighted to compensate for the logarithmic spacing. The contrast was then obtained using Eq. (2.8).

Figure 5.2 shows the contrast plotted against surface position. The marked values, contained in and bounding the thick portions of the contrast curves, were produced by the arrays enclosed in the white regions in Fig. 5.3 (note that each array’s location
Figure 5.3: Ranges of θ_A that fulfil the requirements for increased efficiency of radiation into zone A, specified in Sec. 4.6 (white regions), and arrays corresponding to the marked contrast and effort values in Figs. 5.2 and 5.4 respectively: ♦ (red) Null-Split, ▼ (green) Far-Align, ● (blue) Near-Align.

can be determined by the angles interpreted in either direction from the array axis). Therefore, the thick portions of the curves correspond to the arrays that fulfil the requirements for increased efficiency of radiation into zone A, specified in Sec. 4.6, and thus increase the contribution of sound pressure in zone A to contrast. These arrays achieve high contrasts (over 28 dB). Each technique has a range where it performs better than the others: Far-Align in 0° ≤ α ≤ 80°, Null-Split in 85° ≤ α ≤ 120°, and Near-Align in 125° ≤ α ≤ 180°. Considering the surface positions for which the techniques lead, the following design guidelines can be formulated:

- if the distance of both zones from the surface is the same or similar, the sources should be positioned according to the Null-Split rule;
- if the surface is closer to zone B than zone A, the Far-Align arrangement should be chosen; and
- if zone A is closer to the surface than zone B, the Near-Align technique applies.

For surfaces in ambiguous positions, a closer examination of geometrical options should identify the best technique. In such cases, the arrays that fulfil the positioning requirements for zone A should be the first choice, and the Far- and Near-Align should be
5.1. Null-Split, Far-Align, and Near-Align comparison

preferred over Null-Split. According to Fig. 5.2, combining these selection criteria with the general guidelines allows choosing the best performing arrays for all the examined surface positions, except for $65^\circ \leq \alpha \leq 80^\circ$ where zone A is outside the required range for Far-Align, yet this solution outperforms the Null-Split. This can be attributed to less effective attenuation of sound in zone B by non-endfire Null-Split arrays, as discussed in Sec. 4.5. However, the contrast differences between the two techniques in this range are relatively small (less than 3 dB).

Figure 5.4 shows the equivalent effort plot. As in the case of contrast, the values marked by the thick portions of the curves were produced by the arrays fulfilling the positioning requirements for zone A (i.e. those enclosed in the white regions in Fig. 5.3). These values generally represent the lowest effort for each technique. It is therefore clear that increasing the sound pressure radiated by the cancelling array into zone A significantly lowers the effort. This is certainly the case for Null-Split and Far-Align, except for the $180^\circ$ case where the Null-Split’s effort drops in spite of the fact that zone A is not appropriately positioned. This can be attributed to the array’s symmetry with respect to both setup sensors. Solving the ACC problem in this case gives the BC’s source weights, as shown in Appendix B. Sound cancelling in both zones is therefore avoided, and instead the effort is reduced by steering the main energy beam at zone A.

The Near-Align technique exhibits more effort variability across surface positions than the other techniques. Over 10 dB is reached in the thick-curve region, and lower values can be observed in the thin-curve portion. This can be attributed to the fact that the interference between the direct sound and the reflection at the setup sensor in zone A is more pronounced compared to the other techniques. This makes the estimation of sound pressure based on the direct sound component less reliable. In consequence, satisfying the previously defined positioning requirements for the bright zone may not always give the best result. Nevertheless, the effort in the thick-curve region remains low. Elsewhere, a surge in effort at the $65^\circ$ position can be observed. The peak is due to destructive interference between the direct sound and the reflection, producing a deep null around the setup sensor in zone A in the 2.52 kHz frequency band, whose compensation increases the effort.
Chapter 5. Optimization of source positions: simulations

Figure 5.4: Frequency averaged effort produced by the 2×2 system in the Null-Split, Far-Align, and Near-Align arrangements with a surface in different positions, as shown in Fig. 5.1. The thick portions of the curves correspond to the arrays that fulfil the requirements for increased efficiency of radiation into zone A, specified in Sec. 4.6. The arrays producing the marked effort values are indicated in Fig. 5.3.

5.2 Numerical optimization of source positions

In this section, source positions are optimized by numerical search. Such optimization serves two purposes. First, it provides independent validation of the techniques proposed in Chapter 4. This can be achieved by searching over a large source candidate set, which could reveal alternative techniques potentially overlooked in the analysis. Second, it is used to find optimal source positions for systems with additional sensors, sources and surfaces—a numerical search allows straightforward extension of the optimization techniques to more complex systems.

Two types of optimizations are considered. First, the candidate source locations are such that only compact arrays with regular spacing can be formed. This set allows validating the analytic techniques, which were formulated assuming compact source arrangements. A basic extension to the three-source case is also provided. Second, the candidate set is extended to allow formation of non-compact and irregular arrays, which could provide a useful alternative to compact arrays.

Two zone generation scenarios are considered: anechoic and reflective. The anechoic scenario was included to provide a reference for the analysis of the reflective case. As
5.2. Numerical optimization of source positions

Figure 5.5: Configuration used in the numerical search for optimal source positions. Symbols: □ zone, ☐ source subset, ■ surface; in close-up: ★ fixed source, ■ candidate source.

The anechoic-reflective scenario limits the spectrum of possible solutions to Null-Split and Far-Align, it is not examined. The examined systems are 2×2, 2×50, and 3×50.

5.2.1 Compact arrays with regular spacing

Figure 5.5 shows the geometries used in the search for best-performing compact arrays. There were three surfaces (under reflective conditions), each with $\gamma = 1$, considered either individually or in combination. For the 2×2 system, the same sensor layouts as in Fig. 5.1 were used. For the 2×50 and 3×50 systems, twenty-five setup sensors were utilized in each zone, arranged on square 5 cm grids that were centred at the setup sensor positions used for the 2×2 system. There were thirty-six subsets of candidate sources located around zone B on a 2 m radius circle (10° interval). Each subset contained a fixed source and candidate sources located on the inner and outer arcs around that source. There were thirty-seven candidate sources on each arc (5° separation). In the 2×2 and 2×50 cases only the inner arc was considered, whereas for the 3×50 system the candidates from both arcs were used.

The optimization procedure was based on the principle of a beam search, i.e. developing a small number of solutions in parallel to minimize the search effort [Ow and
Morton, 1988]. First, the subsets were tested in parallel to find the best array orientation within each subset. For the 2×2 and 2×50 systems, this meant choosing the best out of thirty-seven source pairs formed by the fixed source and each of the candidate sources located on the inner arc. In the 3×50 case, the optimization algorithm chose the best of thirty-seven triples formed by the fixed source and aligned pairs of candidates located on the two arcs. The overall best performing configuration was then selected from the thirty-six preselected pairs or triplets. The objective function was the frequency-averaged acoustic contrast, calculated as in Sec. 5.1. The source weights were determined similarly as in Sec. 5.1 and scaled to produce 94 dB SPL in zone A (measured as the average SPL at the setup sensors).

Anechoic conditions

Figure 5.6a shows the optimization results for a 2×2 system under the anechoic zone generation scenario (solid, black line), overlaid on the SPL map generated by the optimized array at 1 kHz. The sources have the endfire orientation with respect to the centre of zone B, and an extended null covers the dark zone. This is consistent with the analytic result from Sec. 4.5.1. It can also be noticed that the cancellation is intensified locally around the setup point. The array does not steer endfire at zone A as it might
be expected. Such an arrangement is shown in the same figure (white, dashed line). This would result in the contrast drop towards the spatial aliasing frequency limit, 3430 Hz in this case, caused by the splitting of the main radiation beam. The array is therefore re-oriented to alleviate this effect. The contrasts produced by the optimal and endfire arrays is plotted in Fig. 5.6b. It can be noticed that the optimal array increases the contrast by approximately 6 dB at the highest frequency considered in the optimization (3175 Hz), which is the reason why it was chosen by the optimization algorithm in favour of the alternative arrangement. As expected, further drop of contrast is observed as the frequency increases towards 3430 Hz (which is limits the x-axis in the figure).

The position and orientation of the optimal array in the 2×50 and 3×50 cases was identical as for the 2×2 system. These arrays are shown in Appendix C.

2×2 system with a single surface

Figure 5.7 shows the optimization results for a 2×2 system with a surface in three different positions, together with the SPL maps at 1 kHz. The Null-Split, Far-Align, and Near-Align arrangements were chosen as optimal for the North, East, and West surfaces respectively, which demonstrates the validity of these techniques. The Null-Split produces a narrower null in zone B than the Far- and Near-Align arrangements, which was indicated by the analytic results in Sec. 4.5. As expected, the technique selection follows the general guidelines from Sec. 5.1 (the guidelines were formulated using similar configurations). It was verified that all three arrays fulfilled the requirements for increased efficiency of radiation into zone A, which confirms that it is an important factor in the selection.

2×50 system with a single surface

For the East and West surfaces, the source positions selected for the 2×2 and 2×50 systems were identical (see Appendix C). This demonstrates the applicability of the analytic Far- and Near-Align solutions to systems with extended setup sensor arrays.
A different solution was chosen in the North case, where the array approximated Far-Align, as shown in Fig. 5.8a. The sources were shifted from the regular Far-Align position towards zone A and rotated to remain endfire at zone B. Neither the Far-Align arrangement nor the adjusted array met the positioning requirements for zone A. However, the adjustment alleviated the effect of unwanted sound attenuation in that zone, increasing contrast at low frequencies. The preference of this solution over the Null-Split arrangement from Fig. 5.7a can be attributed to null broadening, which improved contrast in the middle and high frequency regions, as demonstrated in Fig. 5.8b.

It is also worth investigating why the approximated Far-Align arrangement was not selected as optimal in the North case for the $2 \times 2$ system. The contrast curve produced by the array in this position exhibits regular strong variations across frequency, as shown in Fig. 5.8c. This characteristic does not occur for the Null-Split’s contrast, which is also plotted in Fig. 5.8c. Furthermore, the Null-Split performs better at low frequencies, which overall makes it a more suitable solution in the $2 \times 2$ case.

### 3×50 system with a single surface

For the East and West surfaces, the positions and orientations of the arrays chosen for the $3 \times 50$ system were the same as in the $2 \times 2$ and $2 \times 50$ cases (see Appendix C), and for the North surface the Near-Align arrangement was selected. These results indicate the
5.2. Numerical optimization of source positions

Figure 5.8: Results for systems with a single surface in the North location. (a) Optimized source array for a 2×50 system, and the sound pressure level map at 1 kHz; symbols: □ zone, √ array centre, — array axis, ■ surface. (b) Contrast produced by a 2×50 system, using the optimal approximated Far-Align arrangement from Fig. 5.8a (thick, red) and the Null-Split arrangement from Fig. 5.7a (thin, green). (c) Contrast produced by a 2×2 system, using the optimal Null-Split arrangement from Fig. 5.7a (thick, red) and the approximated Far-Align arrangement from Fig. 5.8a (thin, green).

The suitability of the Far- and Near-Align techniques for larger source arrays. The North case is shown in Fig. 5.9a. The array does not satisfy the positioning requirements for zone A, which means that the direct sound in that zone is excessively attenuated. This attenuation is compensated by strong reflections produced by the array located close to the surface. Interference between the direct sound and reflections results in unwanted pressure nulls occurring in zone A periodically at certain frequencies. These nulls cause the periodic dips in contrast, as shown in Fig. 5.9b. However, the overall contrast is high due to effective cancellation—the Near-Align arrangement facilitates destructive interference between the direct sound and reflection, producing a localized minimum in zone B. As a result, the contrast is higher at most frequencies compared to that produced by the Far-Align arrangement, similar to the one that was optimal in the 2×50 case (Fig. 5.8a).

It is also interesting to investigate why a similar Near-Align configuration was not optimal in the 2×50 case. The contrast for such a configuration is compared with that produced by the optimal approximated Far-Align arrangement in Fig. 5.9c. It can be noticed that the Near-Align solution produces large periodic dips, similarly to
Chapter 5. Optimization of source positions: simulations

Figure 5.9: Results for systems with a single surface in the North location. (a) Optimized source array for a 3×50 system, and the sound pressure level map at 1 kHz; symbols: □ zone, × array centre, — array axis, ■ surface. (b) Contrast produced by a 3×50 system, using the optimal Near-Align arrangement from Fig. 5.9a (thick, red) and the Far-Align arrangement similar to that in Fig. 5.8a (thin, green). (c) Contrast produced by a 2×50 system, using the optimal approximated Far-Align arrangement from Fig. 5.8a (thick, red) and the Near-Align arrangement similar to that in Fig. 5.9a (thin, green).

The 3×50 case. However, local cancellation in the dark zone is not effective enough to increase the overall contrast, making it generally lower than that produced by the optimal arrangement.

Systems with two and three surfaces

Figure 5.10 shows the optimization results for the 2×2 system with two or three surfaces generating first order reflections. Figure 5.10a shows the case with two perpendicular surfaces. This layout combines the Far-Align (non-endfire) and the Null-Split solutions for the East and North surfaces respectively. In Fig. 5.10b, two surfaces are positioned in parallel. The chosen array forms the Far- and Near-Align arrangements for the East and West surfaces respectively. These results demonstrate that combining the Null-Split, Far-Align and Near-Align techniques is a valid optimization approach when two strongly reflecting surfaces coexist. Figure 5.10c shows the case with three surfaces. Since the baseline solutions cannot be combined as in the two-surface cases, a hybrid solution is created to facilitate controlling reflections from all surfaces. The array is rotated from the position in Fig. 5.10b towards the orientation in Fig. 5.10a, but does
5.2. Numerical optimization of source positions

5.2.1 Compact and regular arrays

In the previous section, the candidate source set was restricted to forming compact and regular arrays. In this section, the candidate source set is extended to allow forming non-compact and irregular arrays. As in the previous case, the 2×2, 2×50, and 3×50 systems are examined. The investigations are limited to the reflective scenario with a single surface in three different positions (North, East, West) to seek alternatives to the baseline techniques.

Figure 5.11 shows the geometries used in the search. The surface layout and properties, as well as the sensor configurations, were the same as in Sec. 5.2.1. Again, the search was based on the principle of the beam search. There were thirty-six fixed (seed) sources located around zone B (10° intervals, 2 m radius) and 10076 candidate sources positioned on a toroidal grid extending from the seed sources (2.5 cm spacing, outer radius of 2.45 m). Each search instance considered a different seed source and all of the

Figure 5.10: Optimized source arrays for a 2×2 system with up to three surfaces generating first order reflections, and the sound pressure level maps at 1 kHz: (a) East and North, (b) East and West, and (c) East, North and West surfaces. Symbols: □ zone, × array centre, — array axis, ■ surface.

not reach that position, seeking the geometrical compromise between the two cases.

For 2×50 and 3×50 systems with the same surface combinations, the optimized source arrays had positions and orientations that were similar to the 2×2 case (see Appendix C).

5.2.2 Non-compact and irregular arrays

In this section, the candidate source set is extended to allow forming non-compact and irregular arrays. As in the previous case, the 2×2, 2×50, and 3×50 systems are examined. The investigations are limited to the reflective scenario with a single surface in three different positions (North, East, West) to seek alternatives to the baseline techniques.

Figure 5.11 shows the geometries used in the search. The surface layout and properties, as well as the sensor configurations, were the same as in Sec. 5.2.1. Again, the search was based on the principle of the beam search. There were thirty-six fixed (seed) sources located around zone B (10° intervals, 2 m radius) and 10076 candidate sources positioned on a toroidal grid extending from the seed sources (2.5 cm spacing, outer radius of 2.45 m). Each search instance considered a different seed source and all of the
Chapter 5. Optimization of source positions: simulations

Figure 5.11: Extended configuration used in the numerical search for optimal source positions. Symbols: □ zone, ■ surface, ★ fixed source, ○ toroidal grid of candidate sources (2.5 cm spacing).

Due to a large candidate set, a more advanced search algorithm had to be chosen to handle the source selection. Numerical optimization of source positions is related to the problem of feature selection in pattern recognition. In the considered problem, the algorithm was to search for subsets of features (source positions) out of the total number of available ones, such that the criterion function (acoustic contrast) was maximized. The search algorithms for feature selection can be divided into two main groups: optimal and suboptimal [Devijver and Kittler, 1982, Chapter 5, Secs. 5.4–5.7]. Optimal algorithms, such as exhaustive search and the branch and bound (BAB) allow finding the global optimum in the search space by examining all possible feature combinations. However, these methods are not computationally efficient, which makes them unsuitable for the large scale problem considered here.

Suboptimal methods can be more efficient, but this is achieved at the expense of an increased risk of finding a local optimum. Two notable types of suboptimal methods are the sequential search and genetic algorithms. Sequential forward selection and its backward counterpart are the basic sequential algorithms. They are fast and straight-
5.2. Numerical optimization of source positions

Input:

\[ Y = \{ y_j | j = 1, ..., D \} \]

Output:

\[ X_k = \{ x_j | j = 1, ..., k, x_j \in Y \}, k = 1, ..., D \]

Initialization:

\[ k := 0; X_0 := \emptyset; \text{ go to Step 1} \]

Termination:

Stop when \( k \) equals the number of features required

**Step 1 (Inclusion)**

repeat \( l \) times

\[ x^+ = \arg \max_{x \in Y - X_k} J(X_k + x) \quad \text{the most significant feature with respect to } X_k \]

\[ X_{k+1} := X_k + x^+; k := k + 1 \]

go to Step 2

**Step 2 (Exclusion)**

repeat \( r \) times

\[ x^- = \arg \max_{x \in X_k} J(X_k - x) \quad \text{the least significant feature in } X_k \]

\[ X_{k-1} := X_k - x^-; k := k - 1 \]

go to Step 1

Table 5.2: Plus-\( l \) Take Away-\( r \) algorithm (bottom up search procedure). Reproduced from [Ferri et al., 1994].

Forward in implementation, but suffer from the nesting of the chain of subsets which results in low performance. This problem is partially overcome by the Plus-\( l \) Take Away-\( r \) (PTA(\( l \),r)) algorithm, which allows low level backtracking in the selection process by combining sequential forward and backward selection [Devijver and Kittler, 1982, Chapter 5, pp. 220]. Genetic Algorithm (GA) is known to provide a good trade-off between accuracy and computation time [Kudo and Sklansky, 2000]; however, the result depends on the appropriate choice of parameters which must often be done by trial and error. Therefore, the PTA(\( l \),r) was chosen to carry out the search, due to its speed, good accuracy and relative simplicity.

The PTA(\( l \),r) method is characterized in Table 5.2. The input to the algorithm is the superset of features \( Y = \{ y_j | j = 1, ..., D \} \) where \( D \) is the total number of features. The output is the subset \( X_k = \{ x_j | j = 1, ..., k, x_j \in Y \} \) where \( k = 0, 1, ..., D \). For the bottom up search procedure used here, the initial subset is empty: \( k := 0; X_0 := \emptyset \). The
algorithm begins with adding a feature to the set $X_0$, such that the criterion function $J(X_k)$ is maximized. The process of addition is then repeated $l$ times, and in each iteration the most significant feature with respect to the existing set $X_k$ is added. Subsequently, the least significant feature is excluded from the resulting set iteratively $r$ times. The cycle of inclusion and exclusion is repeated until $k$ reaches the number of features required. In the considered problem, the algorithm stopped when $k = 2 \lor k = 3$ (the target number of sources). The iteration counters were set to $l = 2$ and $r = 1$, which required two and three cycles of the algorithm to obtain the result in the two and three source cases respectively. The criterion function used was frequency-averaged acoustic contrast, calculated as in Sec. 5.1. The frequency range was extended to 250–3564 Hz, giving a total of forty-six frequency bins with one-twelfth octave band spacing. The source weights were determined as in Sec. 5.2.1.

**2×2, 2×50, and 3×50 systems with a single surface**

Figure 5.12 shows the source optimization results for the 2×50, and 3×50 systems with a single surface in the North position, overlaid on the SPL maps at 1 kHz (note that the source positions rather than array centres are indicated). In the 2×2 case, the solution was the same as for 2×50, so it is not shown. The arrangement selected as optimal for both systems is Null-Split. The source positions are similar as in the 2×2 case from Sec. 5.2.1. However, the centre of the array is here shifted away from the surface and rotated accordingly to achieve the Null-Split configuration. In this way, the spacing $d = 3.5 \text{ cm}$ was achieved, which is the minimum for non-orthogonal arrays formed on the considered grid. Small spacing reduced the contrast drop at higher frequencies. In the 2×50, this contrast improvement is the main factor contributing to the preference of Null-Split over approximation to Far-Align, which was previously selected as optimal in Sec. 5.2.1 (the spacing between the sources was then set to $d = 5 \text{ cm}$). In the 3×50 case, the optimal arrangement is Near-Align, similarly as in Sec. 5.2.1, but with wider and irregular spacing: $d_1 = 10 \text{ cm}$ and $d_2 = 5 \text{ cm}$. It is interesting to investigate why this arrangement was preferred over a compact arrangement with minimum spacing ($d = 2.5 \text{ cm}$) and the Null-Split configuration selected for 2×50. The contrasts for the three cases (Near-Align with irregular and wide spacing, Near-Align with compact
5.2. Numerical optimization of source positions

Figure 5.12: Optimized source arrays with a single surface in the North location, and the sound pressure level maps at 1 kHz: (a) 2×50, \( d = 3.5 \) cm, and (b) 3×50, \( d_1 = 10 \) cm, \( d_2 = 5 \) cm. Symbols: \( \square \) zone, \( \times \) source position, — array axis, \( \square \) surface.

(a) 2×50: Null-Split.
Contrast: 35 dB.

(b) 3×50: Near-Align.
Contrast: 77 dB.

The preference of the Null-Split over the compact and widely-spaced Near-Align configurations could be also examined for the 2×50 case. For this system, the contrasts achieved by the Null-Split, and the two Near-Align configurations (\( d = 2.5 \) cm and \( d = 10 \) cm in the compact and widely-spaced cases respectively) are plotted in Fig. 5.13b. It can be noticed that all scores are within a similar range, but the variability of contrast for Null-Split is largely reduced compared to the Near-Align arrangements. Thus, Null-Split achieves the highest mean contrast (34.9 dB, compared to 25.7 dB and 29.4 dB produced by the other arrays).

Figure 5.14 shows optimal source configurations and the corresponding SPL maps at 1 kHz for the 2×50 and 3×50 systems with the surface in the West position (the result for 2×2 and 2×50 were identical, so the former is not shown). Both arrays were formed according to the Near-Align technique. The results are therefore similar to the corresponding case described in Sec. 5.2.1, but the source spacing is larger: \( d = 22.5 \) cm, and \( d_1 = 22.5 \) cm, \( d_2 = 15 \) cm in the 2×50 and 3×50 cases respectively. Interestingly, the smallest possible spacing (2.5 cm) was not optimal. Fig. 5.15a shows the acoustic
Figure 5.13: Contrast produced for the surface in the North location. (a) 3×50 system: Null-Split arrangement similar to that in Fig. 5.12a (thick, red), Near-Align arrangement with compact and regular spacing ($d_1 = d_2 = 2.5$ cm), similar to that shown in Fig. 5.9a (thin, green), and Near-Align arrangement with wide and irregular spacing ($d_1 = 10$ cm, $d_2 = 5$ cm) as in Fig. 5.12b (dashed, blue). (b) 2×50 system: Null-Split arrangement as in Fig. 5.12a (thick, red), Near-Align arrangement with compact spacing ($d = 2.5$ cm), similar to that shown in Fig. 5.9a (thin, green), and Near-Align arrangement with wide spacing ($d = 10$ cm) similar to that shown in Fig. 5.12b (dashed, blue).

contrast produced by such a compact array and the array with optimal spacing (wide) for the 2×50 case. It can be noticed that the wider spacing results in large variability of contrast across frequency, with multiple peaks and dips (a very similar result was observed in the 2×2 case). Due to the peaks, the mean contrast is higher than for the compact array, but the difference is less than 1 dB. The contrast gain from larger spacing is therefore small. Furthermore, the irregularity of contrast with frequency may be undesirable. A more complex cost function could be considered to improve this optimization result—for instance, a standard deviation penalty could be introduced in addition to the mean contrast criterion. A different performance characteristic is observed for the 3×50 system—wide and irregular spacing increases the overall contrast without introducing much variability compared to the compact arrangement, as shown in Fig. 5.15b. As a result, a mean contrast gain of over 3 dB is achieved.

For the surface in the East position, regular and compact arrays formed the Far-Align arrangement for each system, giving similar results as in Sec. 5.2.1 (see for instance Fig. 5.7b), but with a smaller source spacing ($d = 2.5$ cm).
5.3. Summary

In this chapter, simulation results were presented. First, the Null-Split, Far-Align and Near-Align techniques were applied to a 2×2 system with a single surface in various positions and compared in terms of the produced contrast and effort. This allowed for the formulation of guidelines to choose the most suitable technique for a given surface positioning with respect to the zones: (i) Null-Split is a preferred solution when the bright and dark zones are at the same or similar distance from the surface; (ii) Far-Align should be used when the dark zone is closer to the surface than the bright zone; (iii) Near-Align should be used when the bright zone is closer to the surface than the dark zone.

The results of numerical search for optimal source positions were also presented. The search was first carried out on the source candidate set forming compact arrays with regular spacing, using the frequency-averaged acoustic contrast as the objective function. This confirmed that the endfire orientation of the array at the dark zone is optimal under anechoic zone generation conditions. The analytic Null-Split, Far-Align, and Near-Align techniques were also validated by the results of that search for a 2×2 system. Moreover, the requirements for increased efficiency of array radiation into the bright zone, determined in Sec. 4.6, were confirmed to be an important factor in the selection of the most suitable technique. The Far- and Near-Align techniques were

Figure 5.14: Optimized source arrays with a single surface in the West location, and the sound pressure level maps at 1 kHz: (a) 2×50, \( d = 22.5 \) cm, and (b) 3×50, \( d_1 = 22.5 \) cm, \( d_2 = 15 \) cm. Symbols: □ zone, × source position, — array axis, ■ surface.
found to be applicable to systems with additional sensors and sources. Furthermore, a combination of the Null-Split, Far-Align, and Near-Align techniques was found to be a valid approach to optimizing source position with two strongly reflecting surfaces. With three surfaces, the optimal source arrangement was a hybrid that sought geometrical compromise between the analytic techniques applied to the different surfaces.

A source candidate set that allows for non-compact, irregular arrays, was also used in the numerical search for systems with a single surface in three different positions. The PTA(l,r) search algorithm was employed. In all cases, the chosen arrays formed the Null-Split, Far-Align, and Near-Align arrays, giving similar results to the search over the candidate set that allowed only compact arrays with regular spacing. The two-source cases with surfaces in the North (equidistant from the zones) and East (much closer to the dark zone than the bright zone) positions, compact arrays were formed. With the surface located in the West (much closer to the bright zone than the dark zone) position, large source spacing was chosen as optimal. However, the optimal widely-spaced array was found to give marginally better contrast (less than 1 dB) than a more compact arrangement. Furthermore, the latter produced a smoother contrast curve, which is a desirable characteristic. A more advanced objective function could improve this optimization result—for instance additional standard deviation penalty
could lead to the preference of the compact arrangement. A different performance characteristic was observed for three sources with the surfaces in the North and West locations—the arrays with larger, irregular spacing, which were chosen as optimal, gave a notable improvement in contrast (3 dB) over the compact and regular configurations, without excessive contrast variation over frequency.

The Null-Split, Far-Align, and Near-Align techniques are therefore valid source placement techniques for small sound zone systems with up to two strongly-reflecting surfaces, and provide foundation for the optimization of systems with additional surfaces. They may therefore be considered as an alternative to passive acoustic treatment. Furthermore, the techniques may influence the specification of candidate source locations when optimizing larger systems by indicating alternatives to conventional linear or circular arrangements.

Having validated the analytic techniques, the systems that employ the Null-Split, Far-Align, and Near-Align techniques or their numerical approximations can now be evaluated for acoustic contrast and array effort. These results will be compared with the performance of the systems based on alternative approaches to sound zone generation in reflective rooms, thus quantifying the performance gains from using the techniques.
Chapter 5. Optimization of source positions: simulations
Chapter 6

Evaluation of reflection control strategies

In Chapter 4, source position optimization techniques to increase contrast in a room with a strongly reflecting surface were proposed. Their validity for the 2×2 and larger systems with up to two surfaces was demonstrated with the numerical search results presented in Chapter 5.

As discussed in Sections 1.2.1 and 1.2.3, there are limited examples of evaluation of small sound zone systems under reflective conditions. As a result, certain potentially effective approaches to sound zone generation in the reflective sound field have not been evaluated. One such approach could be to employ the SPM method, which minimizes the acoustic power radiated from the array—the method should therefore be robust to the influence of reflections due to a relatively large DRR achieved in the zones. Another approach could be to use the ACC method with the source positions optimized for the direct sound only to generate the directivity pattern that is best suited for a given dark zone position, as discussed in Sec. 4.5.1 and similarly to [Elliott and Jones, 2006] or [Jones and Elliott, 2008]. In Chapter 4, the techniques were proposed to optimize source positions for ACC, taking the first order reflections into account. Application of the techniques can therefore be regarded as yet another approach to sound zone generation in a reflective room environment. Evaluation and ranking of the above approaches could serve as a useful guideline for system designers. So in this chapter, the
ACC-based systems with sources optimized under different conditions are evaluated and compared with the SPM-controlled systems. Non-optimized systems are also examined. In this way, the proposed source optimization techniques can be evaluated against some baseline sound zone reproduction strategies.

As discussed in the introduction to Chapter 4, the influence of first order reflections is of primary concern in sound reproduction in reflective rooms. Therefore, the evaluation focuses on the single- and two-surface scenarios with first order reflections.

The contribution of this chapter can be summarized as follows:

- Comparison of two different approaches to source optimization for the ACC method with the SPM-controlled systems under single- and two-surface scenarios.

### 6.1 Systems and methods under evaluation

In this section, the systems and methods under evaluation are introduced.

The examined systems were 2x2, 2x50, and 3x50, as in in Chapter 5. One or two fully reflective surfaces ($\gamma = 1$) generating first order reflections were considered. The Null-Split, Far-Align and Near-Align techniques were represented by the numerically optimized arrays from Sec. 5.2.1 (examples were shown in Figures 5.7, 5.8a, 5.9a, and 5.10). The assignment of the techniques to particular systems and surfaces is summarized in Table 6.1. The ACC-based systems optimized for the direct sound only were represented by arrays from section “Anechoic conditions” included in Sec. 5.2.1 (a 2x2 array was shown in Fig. 5.6a). To ensure meaningful comparison with the ACC-based systems, the SPM-controlled systems were also based on source positions that have been numerically optimized for maximum contrast. The optimization was carried out based on the candidate set previously used for the optimization for ACC, shown in Fig. 5.5. The search procedure followed the same steps as described in Sec. 5.2.1, and the source weights were calculated using the method described in Sec. 2.3.1. Therefore, all the evaluated arrays were obtained from the numerical search over compact and
6.1. Systems and methods under evaluation

<table>
<thead>
<tr>
<th>Surface</th>
<th>System 2×2</th>
<th>System 2×50</th>
<th>System 3×50</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>Null-Split</td>
<td>approx. Far-Align</td>
<td>Near-Align</td>
</tr>
<tr>
<td>East</td>
<td>Far-Align</td>
<td>Far-Align</td>
<td>Far-Align</td>
</tr>
<tr>
<td>West</td>
<td>Near-Align</td>
<td>Near-Align</td>
<td>Near-Align</td>
</tr>
<tr>
<td>East, North</td>
<td>Far-Align, Null-Split</td>
<td>Far-Align, Null-Split</td>
<td>Far-Align, Null-Split</td>
</tr>
<tr>
<td>East, West</td>
<td>Far-Align, Near-Align</td>
<td>Far-Align, Near-Align</td>
<td>Far-Align, Near-Align</td>
</tr>
</tbody>
</table>

Table 6.1: Assignment of the proposed source optimization techniques to particular systems and surfaces, according to the numerical optimization results from Sec. 5.2.1.

regular candidate arrays as in Sec. 5.2.1, using the frequency-averaged contrast as the objective function. All the optimized ACC and SPM arrays evaluated in this section are shown in Appendix C.

All systems and methods were evaluated by calculating the average acoustic contrast in the frequency range of 250–3175 Hz, using the procedure explained in Sec. 5.1. The array effort was calculated similarly using Eq. 2.11, where the reference monopole was assumed to be located at the geometrical centre of each array. The frequency responses of contrast and effort were also examined.

For each system, three stages can be distinguished in the sound zone generation process: source position optimization, source weights calculation, and sound field evaluation. At each stage, the acoustic conditions can be either anechoic (abbreviated to A) or reflective (R). First, the optimal arrays from Sec. 5.2.1 were evaluated. These arrays are ACC-based, with frequency-independent regularization (β = 10⁻⁶) and anechoic or reflective conditions at all stages, hence they are referred to as ACC A-A-A or ACC R-R-R respectively. Furthermore, the optimal arrangements obtained considering
the direct sound only were used with different conditions at the source weight and sound field evaluation stages, hence ACC A-A-R and ACC A-R-R. The performance of the arrays with non-optimized source positions was represented by the median of contrasts or efforts achieved by all the configurations considered in the optimizations under the reflective conditions (ACC R\textsubscript{M}-R-R). Finally, the SPM method with the reflective conditions at all stages was evaluated (SPM R-R-R).

It is now necessary to discuss the motivation behind each system and method included in this evaluation. The ACC A-A-A arrays have source positions and weights optimized for maximum contrast under anechoic conditions. With matching ideal conditions at the evaluation stage, they define the upper contrast performance limit. Comparing the ACC R\textsubscript{M}-R-R, ACC A-A-R, ACC A-R-R, ACC R-R-R, and SPM R-R-R results with this limit will rank a number of different approaches to sound zone generation under reflective conditions. The ACC R\textsubscript{M}-R-R results will indicate the probable effect of attenuating reflections using an arbitrarily chosen configuration (there is equal chance that contrast will be higher or lower than the median value). Evaluating the ACC A-R-R and ACC A-A-R systems will show the effect of using sources optimized for the direct sound only under reflective conditions, either with (A-R-R) or without (A-A-R) the reflection canceller. By employing optimized sources, these systems are more advanced with respect to ACC R\textsubscript{M}-R-R. Evaluating the ACC R-R-R systems will demonstrate the impact of the proposed Null-Split, Far-Align, and Near-Align techniques, i.e. combining the reflection canceller and sources optimized for the reflective conditions. Finally, the SPM R-R-R results will show the effectiveness of using the power minimization approach to reduce the impact of individual reflections on contrast. As discussed in Sec. 2.3.1, the SPM method adjusts the source weights, taking reflections into account, to minimize the acoustic power radiated by the array while aiming to maintain a high SPL in zone A [Jones and Elliott, 2008; Elliott et al., 2010]. Minimizing the power should reduce the strength of reflections in the dark zone, thus improving contrast. Further improvement should be obtained from optimizing the source positions for maximum contrast.
6.2 Contrast and effort performance

Table 6.2 shows the frequency-averaged contrast for the 2×2, 2×50, and 3×50 systems. In all cases, the ACC R-R-R systems achieved the highest contrasts (results in boldface), exhibiting the smallest degradation of performance with respect to the reference ACC A-A-A case. While an average 18.6 dB contrast loss was observed for ACC R-R-R configurations in the single surface scenario, in the ACC R\textsubscript{M}-R-R, ACC A-A-R, ACC A-R-R, and SPM R-R-R cases the contrast was degraded on average by 58.9 dB, 42.7 dB, 33.9 dB, and 52.9 dB respectively. This demonstrates the benefits of source optimization using the Null-Split, Far-Align, and Near-Align techniques and their approximations. Moreover, these results indicate that the ACC method implemented on a geometrically optimized array, even for the direct sound only, is a more suitable sound zone reproduction strategy than SPM when strong individual reflections occur. Comparison of the ACC A-R-R and ACC R-R-R contrast scores for the systems with two surfaces shows an average 9.6 dB gain when using configurations that combine the Null-Split, Far-Align, and Near-Align solutions instead of sources optimized considering the direct sound only.

It is also worth discussing the frequency characteristics of contrast produced by different methods. Figure 6.1 shows contrast over frequency for the ACC-based systems when the surface is in the North location. In the 2×2 case (Fig. 6.1a), the Null-Split configuration (R-R-R) outperforms the arrays optimized for direct sound only (A-A-R and A-R-R) consistently over frequency. In the the 2×50 case (Fig. 6.1b), the approximated Far-Align (R-R-R) cannot maintain contrast at low frequencies due to excessive self-cancellation of the array which does not fulfil the positioning requirements for increased efficiency of radiation into the bright zone. However, the array’s positioning alleviates the effect of main beam-splitting close to the aliasing limit (3.43 kHz), thus increasing contrast at higher frequencies compared to the Null-Split in the 2×2 case. In the 3×50 case (Fig. 6.1c), the Near-Align (R-R-R) configuration always gives higher contrast than the A-A-R and A-R-R arrays, but strong periodic variations of contrast are observed. The variations are due to the placement of the array at close distance from the surface, which generates strong reflections that interfere with the direct sound.
## Chapter 6. Evaluation of reflection control strategies

<table>
<thead>
<tr>
<th>Surface</th>
<th>Method</th>
<th>Contrast (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2×2</td>
</tr>
<tr>
<td>None</td>
<td>ACC A-A-A</td>
<td>54.6</td>
</tr>
<tr>
<td></td>
<td>ACC R&lt;sub&gt;M&lt;/sub&gt;-R-R</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>ACC A-A-R</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td>ACC A-R-R</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>ACC R-R-R</td>
<td><strong>33.3</strong></td>
</tr>
<tr>
<td></td>
<td>SPM R-R-R</td>
<td>12.1</td>
</tr>
<tr>
<td>North</td>
<td>ACC R&lt;sub&gt;M&lt;/sub&gt;-R-R</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>ACC A-A-R</td>
<td>38.2</td>
</tr>
<tr>
<td></td>
<td>ACC A-R-R</td>
<td>34.4</td>
</tr>
<tr>
<td></td>
<td>ACC R-R-R</td>
<td><strong>42.8</strong></td>
</tr>
<tr>
<td></td>
<td>SPM R-R-R</td>
<td>15.9</td>
</tr>
<tr>
<td>East</td>
<td>ACC R&lt;sub&gt;M&lt;/sub&gt;-R-R</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>ACC A-A-R</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>ACC A-R-R</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>ACC R-R-R</td>
<td><strong>35.3</strong></td>
</tr>
<tr>
<td></td>
<td>SPM R-R-R</td>
<td>13.1</td>
</tr>
<tr>
<td>West</td>
<td>ACC A-A-R</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>ACC R-R-R</td>
<td><strong>30.4</strong></td>
</tr>
<tr>
<td>East, North</td>
<td>ACC A-A-R</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td>ACC R-R-R</td>
<td><strong>20.5</strong></td>
</tr>
</tbody>
</table>

Table 6.2: Frequency averaged acoustic contrast produced by systems based on the acoustic contrast control (ACC) and sound power minimization (SPM) methods. A and R denote anechoic and reflective conditions at source position optimization—source weight calculation—sound field evaluation stages respectively, and the subscript M denotes the median value.
6.2. Contrast and effort performance

Figure 6.1: Contrast produced by ACC-based systems with a single surface in the North location: (a) 2\times2, (b) 2\times50, and (c) 3\times50. Conditions: A-A-A (thin, dot-dash, red), R-R-R (thick, solid, green), A-R-R (thin, solid, magenta), A-A-R (thin, dot-dash, blue), and ACC R_M-R-R (thin, dash, black).

producing pressure nulls in the bright zone at a number of frequencies (two of such pressure nulls are visible in Fig. 5.9a). Such contrast variability may be undesirable when listening to broadband signals, which is almost always the case in a practical system. Therefore, even though the Near-Align configuration produces the largest mean contrast in this case, one may consider to employ Far-Align or Null-Split solutions instead which would provide a more regular contrast response.

In the 2\times2 case, A-R-R provides some benefit over A-A-R at low frequencies, but multiple peaks and dips lower the A-R-R’s contrast in the remainder of the frequency range. The contrast curve is much smoother in the 2\times50 case (Fig. 6.1b) thanks to additional setup sensors; however, the advantage from reflection attenuation is still limited to low frequencies. This is different in the 3\times50 case (Fig. 6.1c), where the additional source allows taking a full advantage from considering the reflections at the source weight calculation stage—the A-R-R array’s contrast is consistently much higher than that produced by the A-A-R array. Considering reflections at the system setup stage is therefore beneficial for contrast, but not when the minimal number of sources and sensors are available.

Fig. 6.2 shows the acoustic contrast produced by the ACC R-R-R configurations com-
Chapter 6. Evaluation of reflection control strategies

Figure 6.2: Contrast produced by systems with a single surface in the East location: (a) 2×2, (b) 2×50, and (c) 3×50. Methods and conditions: ACC R-R-R (thick, green) and SPM R-R-R (thin, dark-red).

Compared to the SPM-based systems, when the surface is located in the East position. SPM exhibits regular variations of contrast over frequency, but the amplitude of these variations is relatively low (2–3 dB). The envelope of contrast remains consistent over frequency, reaching approximately 20 dB. However, not much benefit is gained from adding setup sensors or sources compared to the ACC-based systems (compare Figs. 6.2a, 6.2b, and 6.2c). The ACC R-R-R arrays produce higher contrast consistently over frequency, with some inherent regular variability in the 2×2 case. However, the beam-splitting at higher frequencies and the null approaching the bright zone affect contrast in this range—a particularly steep roll-off is observed in the 3×50 case.

Table 6.3 shows the mean array effort scores achieved by the examined systems. The lowest effort (results in bold) is required by the ACC A-A-R or ACC A-R-R arrays, so those with sources optimized for direct sound control. The low effort can be attributed to a suitable orientation of the arrays, which allows maintaining high efficiency of the direct sound radiation into zone A over the whole frequency range (see example in Fig. 5.6 and the related discussion). In almost every case, the median values (ACC R\textsubscript{M}-R-R) are the largest, which demonstrates that the benefits from geometrical optimization extend to achieving lower effort compared to arbitrarily chosen configurations.

The ACC R-R-R arrays require higher effort than the ACC A-A-R arrays, which shows clearly that increasing contrast closer to the levels achieved by ACC A-A-A configu-
6.2. Contrast and effort performance

Figure 6.3: Optimized ACC-based source arrays with a single surface in the West location, and the sound pressure level maps at 1 kHz: (a) 3×50 and (b) 2×50. Symbols: □ zone, × array centre, — array axis, – surface.

Contractions must come at the expense of larger power consumption. The effort increase due to ACC R-R-R arrangements is particularly large when the arrays do not fulfil the positioning requirements for increased efficiency of radiation into the bright zone, i.e. for the 2×50 and 3×50 systems with the surface in the North position, or when the Near-Align is used with the 3×50 system with the surface in the West location. The last case is particularly problematic—while the Near-Align technique improves sound attenuation in zone B, a similar region of attenuation is created in the opposite direction where zone A is located. This is clearly visible in Fig. 6.3a and does not occur in the 2×50 case, as shown in Fig. 6.3b. This result indicates that the zones should not be located along the same line perpendicular to the surface when the Near-Align technique is used with larger source arrays. Despite these disadvantages, the effort required by the ACC R-R-R arrays is in most cases lower than the ACC R_M-R-R values (except for the 2×50 system, North case).

The array effort required by the SPM arrays is similar across the North, East, and West cases, and is higher than the effort requirements for the ACC R-R-R configurations, except for the three problematic cases discussed above. This again indicates the lack of suitability of the SPM method for the considered acoustic conditions: it achieves relatively low contrast with very little effort benefit compared to the ACC-based optimized arrays.

The results for the cases with two surfaces (East combined with North and East com-
Table 6.3: Frequency averaged array effort produced by systems based on the acoustic contrast control (ACC) and sound power minimization (SPM) methods. A and R denote anechoic and reflective conditions at source position optimization—source weight calculation—sound field evaluation stages respectively, and the subscript M denotes the median value.
6.2. Contrast and effort performance

Figure 6.4: Effort produced by systems with a single surface in the North location: (a) 2×2, (b) 2×50, and (c) 3×50. Methods and conditions: ACC R-R-R (thick, green) and SPM R-R-R (thin, dark-red).

combined with West), when compared with the equivalent contrast results from Table 6.2, indicate that the ACC R-R-R configurations can achieve significant contrast gains with respect to the arrays optimized for direct sound only, while increasing the array effort only by small amount.

Figure 6.4 shows the array effort over frequency for the ACC R-R-R and SPM R-R-R cases with the surface in the North position. In both cases, the effort increases significantly at low frequencies. The methods exhibit rapid effort variability in the 2×2 case (Fig. 6.4a), which is alleviated in the 2×50 case (Fig. 6.4b) with additional setup sensors. For the 3×50 system (Fig. 6.4c), the effort of the ACC R-R-R configuration is highly irregular, with periodic peaks and dips. This characteristics has the same origin as the contrast irregularity observed for ACC R-R-R in Fig. 6.1c. In fact, the contrast dips in Fig. 6.1c correspond to the effort peaks in Fig. 6.4c. This confirms previous observations that the Near-Align arrangement with the array positioned close to the surface may be problematic when the bright and dark zones are located at the similar distance from the surface, and suggests using the Null-Split or Far-Align arrangements to achieve more uniform contrast and effort responses.

The results presented above show that the ACC method, combined with the Null-Split, Far-Align, and Near-Align techniques or their approximations, has the potential to
realize substantial gains in contrast with respect to alternative approaches to sound zone generation when strong isolated reflections are present. These gains are achieved with a reasonable effort cost in most cases. The ACC method optimized using the proposed techniques may therefore be utilized in sound zone systems in situations when it is not possible to apply passive acoustic treatment on the problematic surfaces.

## 6.3 Robustness to implementation errors

A full experimental validation of the proposed techniques is outside the scope of this work. The initial set of validation experiments would require implementing the techniques in a controlled reflective environment, with a single reflecting surface installed in an anechoic environment. Unfortunately, such laboratory conditions were not readily available to the author at the time of completing this work. One of the main aims of such experiments would be to test the techniques’ robustness to implementation errors which inevitably occur in practice. The errors which could affect the performance include source positioning errors, system misplacement with respect to the surface, variations of reflection strength and environmental conditions between the setup and playback stages, and measurement noise. Although the robustness issue can only be fully explored in experimental work, carefully designed simulations offer a feasible and valid alternative to explore the key aspects of the problem.

In this section, the techniques’ sensitivity to source and surface positioning, and to reflectivity errors is examined using simulations. The techniques are represented by 2×2 arrangements optimized for a single surface (ACC R-R-R) from Sec. 6.2, and compared with SPM (SPM R-R-R). Random source and surface position errors were drawn from independent normal distributions for $x$ and $y$ coordinates. Normally-distributed errors to the surface reflection coefficients were introduced similarly. In each case, the procedure was repeated for 100 trials and the mean acoustic contrast over all trials was obtained.

The source position errors had one standard deviation (1 s.d.) between 0.25 mm and 5 mm, which can be regarded as manufacturing tolerance versus manual placement. Surface errors had 1 s.d. between 1 mm and 100 mm, which represent minor and major
displacement in relative position. The reflection coefficient was perturbed from the default value $\gamma = 0.8$ with 1 s.d. between 0.02 and 0.05. Errors were applied either before or after source weight calculation, i.e. calibrated errors at setup or uncalibrated error at playback.

The mean contrasts are shown in Fig. 6.5. It can be noticed that the techniques are generally most sensitive to source position errors, although the contrast degrades gracefully as in anechoic conditions. For instance, for the calibrated Null-Split technique (ACC R-R-R, North), the largest source and surface errors resulted in contrasts of 25.2 dB and 32.4 dB (losses of 8.1 dB and 0.9 dB) respectively. Calibrating the ACC techniques’ source weights typically recovered at least 2 dB contrast with respect to uncalibrated position errors. This gain was most pronounced for Near-Align (ACC R-R-R, West): e.g. the largest source error gave 33.0 dB calibrated (2.3 dB loss) versus 17.0 dB uncalibrated (18.3 dB loss). The largest surface error produced 34.1 dB calibrated (1.2 dB loss) compared to 10.5 dB uncalibrated (24.8 dB loss), as the Near-Align technique relies on destructive interference between direct and reflected sound for cancellation in the dark zone. This characteristic is illustrated by Fig. 6.6—the sensitivity of the technique to uncalibrated surface position errors manifests itself in much poorer sound energy attenuation in zone B compared to the calibrated case.

From its much lower ideal-case performance, SPM was generally less sensitive to errors in terms of contrast than the ACC techniques, which nevertheless outperformed SPM in all error conditions except for uncalibrated surface errors in the ACC R-R-R, West case. Overall, the ACC techniques gave twice the contrast of SPM with large position errors. The effect of reflection coefficient errors was negligible in all cases (less than 1 dB). These results support the previous conclusion that ACC combined with the proposed source optimization techniques is the most effective sound zoning strategy with strong individual reflections for a limited number of sources.

To provide experimental validation of the above conclusion, the methods could be implemented and compared in a room with one or two surfaces much more reflective than other. This would allow to test the techniques’ robustness to the influence of additional reflections, arriving from the directions for which the source arrangement
Figure 6.5: Mean contrast for 2×2 systems subject to different random errors applied either before or after source weight calculation, i.e. calibrated errors at setup or uncalibrated error at playback. Cases: ACC A-A-A (orange), ACC R-R-R North (red), ACC R-R-R East (green), ACC R-R-R West (blue), SPM R-R-R North (magenta), SPM R-R-R East (cyan), and SPM R-R-R West (black). The errors were drawn from normal distributions, with one standard deviation values shown on the $x$ axes. The error bars are 95% confidence intervals.

may be suboptimal. Although an experimental study in such an environment was not carried out due to practical and time constraints, a pilot simulation study on this problem is presented in Appendix D.
6.4. Summary

In this chapter, a number of approaches to sound zone generation under reflective conditions were evaluated in terms of acoustic contrast and array effort. One approach was to use the ACC method, with the sources optimized for the direct sound only. Two variants were considered here: with or without the reflection cancellers. Another approach was to combine ACC with the proposed source optimization techniques, which considered both the direct sound and reflections. Yet another approach was to use the SPM method with source positions optimized for maximum contrast. These approaches were compared against the systems that set the upper performance limit, i.e. the ACC-based systems with the weights and source positions optimized for and operating under anechoic conditions. The performance of non-optimized configurations was represented by the median contrast and effort values calculated from the scores achieved by all the configurations previously considered in the numerical optimization search under reflective conditions. The evaluation was carried out for the 2×2, 2×50, and 3×50 systems, with up to two strongly reflecting surfaces generating first order reflections.

It was found that the ACC method, combined with the Null-Split, Far-Align, and Near-Align techniques, or their approximations, can provide substantial contrast gains.
with respect to other systems, largely reducing the gap to the upper performance limit (to 19 dB on average). Compared to the SPM-controlled systems, the ACC-based arrangements optimized using the proposed techniques produced on average 34 dB higher contrast. Furthermore, it was shown that the ACC method with the sources optimized considering the direct sound only can also outperform the SPM method—the average contrast degradation with respect to the upper performance limit was 33.9 dB for ACC (with the reflection canceller employed) compared to 52.9 dB for SPM.

The $3 \times 50$ system optimized with the proposed techniques exhibited undesirable frequency characteristics in the case when the zones were at the equal distance from the surface (the North case). The optimal array generated large periodic variations of contrast across frequency. This suggests replacing the Near-Align solution with the Far-Align or Null-Split techniques in similar cases, which will trade higher average contrast for better consistency of contrast across frequency.

The systems’ performance in terms of the array effort was also examined. The lowest effort was required by the ACC-based systems with sources optimized considering only the direct sound. It was also demonstrated that the SPM method generally requires larger array effort than the ACC-based systems optimized with the proposed techniques. This provided further evidence that the SPM method is not a suitable method for controlling the sound field dominated by the direct sound and strong individual reflections. The effort required by systems based on the proposed techniques was high in some problematic cases. For instance when the sources and both zones were aligned at the right angle to the surface (the West case), the Near-Align configuration was increased effort for the $3 \times 50$ system due to symmetric sound attenuation affecting the bright zone’s sound field. This suggested that zone alignment at the right angle with respect to the surface should be avoided when the Near-Align technique is used with larger source arrays. In general, the array effort required by the proposed techniques was found to be lower (by 6.7 dB on average) than the median value representing the probable effort that can be achieved by an arbitrarily chosen source arrangement. The effort results for the cases with two surfaces generating strong first order reflections showed that the Null-Split, Far-Align and Near-Align techniques can significantly increase the contrast at very little effort cost.
Finally, the techniques’ robustness to random source and surface position, as well as surface reflectivity errors, was examined. It was found that the ACC systems optimized using the techniques are generally most sensitive to source position errors. However, calibrating the techniques’ source weights recovered contrast (typically at least 2 dB) with respect to uncalibrated errors. The Near-Align technique was found particularly sensitive to uncalibrated surface position errors. The SPM configurations were generally less sensitive to source and surface position errors than ACC. Nevertheless, the ACC systems outperformed the SPM-based configurations—on average, the former produced twice the contrast of the latter. The effect of reflection coefficient errors was negligible in all cases.

The results presented in this chapter demonstrated the potential of the proposed source optimization techniques for controlling strong individual reflections in small practical systems used for sound zone reproduction in domestic rooms. A pilot study investigating the application of the techniques to such systems is presented in Appendix D.
Chapter 6. Evaluation of reflection control strategies
Chapter 7

Conclusions and further work

This thesis concerned the problem of sound zone reproduction in reflective rooms with a limited number of sources. The primary focus of the thesis was to devise and evaluate the methods for overcoming the practical limitations on the number of sources in a sound zone system, so that the psychoacoustic requirements previously indicated in the literature can be satisfied. The following research questions were addressed:

1. What is the most suitable method and source arrangement for a domestic sound zone system with practical limitations?

2. How can the source positions be optimized for a sound zone system with a limited number of sources, operating in a reflective room?

3. How do different strategies for sound zone generation under reflective conditions compare?

In Chapter 2, the existing sound zone methods were reviewed. Examples of method evaluation available in the literature were discussed and method characteristics scrutinized in the context of small systems and reflective rooms. It was concluded that although a large ensemble of different sound zoning methods is available, none can satisfy the performance requirements using a system with practical limitations without additional optimization. One important aspect of system optimization is the placement of loudspeakers with respect to the zones and room surfaces. Chapter 3 provided
a necessary background for investigating this problem. Key sound zone methods were implemented on two different loudspeakers arrays, compact (linear) and spatially distributed (circular), in a moderately reflecting room. The study revealed that the ACC method implemented on the compact line array produced the largest acoustic contrast between the zones, while maintaining relatively uniform sound pressure levels across the bright zone and high planarity. This result, together with the conclusions from the literature review, indicated the answer to the first research question. The optimization of source positions for ACC implemented on compact arrays was therefore the main focus of Chapters 4 and 5.

In Chapter 4, the source optimization problem was studied analytically. A very small system (2×2) with a single reflective surface was considered. The motivation for such a study was to examine the fundamental relationship between system performance and its geometry, so that source optimization techniques that address the principal problems related to reflections could be devised. Such techniques are applicable to practical problems—a single strongly reflecting surface can be regarded as an approximation of the sound field in many domestic rooms, where at listening positions the direct sound and strong low order reflections are often dominant. The study resulted in three source optimization techniques and guidelines for zone positioning for maximizing the acoustic contrast and reducing the array effort. In Chapter 5, guidelines for selecting the most suitable technique for a given surface position were provided based on simulations. Moreover, source optimization solutions were extended to larger systems using the numerical search. The numerical results demonstrated the applicability of the proposed techniques to larger systems and two-surface scenarios. In this way, Chapters 4 and 5 formed a contribution which is an important step in the search for the answer to the second research question.

In Chapter 6, the proposed ACC-based source optimization techniques were evaluated and compared with alternative approaches to sound zone generation under reflective conditions. One of the approaches was to use the SPM method, which inherently reduces the radiated sound power, thus limiting the influence of reflections on performance. The effects of using ACC with arbitrarily chosen arrays and sources optimized for the direct sound only were also investigated. Performance gains offered by the pro-
posed techniques with respect to arbitrarily chosen source arrangements and the arrays controlled using existing state of the art techniques could therefore be observed. The obtained acoustic contrast and array effort results provided a useful overview of the capabilities of different methods and systems operating in the sound field dominated by the direct sound and strong first order reflections. Therefore, the material presented in Chapter 6 advanced the knowledge about performance of sound zone methods under reflective conditions, and contributed to solving system design problems related to the third research question.

The following section contains a more detailed summary of the work presented in this thesis, organized according to the contributions outlined above.

7.1 Summary

This section recapitulates on the content of this thesis and highlights the most important findings.

Chapter 3: Practical implementation study

The literature review (Chapter 2) provided insight into the characteristics and capabilities of different sound zone methods. However, it did not provide complete information about the performance of systems with a limited number of loudspeakers operating in a reflective room. For instance, the methods have not been compared on different array geometries in the domestic type of room. Moreover, some important aspects of performance, such as uniformity of the bright zone’s sound field, have not been evaluated for such systems. Therefore, compact (linear) and distributed (circular) arrays, each containing twenty-four loudspeakers, were used in a reflective room to compare the performance of the DS, ACC, and PM methods using the acoustic contrast and planarity metrics. The aim of each system was to produce high acoustic contrast between two square zones (53 cm×53 cm), while achieving high planarity and uniform sound pressure level distribution in the bright zone. The sound zone filters were band-limited to the range 0.3–3.5 kHz and the performance was evaluated by convolving the filters with an
MLS and measuring the total system response at the microphones located in the zones. The room was equipped with some *ad hoc* acoustic treatment on the surfaces, and so it was moderately reflective (with reverberation of 0.28 s time in the 1 kHz band).

The ACC method implemented on the compact line array produced the highest overall contrast, achieving approximately 20 dB in the 1–2.5 kHz range. The method maintained high bright zone planarity—over 80% in the 0.6–3 kHz range. This contributed to good uniformity of the bright zone’s sound field in this frequency range. DS performed similarly in terms of planarity, but the produced contrast was lower, particularly at low frequencies. The ACC method implemented on the circular arrangements of loudspeakers produced the largest contrast at the low end (below 500 Hz), but the contrast rolled off significantly at higher frequencies where the wavelengths became short compared to the loudspeaker spacing (45.7 cm). Moreover, the method produced low planarity and a highly non-uniform bright zone. Such a performance characteristic is undesirable, as it makes the perception of the reproduced sound programme dependent on the position of the listener in the zone. This extends the observations related to the lack of sound pressure phase control inherent in the ACC method, reported in the context of a large loudspeaker array under anechoic conditions by Coleman et al. [2014a] and in a reflective room by Coleman [2014, Chapter 3], to a smaller loudspeaker array. PM succeeded in producing a plane wave in the bright zone in the majority of the frequency range, thus producing a uniform zone. However, the contrast dropped rapidly above the spatial aliasing limit of the array, specified for the sound reproduction region encompassing the zones (508 Hz). A similar characteristic was observed for larger arrays and under anechoic conditions by Coleman et al. [2014a] and in a different room by Coleman [2014, Chapter 3].

An examination of the computational complexity of the methods revealed that DS and PM had the shortest and the longest source weight calculation times respectively, with ACC ranking between these two methods. The relatively poor computational efficiency of PM was attributed to a more complex regularization technique used.

The study indicated that the ACC method implemented on the line array is the most suitable control method for systems with a limited number of loudspeakers operating
in reflective rooms. Therefore, the method was selected for further examination and optimization in the following chapters of the thesis.

### Chapters 4 and 5: Optimization of source positions—analysis and simulations

The main aim of Chapters 4 and 5 was to devise source optimization techniques for small systems operating in a reflective room. Results of numerical optimization under such conditions were presented by Coleman et al. [2012] and Coleman [2014, Chapter 6] ([cf. Coleman et al., 2014b]). However, these results were enclosure-specific and limited to a particular source candidate set, and a systematic examination of the problem was required to devise more generally applicable solutions. Based on the results and observations in the literature, it was noted that it is often the case in domestic rooms that the sound field is dominated by the direct sound and strong low order reflections. It was therefore concluded that an investigation into this kind of sound environment could offer techniques that will be effective in typical domestic listening rooms.

Chapter 4 focused on the analysis of a small, 2×2 system, based on the ACC method. The aim of this study was to provide the algebraic framework for the analysis of system geometry. This allowed establishing the relationship between the system’s performance in terms of acoustic contrast and array effort and the geometrical parameters of the system operating in the presence of a reflecting surface.

The acoustic contrast expression was first analysed to devise geometrical methods of increasing the contrast through minimization of the sound pressure in the dark zone. First, the array directivity was optimized, resulting in two source positioning techniques: Null-Split (pointing the directivity null at the surface and exploiting the array symmetry) and Far-Align (null sharing between the dark zone and surface). The third technique, Near-Align, optimized the operation of the reflection canceller by taking advantage of the spatial match between the sources and their images with respect to the dark zone. The endfire orientation of the array, inherent in the Far- and Near-Align techniques, provided better control over the direct sound component than in the Null-Split case. Furthermore, it was shown that enhancement of contrast is possible
by ensuring appropriate positioning of the bright zone with respect to the optimized source array. Such positioning can also lead to a significant reduction of the array effort, as demonstrated by the analysis of the array effort expression.

In Chapter 5, simulations were employed to evaluate the proposed source optimization techniques in a number of geometrical scenarios and formulate guidelines for choosing the most suitable one for a given positioning of the surface with respect to the zones. It was found that the bright zone positioning requirements for enhanced contrast and low effort, formulated based on the analysis in Section 4.6, were an important selection criterion. The following conclusions were drawn:

- if the distance of both zones from the surface is the same or similar, the sources should be positioned according to the Null-Split rule;
- if the surface is closer to zone B than zone A, the Far-Align arrangement should be chosen; and
- if zone A is closer to the surface than zone B, the Near-Align technique applies.

Further in Chapter 5, a numerical search was carried out to validate the proposed techniques and test their applicability to larger systems and systems with additional surfaces. The 2×2, 2×50, and 3×50 systems with up to three surfaces were considered. The objective function for the search was the frequency-averaged contrast. Initially, the search was restricted to compact and regular arrays. The selected configurations were in good agreement with the analytic results in all single-surface scenarios. With two surfaces, the optimal arrays were shown to form a combination of the proposed techniques to minimize the influence of each of the surfaces on performance. With three surfaces, the selected arrays sought a geometrical compromise between the three techniques. The search on the expanded source candidate set with a single surface increased the possibility of finding alternative solutions (e.g. non-compact, irregular arrays) which might have been overlooked when the smaller candidate set was used. The results confirmed the optimality of the proposed techniques. However, the optimal arrays had wide and/or irregular spacing in a number of cases—such spacings were not predicted from the analysis of the contrast and effort expressions. A close analysis
revealed that larger and irregular spacing was beneficial in shaping the frequency response of contrast for the three-source array. In contrast, wide spacing did not provide much benefit when two sources were used.

Chapter 6: Evaluation of reflection control strategies

The aim of Chapter 6 was to quantify the performance gains from using the proposed source optimization techniques and compare them with other approaches to sound zone generation under reflective conditions, including the SPM method which can be regarded as an alternative strategy for limiting the impact of reflections on performance. Different sound zone reproduction strategies have not been evaluated in detail under reflective conditions in the literature. Therefore, the evaluation and ranking of such strategies would provide useful information for a system designer. The findings of this chapter have therefore practical significance.

The evaluation focused on the contrast and effort performance of the 2×2, 2×50, and 3×50 systems. The upper performance limit was provided by the ACC-based arrays with anechoic conditions at setup and playback, and source positions optimized for such conditions. The remaining systems were evaluated with one or two surfaces present. The arrays identified as optimal by the numerical search under reflective conditions at setup and playback (Chapter 5) represented the proposed techniques in the evaluation. The techniques were compared with the ACC-based systems with sources optimized for anechoic setup and playback (i.e. the direct sound control), with or without the reflections canceller. In this way, the effect of taking into account the reflections in the source optimization process could have been better observed. Moreover, the median of contrasts and efforts, produced by all the configurations considered in the numerical search in each case, was calculated to provide an indication of the typical performance when arbitrary source positions are chosen. The SPM-based arrangements were optimized for maximum contrast for meaningful comparison with the ACC-based arrays.

It was found that the proposed techniques significantly reduced the contrast loss with respect to the upper performance limit: an average loss of 18.6 dB was observed in a single-surface case. In comparison, for the SPM-controlled systems the loss was
52.9 dB on average. The techniques also outperformed other ACC-based systems: for instance, the system with sources optimized for the direct sound only and employing the reflection canceller lost on average 33.9 dB of contrast with respect to the reference. The arrays optimized with the proposed techniques exhibited larger array effort than most other systems, but apart from three problematic cases the effort increase was relatively small. For the 2×50 and 3×50 systems with the surface equidistant from both zones, the position of the bright zone with respect to the optimized arrays did not fulfil the requirements for efficient radiation into that zone (defined in Section 4.6), which was the main reason for the effort increase. Another case that was characterized with large effort was the 3×50 system in which the sources were optimized using the Near-Align technique and aligned with both zones. The produced symmetrical attenuation pattern affected the sound pressure levels in the bright zone. This result lead to the conclusion that in order to use the Near-Align technique with more than two sources, the array should not be aligned with both zones for best performance. Examination of the results for the two-surface cases showed that the arrays optimized using the combination of the proposed techniques can achieve significant contrast gains with respect to the arrays optimized for direct sound only, while increasing the array effort only by small amount.

The examination of the techniques’ robustness to different implementation errors showed that they are generally most sensitive to source position errors. Despite contrast losses due to errors, the ACC systems that employed the techniques outperformed the SPM-based systems—on average, the former produced twice the contrast of the latter.

7.2 Further work

In this section, further work that arises from this thesis is outlined. Three main areas of possible development are identified: experimental work on source optimization, advanced numerical optimization of source positions and source optimization techniques for complex reflective environments.
7.2. Further work

7.2.1 Experimental work on source optimization

One obvious extension of the source optimization work presented in this thesis is the experimental validation of the proposed techniques. The Null-Split, Far-Align, and Near-Align arrangements were devised to handle first order reflections from one or two strongly reflecting surfaces, with the inherent assumption that other reflections are less problematic for the system’s performance. Although it was established that such acoustic conditions are likely to occur in typical domestic rooms, the influence of the additional weaker reflections was not investigated on a practical system (some initial simulation results are included in Appendix D). It is therefore essential to test the techniques in practical sound zone reproduction scenarios using a physical system in a room.

7.2.2 Advanced numerical optimization of source positions

The source optimization results presented in Chapter 5 were based on candidate sources located around the dark zone. Although the choice of such locations is rooted in the analytic findings, other locations could also be considered, for instance in the area between the zones or around the bright zone. This would increase the possibility of finding improvements in the numerically optimal solution. The results from a pilot study with a candidate set extended in this way showed that main improvements are related to lower control effort, which is achieved by allowing the sources to be located closer to the bright zone. Since the optimal positions still formed compact line arrays that approximated the Null-Split, Far-Align and Near-Align techniques, the contrast improvement with respect to the solutions based on a more restricted candidate set was not significant. However, an extension of the cancellation region was possible due to a larger distance of the compact arrays from the dark zone, which is a desirable performance characteristic. Therefore, optimal arrays based on the extended numerical search may potentially offer some performance benefits; however, locating the sources close to the bright zone may not always be feasible.

For the numerical optimization search based on the extended source candidate set, the PTA(l,r) algorithm was employed in this thesis. The method is much more com-
putationally efficient than exhaustive search or a BAB algorithm (see discussion in Section 5.2.2), which were not tractable in the considered problem. PTA(l,r) increases the chances of finding a good solution compared to the basic sequential search techniques; however, the risk of finding a local optimum could be potentially decreased by a more advanced search algorithm that operates with a similar computational effort. For instance, a genetic algorithm (GA) could be employed [Kudo and Sklansky, 2000]. GA has been successfully applied to source optimization in a related problem of binaural reproduction over loudspeakers by Bai et al. [2005].

7.2.3 Source optimization for complex reflective environments

The source optimization techniques proposed in this thesis considered only first order reflections and up to two surfaces. As discussed in Chapter 4, in a typical domestic room the majority of the reflected sound energy at a listening position comes from low order reflections. The optimization results presented in this thesis are therefore an important step in improving the performance of domestic sound zone systems. However, higher order reflections and additional surfaces should be considered to address the problem of reverberation for further performance gains. The extensions could be based on more advanced analytic and numerical investigations.

A natural step forward would be to examine systems with three strongly reflecting surfaces generating reflections of up to the second order. Initial investigations into the three-surface scenario in Section 5.2.1 indicated that geometrical adjustments to the proposed techniques are required to optimize performance. The exact requirements for such adjustments to reduce the influence of first order reflections could be determined analytically. For higher order reflections, numerical optimization could be employed to reveal some performance patterns on which advanced source positioning techniques could be based. Numerical studies could also serve as a platform for quantitative characterization of the effects of the new techniques on performance.
7.3 Conclusions

The material presented in this thesis constitutes a substantive contribution to the existing body of literature on the personal audio problem. First, the key sound zone methods have been implemented and evaluated using perceptually related physical measures in a reflective room, using loudspeaker arrays with a restricted number of sources. This study provides useful information for system designers about the performance characteristics of the main methods implemented on compact and distributed loudspeaker arrangements. Second, source optimization techniques have been proposed to increase acoustic contrast and lower the array effort produced by small systems that are influenced by strong individual reflections. This provides a significant contribution to the source optimization problem, which has previously been examined in case studies. Finally, an ensemble of methods and systems was evaluated in simulated reflective conditions, resulting in a previously unpublished performance ranking which provides valuable guidelines for system designers.

As outlined in the previous section, further work in the areas of experimental and advanced numerical source optimization, as well as source optimization in complex reflective environments is required for a widespread success of the personal sound technology for domestic listeners. The research presented in this thesis has provided a solid foundation for such work.
Appendix A

Derivation of contrast and effort equations

In Sec. 4.2, analytical expressions for the source weights of the ACC-based 2×2 system were derived. In Sections 4.3 and 4.4, the acoustic contrast and array effort expressions for this system were presented. In the following, the detailed derivation of the expressions is carried out.

A.1 Definitions

A.1.1 Transfer functions: general definitions

Under anechoic conditions, the transfer function between the $l$th monopole source at $\mathbf{x}_l$ and the $m$th setup sensor in zone B at $\mathbf{x}_{Bm}$ is

$$G_{Bml} = K g_{Bml},$$  \hspace{1cm} (A.1)

where $K = j\rho ck/4\pi$ ($\rho$—air density, $c$—speed of sound, $k$—wavenumber), $g_{Bml} = e^{-jk r_{Bml}}/r_{Bml}$, and $r_{Bml} = |\mathbf{x}_{Bm} - \mathbf{y}_l|$. With a single rigid surface present, the transfer function becomes

$$G_{Bml} = K \left( g_{Bml} + g'_{Bml} \right),$$  \hspace{1cm} (A.2)
where \( g'_{Bml} = \gamma e^{-jk'r'_{Bml}} / r'_{Bml} \), \( \gamma \) is the magnitude of the surface’s reflection coefficient, \( r'_{Bml} = |x_{Bm} - y'_l| \), and \( y'_l \) is the position of the image source corresponding to the \( l \)th source. By analogy, under anechoic conditions, the transfer function between the \( l \)th source at \( x_l \) and the \( n \)th monitor sensor in zone B at \( x_{Bn} \) is

\[
\Omega_{Bnl} = K \omega_{Bnl}, \tag{A.3}
\]

where \( \omega_{Bnl} = e^{-jks_{Bnl}} / s_{Bnl} \), and \( s_{Bnl} = |x_{Bn} - y'_l| \). With a rigid surface, the transfer function expands to read

\[
\Omega_{Bnl} = K \left( \omega_{Bnl} + \omega'_{Bnl} \right). \tag{A.4}
\]

where \( \omega'_{Bnl} = \gamma e^{-jk's'_{Bnl}} / s'_{Bnl} \), \( s'_{Bnl} = |x_{Bn} - y'_l| \), and \( y'_l \) is the position of the image source corresponding to the \( l \)th source. Transfer functions for zone A are defined similarly.

### A.1.2 Transfer functions: a 2×2 system

For a 2×2 system, the vector of transfer functions between the sources and the setup sensor in zone B is written as

\[
G_B = \begin{bmatrix} G_{B1} & G_{B2} \end{bmatrix}, \tag{A.5}
\]

where the sensor index was omitted in the subscript to simplify notation. \( G_{B1} \) and \( G_{B2} \) are defined by Eq. (A.1) under anechoic conditions, and Eq. (A.2) when the surface is present. By analogy, the vector of transfer functions between the sources and the \( n \)th monitor sensor in zone B is

\[
\Omega_B = \begin{bmatrix} \Omega_{Bn1} & \Omega_{Bn2} \end{bmatrix}, \tag{A.6}
\]

where \( \Omega_{Bn1} \), \( \Omega_{Bn2} \) are defined by Eq. (A.3) under anechoic conditions, and Eq. (A.4) when the surface is present. Assuming that the sources and their images are far from the zone B’s setup sensor (see Fig. 4.1), the components of the transfer functions between each source and that sensor can be written as

\[
\begin{align*}
g_{B1} &= \frac{e^{-jk(r_B + d \cos \phi_B)/2}}{r_B}, & g_{B2} &= \frac{e^{-jk(r_B - d \cos \phi_B)/2}}{r_B} , \\
g'_{B1} &= \frac{\gamma e^{-jk(r'_B + d \cos \phi'_B)/2}}{r'_B}, & g'_{B2} &= \frac{\gamma e^{-jk(r'_B - d \cos \phi'_B)/2}}{r'_B}.
\end{align*} \tag{A.7}
\]
A.2. Derivations

where \( r_B \) and \( r'_B \) are the distances between the sensor and the physical and image source array centres respectively; \( \phi_B \) and \( \phi'_B \) are the angles between the array axes and the lines between the sensor and the array centres, for the physical and image arrays respectively; and \( d \) is the source spacing. Note that there is no source indexing in the symbols of the distances and angles as they are now measured with respect to the array centre. By analogy, the components of the transfer functions between each source and a monitor sensor in zone B can be written as

\[
\omega_{Bn1} = \frac{e^{-jk(s_{Bn}+d\cos\theta_{Bn}/2)}}{s_{Bn}}, \quad \omega_{Bn2} = \frac{e^{-jk(s_{Bn}-d\cos\theta_{Bn}/2)}}{s_{Bn}},
\]

\[
\omega'_{Bn1} = \frac{\gamma e^{-jk(s'_{Bn}+d\cos\theta'_{Bn}/2)}}{s'_{Bn}}, \quad \omega'_{Bn2} = \frac{\gamma e^{-jk(s'_{Bn}-d\cos\theta'_{Bn}/2)}}{s'_{Bn}},
\]

where \( s_{Bn}, s'_{Bn}, \theta_{Bn}, \) and \( \theta'_{Bn} \) are the monitor sensor’s equivalents to \( r_B, r'_B, \phi_B, \) and \( \phi'_B \) respectively. Transfer function components for sensors in zone A are defined similarly.

A.2 Derivations

A.2.1 Optimal source weights

The solution to the regularized ACC problem for a 2×2 system is the eigenvector \( \hat{q} \) corresponding to the largest eigenvalue of the matrix \( (G_B^H G_B + \beta I)^{-1} G_A^H G_A \) (see sub-section “Acoustic contrast control” in Section 2.3.1). The component matrices are defined as

\[
(G_B^H G_B + \beta I)^{-1} = \frac{1}{\sigma} \begin{bmatrix} G_{B2}^* G_{B2} + \beta & -G_{B1}^* G_{B1} \\ -G_{B1}^* G_{B1} & G_{B1}^* G_{B1} + \beta \end{bmatrix},
\]

\[
G_A^H G_A = \begin{bmatrix} G_{A1}^* G_{A1} & G_{A1}^* G_{A2} \\ G_{A2}^* G_{A1} & G_{A2}^* G_{A2} \end{bmatrix},
\]

where \( \sigma = (G_{B1}^* G_{B1} + \beta) (G_{B2}^* G_{B2} + \beta) - G_{B1}^* G_{B2} G_{B2}^* G_{B1} \). For convenience, the following notation can be used:

\[
(G_B^H G_B + \beta I)^{-1} = G^B = \begin{bmatrix} G_1^B & G_2^B \\ G_3^B & G_4^B \end{bmatrix},
\]
The optimal vector \( \hat{q} \) is

\[
G_A^H G_A = G^A = \begin{bmatrix} G_1^A & G_2^A \\ G_3^A & G_4^A \end{bmatrix}.
\] (A.12)

The product of the matrices \( G^B \) and \( G^A \) is

\[
G^B G^A = \begin{bmatrix} G_1^B & G_2^B \\ G_3^B & G_4^B \end{bmatrix} \begin{bmatrix} G_1^A & G_2^A \\ G_3^A & G_4^A \end{bmatrix} = \begin{bmatrix} G_1^B G_1^A + G_2^B G_2^A & G_1^B G_2^A + G_2^B G_4^A \\ G_3^B G_1^A + G_4^B G_2^A & G_3^B G_2^A + G_4^B G_4^A \end{bmatrix}.
\] (A.13)

The eigenvalues \( \kappa \) of \( G^B G^A \) are the roots of the following characteristic equation:

\[
\det(G^B G^A - \kappa I) = 0.
\] (A.14)

Expanding the LHS:

\[
\begin{align*}
(G_1^B G_1^A + G_2^B G_3^A - \kappa) & (G_3^B G_2^A + G_4^B G_4^A - \kappa) - (G_1^B G_2^A + G_2^B G_4^A) (G_3^B G_1^A + G_4^B G_3^A) = 0 \\
\therefore \quad G_1^B G_1^A G_3^B G_2^A + G_1^B G_1^A G_4^B G_4^A - \kappa G_1^B G_1^A + G_2^B G_3^B G_3^A G_2^A + G_2^B G_3^B G_4^A G_4^A - \kappa G_2^B G_3^A \\
- \kappa G_3^B G_2^A - \kappa G_4^B G_4^A + \kappa^2 - G_1^B G_2^A G_3^B G_1^A - G_1^B G_2^A G_4^B G_3^A - G_2^B G_4^B G_3^A G_1^A \\
- G_2^B G_4^A G_4^B G_3^3 & = 0
\end{align*}
\] (A.15)

According to Eqs. (A.10) and (A.12), \( G_1^A G_4^A = G_2^A G_3^A \), and so Eq. (A.15) simplifies to

\[
\kappa^2 + (-G_1^B G_1^A - G_2^B G_3^A - G_3^B G_2^A - G_4^B G_4^A) \kappa = 0.
\] (A.16)

The roots of this equation are:

\[
\kappa_1 = 0
\] (A.17)

\[
\kappa_2 = G_1^B G_1^A + G_2^B G_3^A + G_3^B G_2^A + G_4^B G_4^A.
\]

The optimal vector \( \hat{q} \) corresponds to the largest eigenvalue \( \kappa_2 \). To find the elements of \( \hat{q} \), either equation from the following system must be solved:

\[
[G^B G^A - \kappa_2 I] \hat{q} = 0,
\] (A.18)

\[
\begin{align*}
G_1^B G_1^A + G_2^B G_3^A - \kappa_2 & \quad G_1^B G_2^A + G_2^B G_4^A \\
G_3^B G_1^A + G_4^B G_3^A & \quad G_3^B G_2^A + G_4^B G_4^A - \kappa_2
\end{align*}
\]

The first of the set of equations is

\[
(G_1^B G_1^A + G_2^B G_3^A - \kappa_2) \hat{q}_1 + (G_1^B G_2^A + G_2^B G_4^A) \hat{q}_2 = 0.
\] (A.19)
Substituting $κ_2$ and rearranging yields


Substituting the matrix elements from Eqs. (A.9) and (A.10) yields

$$\frac{\hat{q}_2}{\hat{q}_1} = \frac{G_B G_A^2 G_A^1 G_A^2 - G_B G_A^2 G_A^1 G_A^2}{-G_B G_A^2 G_A^1 G_A^2 + G_B G_A^2 G_A^1 G_A^2},$$

$$\therefore \frac{\hat{q}_2}{\hat{q}_1} = \frac{G_B G_A^2 G_A^1 G_A^2 - G_B G_A^2 G_A^1 G_A^2 - \beta^2 G_A^2 G_A^2}{G_B G_A^2 G_A^1 G_A^2 - G_B G_A^2 G_A^1 G_A^2}.$$  \hfill (A.21)

With $β \to 0$, i.e., with negligible regularization, Eq. (A.21) simplifies to

$$\frac{\hat{q}_2}{\hat{q}_1} = \frac{G_B G_A^1 G_A^2}{G_B G_A^2}.$$  \hfill (A.22)

The normalised (unit) optimal source weight vector can therefore be written as

$$\hat{q} = \frac{1}{\Lambda} \begin{bmatrix} G_B G_A^2 \\ -G_B G_A^1 \end{bmatrix},$$  \hfill (A.23)

where $\Lambda = \sqrt{G_B G_A^2 + G_B G_A^1}$. Using the anechoic transfer function definitions from Eq. (A.1) and the far-field approximation from Eq. (A.7), Eq. (A.23) becomes

$$\hat{q} = \frac{K}{\Lambda} \begin{bmatrix} e^{-j k (r_B - d \cos φ_B')/2} \\ -e^{-j k (r_B + d \cos φ_B)/2} \end{bmatrix},$$  \hfill (A.24)

where

$$\Lambda = \sqrt{2 |K| / r_B}.$$  \hfill (A.25)

Using the transfer function definitions from Eq. (A.2), i.e. for the case with the surface present, and the far-field approximation from Eq. (A.7), Eq. (A.23) expands to

$$\hat{q} = \frac{K}{\Lambda} \begin{bmatrix} e^{-j k (r_B - d \cos φ_B')/2} \\ -e^{-j k (r_B + d \cos φ_B)/2} \end{bmatrix},$$  \hfill (A.26)

where

$$\Lambda = \sqrt{2 |K| \sqrt{\frac{1}{r_B^2} + \frac{2γ cos(kd (cos φ_B - cos φ_B')/2) cos(k(r_B - r_B')/2)}{r_B r_B'}} + \frac{γ^2}{r_B^2}}.$$  \hfill (A.27)
A.2.2 Sound pressure

According to the principle of superposition, complex sound pressure at the nth monitor sensor in zone B, generated by the 2×2 system using the optimal source weights, can be written as

\[ \hat{o}_{Bn} = \Omega_B \hat{q}. \quad (A.28) \]

Substituting Eqs. (A.6) and (A.23) yields

\[ \hat{o}_{Bn} = \frac{1}{\Lambda} \left( \Omega_{Bn1} G_{B2} - \Omega_{Bn2} G_{B1} \right). \quad (A.29) \]

Complex sound pressure at a monitor sensor in zone A is defined similarly.

Anechoic zone generation scenario

Substituting Eqs. (A.1) and (A.3) to Eq. (A.29) yields the complex sound pressure at the nth monitor sensor in zone B, generated by under the anechoic zone generation scenario (anechoic conditions at the setup and playback stages):

\[ \hat{o}_{Bn} = \frac{K^2}{\Lambda} \left( \omega_{Bn1} g_{B2} - \omega_{Bn2} g_{B1} \right) = \frac{K^2}{\Lambda} \hat{O}_{Bn}^D, \quad (A.30) \]

where \( \Lambda \) is the normalization coefficient for the anechoic conditions at setup, and \( \hat{O}_{Bn}^D \) is the complex direct sound pressure component due to the operation of the direct sound canceller. With the far-field approximation from Eqs. (A.7) and (A.8), \( \Lambda \) is defined by Eq. (A.25), and \( \hat{O}_{Bn}^D \) is

\[
\begin{align*}
\hat{O}_{Bn}^D &= \omega_{Bn1} g_{B2} - \omega_{Bn2} g_{B1} \\
&= \frac{s_{Bn1} r_B}{e^{-jk(s_{Bn}+r_B)}} - \frac{s_{Bn2} r_B}{e^{-jk(-d\cos\theta_{Bn}/2 + r_B + d\cos\phi_{Bn}/2)}} \\
&= \frac{2je^{-jk(s_{Bn}+r_B)} \sin(kd(\cos\phi_{B} - \cos\theta_{Bn})/2)}{s_{Bn} r_B}. \\
\end{align*}
\]

Eq. A.30 can be used to define the modulus squared sound pressure at the nth monitor sensor in zone B:

\[
|\hat{o}_{Bn}|^2 = \hat{o}_{Bn} \hat{o}_{Bn}^* = \frac{|K|^4}{|\Lambda|^2} \hat{O}_{Bn}^D \hat{O}_{Bn}^{D^*} = \frac{4|K|^4}{|\Lambda|^2} (O_{Bn}^D)^2, \quad (A.32)
\]
where $(\hat{O}_{\text{Bn}}^D)^2$ is the squared magnitude component for $\hat{O}_{\text{Bn}}^D$, defined in the far field as

$$
(\hat{O}_{\text{Bn}}^D)^2 = (\hat{O}_{\text{Bn}}^D \hat{O}_{\text{Bn}}^{D*}) / 4 = \frac{\sin^2 (kd (\cos \phi_B - \cos \theta_{\text{Bn}}) / 2)}{(s_{\text{Bn}r_B})^2}. \tag{A.33}
$$

Sound pressures in zone A are defined similarly.

### Anechoic-reflective zone generation scenario

Substituting Eqs. (A.1) and (A.4) to Eq. (A.29) yields the complex sound pressure at the $n$th monitor sensor in zone B, generated under the anechoic-reflective zone generation scenario (anechoic conditions at the setup stage and reflective conditions at the playback stage):

$$
\hat{O}_{\text{Bn}} = \frac{K^2}{\Lambda} \left[ (\omega_{\text{Bn}1} + \omega_{\text{Bn}1}') g_{\text{B2}} - (\omega_{\text{Bn}2} + \omega_{\text{Bn}2}') g_{\text{B1}} \right] = \frac{K^2}{\Lambda} \left[ (\omega_{\text{Bn}1}g_{\text{B2}} - \omega_{\text{Bn}2}g_{\text{B1}}) + (\omega_{\text{Bn}1}'g_{\text{B2}} - \omega_{\text{Bn}2}'g_{\text{B1}}) \right] \tag{A.34}
$$

where $\Lambda$ is defined for the anechoic conditions at setup, and $\hat{O}_{\text{Bn}}^{D'}$ is the complex reflected sound pressure component due to the operation of the direct sound canceller. With the far-field approximation from Eqs. (A.7) and (A.8), $\Lambda$ is defined by Eq. (A.25), and $\hat{O}_{\text{Bn}}^{D'}$ is

$$
\hat{O}_{\text{Bn}}^{D'} = \omega_{\text{Bn}1}g_{\text{B2}} - \omega_{\text{Bn}2}g_{\text{B1}},
$$

so

$$
\hat{O}_{\text{Bn}}^{D'} = \omega_{\text{Bn}1}g_{\text{B2}} - \omega_{\text{Bn}2}g_{\text{B1}} = \frac{2 j \gamma e^{-jk(s_{\text{Bn}r_B}^2 + s_{\text{Bn}2}^2)}}{s_{\text{Bn}r_B}} \sin \left( kd \left( \cos \phi_B - \cos \theta_{\text{Bn}}' \right) / 2 \right), \tag{A.35}
$$

Eq. A.34 can be used to define the modulus squared sound pressure at the $n$th monitor sensor in zone B:

$$
|\hat{O}_{\text{Bn}}|^2 = \hat{O}_{\text{Bn}} \hat{O}_{\text{Bn}}^* = \frac{K^2 (K^*)^2}{\Lambda \Lambda^*} \left[ \hat{O}_{\text{Bn}}^{D'} + \hat{O}_{\text{Bn}}^{D'*} \right] \left[ \hat{O}_{\text{Bn}}^D + \hat{O}_{\text{Bn}}^{D*} \right]^* \tag{A.36}
$$

$$
= \frac{|K|^4}{|\Lambda|^2} \left[ \hat{O}_{\text{Bn}}^D \hat{O}_{\text{Bn}}^D* + \hat{O}_{\text{Bn}}^{D'} \hat{O}_{\text{Bn}}^{D'*} + \hat{O}_{\text{Bn}}^D \hat{O}_{\text{Bn}}^{D'}* + \hat{O}_{\text{Bn}}^{D'} \hat{O}_{\text{Bn}}^D* \right] \tag{A.36}
$$

$$
= \frac{4|K|^4}{|\Lambda|^2} \left[ (\hat{O}_{\text{Bn}}^D)^2 + (\hat{O}_{\text{Bn}}^{D'})^2 + \hat{O}_{\text{Bn}}^{D*} \right].
$$
where \((O_{Bn}'^2)\) is defined by Eq. (A.33), \((O_{Bn}'^2)\) is the squared magnitude component for \(O_{Bn}'\), and \(O_{Bn}^{DD'}\) is the interaction component for \(O_{Bn}'\) and \(O_{Bn}'\). The far-field definitions of components \((O_{Bn}'^2)\) and \(O_{Bn}^{DD'}\) are

\[
\left( O_{Bn}'^2 \right) = \left( \hat{O}_{Bn}' \hat{O}_{Bn}'^* \right) / 4 = \frac{\gamma^2 \sin^2 (kd (\cos \phi_B - \cos \theta_B') / 2)}{(s_B^n r_B)^2},
\]

(A.37)

and

\[
O_{Bn}^{DD'} = \left( \hat{O}_{Bn} \hat{O}_{Bn}' + \hat{O}_{Bn}' \hat{O}_{Bn}'^* / 4 \right.
\]

\[
= \frac{\gamma e^{-jk(s_B^n + r_B)} e^{jk(s_B^n + r_B)} \sin (kd (\cos \phi_B - \cos \theta_B') / 2) \sin (kd (\cos \phi_B - \cos \theta_B') / 2)}{r_B^2 s_B^n s_B'^n}
\]

\[
+ \frac{\gamma e^{-jk(s_B^n + r_B)} e^{jk(s_B^n + r_B)} \sin (kd (\cos \phi_B - \cos \theta_B') / 2) \sin (kd (\cos \phi_B - \cos \theta_B') / 2)}{r_B^2 s_B^n s_B'^n}
\]

\[
= O_{Bn}^D O_{Bn}'^D \left( e^{-jk(s_B^n - s_B'^n)} + e^{jk(s_B^n - s_B'^n)} \right) = 2O_{Bn}^D O_{Bn}'^D \cos (k (s_B^n - s_B'^n)).
\]

(A.38)

Sound pressures in zone A are defined similarly.

### Reflective zone generation scenario

Substituting Eqs. (A.2) and (A.4) to Eq. (A.29) yields the complex sound pressure at the \(n\)th monitor sensor in zone B under the reflective zone generation scenario (reflective conditions at the setup and playback):

\[
\hat{O}_{Bn} = \frac{K^2}{\Lambda} \left[ (\omega_{Bn1} + \omega_{Bn2}) (g_{B1} + g_{B2}') - (\omega_{Bn2} + \omega_{Bn1}) (g_{B1} + g_{B2}) \right]
\]

\[
= \frac{K^2}{\Lambda} \left[ (\omega_{Bn1} g_{B2} - \omega_{Bn2} g_{B1}) + (\omega_{Bn1} g_{B2} - \omega_{Bn2} g_{B1}) \right.
\]

\[
+ (\omega_{Bn1} g_{B2}' - \omega_{Bn2} g_{B1}') + (\omega_{Bn1} g_{B2}' - \omega_{Bn2} g_{B1}') \]

\[
= \frac{K^2}{\Lambda} \left[ \hat{O}_{Bn}^D + \hat{O}_{Bn}'^D + \hat{O}_{Bn}^R + \hat{O}_{Bn}'^R \right]
\]

(A.39)

where \(\Lambda\) is defined for the reflective conditions at setup, and \(\hat{O}_{Bn}^D\) and \(\hat{O}_{Bn}'^D\) are the direct and reflected complex sound pressure components due to the reflection canceller respectively. With the far-field approximation from Eqs. (A.7) and (A.8), \(\Lambda\) is defined
by Eq. (A.27), and $\hat{O}_{Bn}^R$ and $\hat{O}_{Bn}^{R'}$ are

$$\hat{O}_{Bn}^R = (\omega_{Bn1}g_{B2}' - \omega_{Bn2}g_{B1}')$$
$$= \gamma \left( \frac{e^{-jk(s_{Bn}d\cos\theta_{Bn} + 2r_B'\cos\phi_B')/2}}{s_{Bn}r_B'} - \frac{e^{-jk(s_{Bn} - d\cos\theta_{Bn} + 2r_B'\cos\phi_B')/2}}{s_{Bn}r_B'} \right)$$
$$= \frac{2j\gamma e^{-j(k(s_{Bn} + r_B')) \sin (kd(\cos\phi_B' - \cos\theta_{Bn})/2)}}{s_{Bn}r_B'},$$

and

$$\hat{O}_{Bn}^{R'} = (\omega_{Bn1}g_{B2}' - \omega_{Bn2}g_{B1}')$$
$$= \gamma^2 \left( \frac{e^{-jk(s_{Bn}d\cos\theta_{Bn} + 2r_B'\cos\phi_B')/2}}{s_{Bn}r_B'} - \frac{e^{-jk(s_{Bn} - d\cos\theta_{Bn} + 2r_B'\cos\phi_B')/2}}{s_{Bn}r_B'} \right)$$
$$= \frac{2j\gamma^2 e^{-j(k(s_{Bn} + r_B')) \sin (kd(\cos\phi_B' - \cos\theta_{Bn})/2)}}{s_{Bn}r_B'}.$$
\[ O_{Bn}^{RR'} =\]
\[
\left( O_{Bn}^R \right)^2 = \left( \hat{O}_{Bn}^R \hat{O}_{Bn}^{R*} \right) / 4 = \frac{\gamma^2 \sin^2 \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 r_B^2}, \quad (A.43)
\]

\[
\left( O_{Bn}^{R'} \right)^2 = \left( \hat{O}_{Bn}^{R'} \hat{O}_{Bn}^{R'*} \right) / 4 = \frac{\gamma^4 \sin^2 \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 r_B^2}, \quad (A.44)
\]

\[
O_{Bn}^{DR} = \left( \hat{O}_{Bn}^D \hat{O}_{Bn}^{R*} + \hat{O}_{Bn}^R \hat{O}_{Bn}^{D*} \right) / 4
\]
\[
= \frac{\gamma e^{-jk(s_{Bn}+r_B)} e^{jk(s_{Bn}+r'_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn} \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn}' \right) / 2 \right)}{s_B^2 r_B r'_B}
+ \frac{\gamma e^{-jk(s_{Bn}+r'_B)} e^{jk(s_{Bn}+r_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn} \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn}' \right) / 2 \right)}{s_B^2 r_B r'_B}
\]
\[
= O_{Bn}^D O_{Bn}^R \left( e^{-jk(r_B-r'_B)} + e^{jk(r_B-r'_B)} \right) = 2O_{Bn}^D O_{Bn}^R \cos \left( k \left( r_B - r'_B \right) \right), \quad (A.45)
\]

\[
O_{Bn}^{DR'} = \left( \hat{O}_{Bn}^D \hat{O}_{Bn}^{R'*} + \hat{O}_{Bn}^R \hat{O}_{Bn}^{D'*} \right) / 4
\]
\[
= \frac{\gamma e^{-jk(s_{Bn}+r_B)} e^{jk(s_{Bn}'+r'_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn}' \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 s_{Bn}^' r_B r'_B}
+ \frac{\gamma e^{-jk(s_{Bn}'+r_B)} e^{jk(s_{Bn}+r'_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn}' \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 s_{Bn}^' r_B r'_B}
\]
\[
= O_{Bn}^D O_{Bn}^{R'} \left( e^{-jk(s_{Bn}-s_{Bn}'+r_B-r'_B)} + e^{jk(s_{Bn}-s_{Bn}'+r_B-r'_B)} \right)
= 2O_{Bn}^D O_{Bn}^{R'} \cos \left( k \left( s_{Bn} - s_{Bn}' \right) + r_B - r'_B \right), \quad (A.46)
\]

\[
O_{Bn}^{DR'} = \left( \hat{O}_{Bn}^D \hat{O}_{Bn}^{R*} + \hat{O}_{Bn}^R \hat{O}_{Bn}^{D'*} \right) / 4
\]
\[
= \frac{\gamma e^{-jk(s_{Bn}+r_B)} e^{jk(s_{Bn}'+r'_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn}' \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 s_{Bn}^' r_B r'_B}
+ \frac{\gamma e^{-jk(s_{Bn}'+r_B)} e^{jk(s_{Bn}+r'_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn}' \right) / 2 \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} \right) / 2 \right)}{s_{Bn}^2 s_{Bn}^' r_B r'_B}
\]
\[
= O_{Bn}^{D'} O_{Bn}^R \left( e^{-jk(s_{Bn}-s_{Bn}'+r_B-r'_B)} + e^{jk(s_{Bn}-s_{Bn}'+r_B-r'_B)} \right)
= 2O_{Bn}^{D'} O_{Bn}^R \cos \left( k \left( s_{Bn} - s_{Bn}' \right) + r_B - r'_B \right), \quad (A.47)
\]
A.2. Derivations

\[ O_{Bn}^{DR'} = \left( \hat{O}_{Bn}^{DR} \hat{O}_{Bn}^{R's} + \hat{O}_{Bn}^{R} \hat{O}_{Bn}^{DR'} \right) / 4 \]
\[ = \gamma^3 e^{-jk(s_{Bn}+r_B)} e^{jk(s_{Bn}'+r_B')} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn}' / 2 \right) \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn} / 2 \right) \right) \]
\[ + \gamma^3 e^{-jk(s_{Bn}+r_B')} e^{jk(s_{Bn}'+r_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn} / 2 \right) \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn}' / 2 \right) \right) \]
\[ = O_{Bn}^{DR} O_{Bn}^{R'} \left( e^{-jk(r_B-r_B')} + e^{jk(r_B-r_B')} \right) = 2O_{Bn}^{DR} O_{Bn}^{R'} \cos \left( k \left( r_B - r_B' \right) \right), \quad (A.48) \]

and

\[ O_{Bn}^{RR'} = \left( \hat{O}_{Bn}^{R} \hat{O}_{Bn}^{R's} + \hat{O}_{Bn}^{R'} \hat{O}_{Bn}^{R'} \right) / 4 \]
\[ = \gamma^3 e^{-jk(s_{Bn}+r_B)} e^{jk(s_{Bn}'+r_B')} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn} / 2 \right) \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn}' / 2 \right) \right) \]
\[ + \gamma^3 e^{-jk(s_{Bn}+r_B')} e^{jk(s_{Bn}'+r_B)} \sin \left( kd \left( \cos \phi_B - \cos \theta_{Bn} / 2 \right) \right) \sin \left( kd \left( \cos \phi_B' - \cos \theta_{Bn}' / 2 \right) \right) \]
\[ = O_{Bn}^{R} O_{Bn}^{R'} \left( e^{-jk(s_{Bn}-s_{Bn}')} + e^{jk(s_{Bn}-s_{Bn}')} \right) = 2O_{Bn}^{R} O_{Bn}^{R'} \cos \left( k \left( s_{Bn} - s_{Bn}' \right) \right). \quad (A.49) \]

Sound pressures in zone A are defined similarly.

### A.2.3 Array effort

**Anechoic and anechoic-reflective zone generation scenarios**

Under the anechoic and the anechoic-reflective zone generation scenarios, the array effort is given by Eq. 4.14. Recall that the general expression for array effort is provided by Eq. 2.11, which is

\[ \text{Effort}_A = 10 \log_{10} \left( \frac{q_A^H q_A}{|q_A|^2} \right), \quad (A.50) \]

The scaled source weights under anechoic conditions are expressed as

\[ q_A = \frac{a}{\Lambda} \begin{bmatrix} G_{B2} \\ -G_{B1} \end{bmatrix}, \quad (A.51) \]

where \( \Lambda \) is given by Eq. (A.25), and the scaling factor \( a \) is expressed by

\[ a = \sqrt{10^{q_A/10} \frac{p_0^2}{|\tilde{p}_A|^2}}, \quad (A.52) \]
with the squared magnitude of the (unscaled) complex sound pressure at the setup sensor in zone A, \(|\hat{p}_A|^2\), defined by

\[ |\hat{p}_A|^2 = 4|K|^4 \left( \frac{P_D^D}{|\Lambda|^2} \right)^2. \]  

(A.53)

\((P_D^D)^2\) is related to \(\hat{p}_A^D\), which is the direct complex sound pressure component due to the operation of the direct sound canceller, observed at the setup sensor in zone A. \(\hat{p}_A^D\) and \((P_D^D)^2\) can be defined similarly as the equivalent components for a monitor sensor from Eqs. (A.31) and (A.33) respectively:

\[
\begin{align*}
\hat{p}_A^D &= g_{A1}g_{B2} - g_{A2}g_{B1} \\
&= e^{-jk(r_A^D+\cos \phi_A^D/2)} - e^{-jk(r_A^D-\cos \phi_A^D/2)} \frac{r_A \tau_B}{r_A \tau_B} \\
&= 2je^{-jk(r_A^D+r_B)} \sin \left( kd\left( \cos \phi_B - \cos \phi_A \right) / 2 \right) \frac{r_A \tau_B}{r_A \tau_B},
\end{align*}
\]

(A.54)

and

\[
(P_D^D)^2 = (\hat{p}_A^D \hat{p}_A^{D^*})^2 / 4 = \frac{\sin^2 \left( kd\left( \cos \phi_B - \cos \phi_A \right) / 2 \right)}{(r_A \tau_B)^2}.
\]

(A.55)

Since \(\|\hat{q}_A\| = 1\), it is clear that \(q^L_{\hat{q}_A} q_A = a^2\). The denominator of Eq. A.50 is

\[ |q_r|^2 = \frac{10^{T_{\Lambda}/10} r_A^2 P_0^2}{|K|^2}. \]

(A.56)

Substituting the above results to Eq. A.50 yields the final effort expression of Eq. 4.14:

\[
\text{Effort}^D_A = -3 + 10 \log_{10} \left( \frac{|\Lambda|^2}{2|K|^2 (P_D^D)^2 r_A^2} \right).
\]

(A.57)

**Reflective zone generation scenario**

As in the anechoic case, the general effort expression from Eq. A.50 is a starting point in the derivation. Similarly, the scaled source weights are expressed by Eq. A.51, with \(\Lambda\) given by Eq. (A.27) and the scaling factor \(a\) defined as in Eq. (A.52), where

\[
|\hat{p}_A|^2 = \frac{4|K|^4}{|\Lambda|^2} \left[ \left( P_D^D \right)^2 + \left( P_D^{D'} \right)^2 + \left( P_A^R \right)^2 + \left( P_A^{R'} \right)^2 + P_A^{DD'} + P_A^{DR} + P_A^{DR'} + P_A^{D'R} + P_A^{DR'R} \right].
\]

(A.58)

\((P_A^D)^2, (P_A^{D'})^2, (P_A^R)^2, (P_A^{D'd})^2, (P_A^{D'R})^2, (P_A^{D'R'})^2, (P_A^{D'R'R})^2\) are related to the four complex sound pressure components observed at the setup sensor in zone A: \(\hat{p}_A^D\) — direct sound pressure component due to the operation of the direct sound canceller,
A.2. Derivations

\( \hat{P}_A^{D'} \) — reflected sound pressure component due to the operation of the direct sound canceller, \( \hat{P}_A^R \) — direct sound pressure component due to the operation of the reflected sound canceller, and \( \hat{P}_A^{R'} \) — reflected sound pressure component due to the operation of the reflected sound canceller. \( \hat{P}_A^{D'}, \hat{P}_A^R, \) and \( \hat{P}_A^{R'} \) are defined similarly as the equivalent components for a monitor sensor from Eqs. (A.35), (A.40), and (A.41) respectively:

\[
\hat{P}_A^{D'} = g'_{A1}g_{B2} - g'_{A2}g_{B1} \\
= \frac{\gamma}{r'_{A}r_{B}} \left( e^{-jk(r'_{A} + d \cos \phi'_{A}/2 + r'_{B} - d \cos \phi'_{B}/2)} - e^{-jk(r'_{A} - d \cos \phi'_{A}/2 + r'_{B} + d \cos \phi'_{B}/2)} \right) \\
= \frac{2j\gamma e^{-jk(r'_{A} + r'_{B})} \sin (kd \cos \phi'_{B} - \cos \phi'_{A}/2)}{r'_{A}r_{B}},
\]

\[
\hat{P}_A^R = g_{A1}g'_{B2} - g_{A2}g'_{B1} \\
= \frac{\gamma}{r_{A}'r_{B}} \left( e^{-jk(r_{A} + d \cos \phi_{A}/2 + r'_{B} - d \cos \phi'_{B}/2)} - e^{-jk(r_{A} - d \cos \phi_{A}/2 + r'_{B} + d \cos \phi'_{B}/2)} \right) \\
= \frac{2j\gamma e^{-jk(r_{A} + r'_{B})} \sin (kd \cos \phi'_{B} - \cos \phi_{A}/2)}{r_{A}'r_{B}},
\]

and

\[
\hat{P}_A^{R'} = g'_{A1}g_{B2} - g'_{A2}g_{B1} \\
= \frac{\gamma}{r'_{A}r_{B}} \left( e^{-jk(r'_{A} + d \cos \phi'_{A}/2 + r'_{B} - d \cos \phi'_{B}/2)} - e^{-jk(r'_{A} - d \cos \phi'_{A}/2 + r'_{B} + d \cos \phi'_{B}/2)} \right) \\
= \frac{2j\gamma e^{-jk(r'_{A} + r'_{B})} \sin (kd \cos \phi'_{B} - \cos \phi'_{A}/2)}{r'_{A}r_{B}},
\]

\( (P_A^{D'})^2, (P_A^R)^2, (P_A^{R'})^2, P_A^{D'D'}, P_A^{D'R'}, P_A^D, P_A^{R'R'}, P_A^{D'D'}, P_A^{R'R'}, P_A^{R'R'}, P_A^{R'R'}, \) and \( P_A^{R'R'} \) can be defined similarly as the equivalent components for a monitor sensor from Eqs. (A.37), (A.43), (A.44), (A.38), (A.45), (A.46), (A.47), (A.48), and (A.49) respectively:

\[
(P_A^{D'})^2 = \left( \hat{P}_A^{D'} \hat{P}_A^{D'*} \right) / 4 = \frac{\gamma^2 \sin^2 (kd (\cos \phi'_{B} - \cos \phi'_{A}) / 2)}{(r'_{A}r_{B})^2},
\]

\[
(P_A^R)^2 = \left( \hat{P}_A^R \hat{P}_A^{R'*} \right) / 4 = \frac{\gamma^2 \sin^2 (kd (\cos \phi'_{B} - \cos \phi_{A}) / 2)}{(r_{A}'r_{B})^2},
\]

\[
(P_A^{R'})^2 = \left( \hat{P}_A^{R'} \hat{P}_A^{R'*} \right) / 4 = \frac{\gamma^4 \sin^2 (kd (\cos \phi'_{B} - \cos \phi'_{A}) / 2)}{(r'_{A}r_{B})^2},
\]

\[
P_A^{D'D'} = \left( \hat{P}_A^{D'} \hat{P}_A^{D'*} + \hat{P}_A^{D'} \hat{P}_A^{D'*} \right) / 4 = 2P_A^{D'}P_A^{D'} \cos (k (r_A - r'_A)),
\]
Substituting these results to Eq. A.50 gives the final effort expression of Eq. 4.15:

\[ P_{A}^{DR} = \left( \hat{P}_{A}^{D} \hat{P}_{A}^{R*} + \hat{P}_{A}^{R} \hat{P}_{A}^{D*} \right) / 4 = 2P_{A}^{D}P_{A}^{R} \cos \left( k \left( r_{B} - r'_{B} \right) \right), \]  

(A.66)

\[ P_{A}^{DR'} = \left( \hat{P}_{A}^{D'} \hat{P}_{A}^{R*} + \hat{P}_{A}^{R'} \hat{P}_{A}^{D*} \right) / 4 = 2P_{A}^{D'}P_{A}^{R} \cos \left( k \left( r_{A} - r'_{A} + r_{B} - r'_{B} \right) \right), \]  

(A.67)

\[ P_{A}^{D'R} = \left( \hat{P}_{A}^{D} \hat{P}_{A}^{R*} + \hat{P}_{A}^{R} \hat{P}_{A}^{D*} \right) / 4 = 2P_{A}^{D'}P_{A}^{R} \cos \left( k \left( r_{A} - r'_{A} - r_{B} + r'_{B} \right) \right), \]  

(A.68)

\[ P_{A}^{D'R'} = \left( \hat{P}_{A}^{D'} \hat{P}_{A}^{R*} + \hat{P}_{A}^{R'} \hat{P}_{A}^{D*} \right) / 4 = 2P_{A}^{D'}P_{A}^{R} \cos \left( k \left( r_{B} - r'_{B} \right) \right), \]  

(A.69)

and

\[ P_{A}^{RR'} = \left( \hat{P}_{A}^{R} \hat{P}_{A}^{R*} + \hat{P}_{A}^{R'} \hat{P}_{A}^{R*} \right) / 4 = 2P_{A}^{R}P_{A}^{R} \cos \left( k \left( r_{A} - r'_{A} \right) \right). \]  

(A.70)

As in the anechoic zone generation scenarios, \( q_{a}H q_{A} = a^{2} \). The denominator of Eq. A.50 is

\[ |q_{r}|^{2} = \frac{10T_{A}/10p_{0}^{2}}{|K|^{2} \left( \frac{1}{r_{A}^{2}} + \frac{2\gamma \cos \left( k \left( r_{A} - r'_{A} \right) \right)}{r_{A}r'_{A}} + \frac{\gamma^2}{r_{A}^{2}} \right)}. \]  

(A.71)

Substituting these results to Eq. A.50 gives the final effort expression of Eq. 4.15:

\[ \text{Effort}_{A}^{R} = -3 + 10 \log_{10} \left[ \left| A \right|^{2} \left( \frac{1}{r_{A}^{2}} + \frac{2\gamma \cos \left( k \left( r_{A} - r'_{A} \right) \right)}{r_{A}r'_{A}} + \frac{\gamma^2}{r_{A}^{2}} \right) / \right. \]

\[ 2|K|^{2} \left( \left( P_{A}^{D} \right)^{2} + \left( P_{A}^{D'} \right)^{2} + \left( P_{A}^{R} \right)^{2} + \left( P_{A}^{R'} \right)^{2} + P_{A}^{DD'} + P_{A}^{DR} + P_{A}^{DR'} + P_{A}^{D'R} + P_{A}^{D'R'} \right) \right]. \]  

(A.72)
Appendix B

Comparison of the BC and ACC solutions

In Sec. 5.1 it was pointed out that for an array symmetrical with respect to setup sensors in zones A and B in a 2×2 system, the solution to the regularized ACC problem is equivalent to the source weights obtained with the BC method. This is demonstrated below.

According to Eq. 2.14 pertaining to the BC method, the optimal source weight vector is proportional to the eigenvector corresponding to the largest eigenvalue of the matrix $G_A^H G_A$. This matrix is defined as

$$G_A^H G_A = \begin{bmatrix} G_{A1}^* G_{A1} & G_{A1}^* G_{A2} \\ G_{A2}^* G_{A1} & G_{A2}^* G_{A2} \end{bmatrix},$$

which can be written more compactly as

$$G_A^H G_A = G_A = \begin{bmatrix} G_A^1 & G_A^2 \\ G_A^3 & G_A^4 \end{bmatrix}.$$  \hspace{1cm} (B.1)

The eigenvalues of $G_A$ are the roots of the following equation:

$$\det(G_A - \kappa I) = 0,$$  \hspace{1cm} (B.3)

where $\kappa$ denotes an eigenvalue. Evaluation of the determinant yields

$$\kappa^2 + \kappa \left(-G_A^1 - G_A^4\right) + G_A^1 G_A^2 - G_A^2 G_A^3 = 0.$$  \hspace{1cm} (B.4)
Since $G_1^AG_4^A = G_2^AG_3^A$, the above equation simplifies to
\[ \kappa^2 + \kappa \left( -G_1^A - G_4^A \right) = 0. \] (B.5)

The roots of this equation are:
\[ \kappa_1 = 0 \]
\[ \kappa_2 = G_1^A + G_4^A. \] (B.6)

The optimal vector $\hat{q}$ corresponds to the largest eigenvalue $\kappa_2$. To find the elements of $\hat{q}$, either equation from the following system must be solved:
\[
\begin{bmatrix}
G^A - \kappa_2 I
\end{bmatrix}
\begin{bmatrix}
\hat{q}_1 \\
\hat{q}_2
\end{bmatrix} = \begin{bmatrix}
0 \\
0
\end{bmatrix}.
\] (B.7)

The first of the set of equations is
\[ (G_1^A - \kappa_2) \hat{q}_1 + G_2^A \hat{q}_2 = 0. \] (B.8)

Substituting $\kappa_2$ and rearranging yields
\[
\frac{\hat{q}_2}{\hat{q}_1} = \frac{G_1^A}{G_2^A} = \frac{G_2^A G_2^A}{G_1^A G_1^A} = \frac{G_2^A}{G_1^A}. \] (B.9)

Now, let’s compare this solution with the one for ACC when the source array is symmetrical with respect to setup sensors in zones A and B. The regularized solution is given by Eq. A.21. In the symmetrical case, we observe the following relationships:
$G_{B1} = G_{A1}$, $G_{B2} = G_{A2}$, $G_{B1}^* = G_{A1}^*$ and $G_{B2}^* = G_{A2}^*$. Therefore, Eq. A.21 simplifies to
\[
\frac{\hat{q}_2}{\hat{q}_1} = \frac{G_2^* A_1 G_1^* A_2 - G_2^* A_2 G_2^* A_2 - G_2^* A_2 G_2^* A_2}{G_1^* A_1 G_2^* A_2 - G_2^* A_2 G_2^* A_2 - G_2^* A_2 G_2^* A_2} \frac{G_2^* A_2}{G_1^* A_1}. \] (B.10)

Comparison of this result with Eq. B.9 shows the equivalence of the BC and the regularized ACC solutions.
Appendix C

Optimal source positions for compact arrays

In the following, the results of the numerical optimization of source position for compact arrays are presented in full. The optimization procedure and the employed configurations were described in Sec. 5.2.1. The results presented below supplement the material included in Sec. 5.2.1 and Chapter 6.

Figure C.1 shows the optimization results for for $2 \times 2$, $2 \times 50$, and $3 \times 50$ systems based on ACC under anechoic conditions (ACC A-A-A systems in Chapter 6), overlaid on the SPL maps at 1 kHz. As pointed out in Sec. 5.2.1, the positions and orientations of the optimal arrays are identical in all cases. The SPL maps clearly demonstrate that the large contrast observed for the ACC A-A-A, $3 \times 50$ case in Table 6.2, can be attributed to very effective attenuation of sound energy around zone B.

Figure C.2 shows the optimization results for for $2 \times 2$, $2 \times 50$, and $3 \times 50$ systems based on ACC and SPM with a single surface in different locations (ACC R-R-R and SPM R-R-R systems from Chapter 6). For ACC, the array positions and orientations are identical for all systems with the surfaces in the East and West locations, and differences are observed in the North case, as discussed in Sec. 5.2.1. Similar characteristics can be observed for SPM; however, the optimal arrays in the East and West cases differ between the two- and three-source systems. The SPL maps show that the SPM arrangements
steer the directivity nulls at zone B to maximize the contrast. Furthermore, the sound energy tends to be directed away from the surface to reduce the magnitude of reflections in zone B.

Figure C.3 shows the results for the ACC-based 2×2, 2×50, and 3×50 systems with two surfaces generating first order reflections (ACC R-R-R systems from Chapter 6). As discussed in Sec. 5.2.1, the positions and orientations of the optimal arrays are similar for all systems with a given surface combination. As with a single surface, the 3×50 attenuates sound energy very effectively in zone B with the surfaces in the East and North positions, which contributes to the large frequency-averaged contrast shown in Table 6.2.

Figure C.1: Optimized compact source arrays for 2×2 (left), 2×50 (middle), and 3×50 (right) systems under anechoic conditions, overlaid on the sound pressure level maps at 1 kHz. Symbols: □ zone, × array centre, — array axis.
Figure C.2: Optimized compact source arrays for 2×2 (left columns), 2×50 (middle columns), and 3×50 (right columns) systems based on ACC (top rows) and SPM (bottom rows) methods with a single surface, overlaid on the sound pressure level maps at 1 kHz. (a) North, (b) East, and (c) West surface locations. Symbols: □ zone, × array centre, — array axis, ■ surface.
Appendix C. Optimal source positions for compact arrays

Figure C.3: Optimized compact source arrays for 2×2 (left column), 2×50 (middle column), and 3×50 (right column) systems based on ACC with two surfaces in the East and North (top row), and the East and West locations (bottom row), generating first order reflections. Results overlaid on the sound pressure level maps at 1 kHz. Symbols as in Fig. C.2.
Appendix D

Source optimization for practical systems: pilot study

A short pilot study has been carried out on the application of the proposed source optimization techniques to systems with additional sources, located in a reflective enclosure. The main goal of this study was to investigate how the proposed techniques could be adopted to a multi-surface reflective environment, larger source arrays and two-programme scenarios. The performance of the optimized and non-optimized systems was compared.

Figures D.1a and D.1b show the test configurations in the plan and side views respectively. There were two square zones, A and B, each with the dimensions 40 cm × 40 cm. There were eighty-one setup sensors in each zone, forming a grid with 5 cm spacing. The performance was evaluated using a geometrically mismatched grid of monitor sensors. Eight sources were used at a time. The sources formed one optimized array (two lines of four sources) and seven different non-optimized arrays (a single line of eight sources in different positions). The room was simulated using the ISM technique, with up to sixth order reflections allowed. There were four walls, three with large absorption ($\gamma = 0.3$) and one with small absorption ($\gamma = 0.9$). The ceiling and floor were also highly absorptive ($\gamma = 0.3$). In this way, the proposed techniques could be tested in an enclosure with unevenly distributed absorption, for which they were designed. The control method was ACC with frequency independent regularization $\beta = 10^{-6}$. Full transfer
Appendix D. Source optimization for practical systems: pilot study

Figure D.1: Configurations used to test the proposed source optimization techniques on practical systems: (a) plan view, (b) side view.
functions, including all the reflections, were utilized for calculating the source weights at the system setup stage.

The techniques were applied with respect to the strongly-reflecting surface (East). This required the Far-Align technique for dark zone B, and Near-Align for dark zone A. It can be noticed that the zones are arranged in a slightly different way than in the simulations reported in Chapters 4, 5 and 6. This was to minimize the influence of the symmetrical attenuation on the bright zone sound field - symmetrical nulls are formed around the array when the Near-Align technique is used with larger source arrays (see discussion concerning the 3\times50 system, the West case, in Sec. 6.2). The optimal array had to be split into two subarrays to facilitate the application of the optimization techniques to both dark zones. However, in order to achieve satisfying performance at low frequencies, both subarrays were active when reproducing each target programme. The positioning of the subarrays along the x-axis was established empirically so that the peak of radiation coincides with the bright zone in each case. The positioning of the non-optimized arrays was chosen to represent a range of locations across the room.

For each array, the acoustic contrast was evaluated in the frequency range 250–3500 Hz, for both target programmes. For the non-optimized arrays, a median of all contrast values was calculated at each frequency to estimate the probable effect of using an arbitrarily chosen configuration (see discussion in Sec. 6.1). The contrast and median values were also calculated as averages over frequency. Two scenarios were considered: the room without the floor and ceiling and the complete room.

Figure D.2a shows the contrast produced by the optimized configuration and the median contrast calculated for non-optimized arrays for the case with walls only. The top and bottom plots are for target zones A and B respectively. The optimal arrays produced higher contrast in both cases: an average contrast gain of 7 dB was achieved. This result indicates that the optimal array is likely to outperform any of the chosen arbitrary arrays. Figure D.2b shows the equivalent contrast results for the complete room. Similarly, the optimized array produced larger contrast in the majority of the frequency range. The average contrast gain was 5 dB in this case.

It is also interesting to compare the SPL maps of the sound fields produced by the
Appendix D. Source optimization for practical systems: pilot study

Figure D.2: Contrast produced by the ACC-based optimized array (thick, blue) and the median contrast calculated for all non-optimized configurations (thin, green): (a) scenario with walls only, (b) the complete room.

optimized and non-optimized arrays. Figure D.3 shows the maps for the walls-only scenario at 1 kHz (an example non-optimized array is shown). Figure D.3a shows the case with target zone A. It can be noticed that the optimized array is more effective in reducing the sound energy around zone B. Although the Far-Align technique is used with respect to the strongly reflecting surface (East), the symmetric attenuation and beam splitting, characteristic for Near-Align, is also observed. This indicates the Near-Align is in operation for the reflections from the moderately reflective, opposite (West) surface. It is also noteworthy that the proposed optimal array reduces the sound energy leakage to the remainder of the room, particularly around zone B. Figure D.3b shows the maps for target zone B. The sound field produced by the optimal array is very similar to the case with target zone A. In fact, the contrast is identical for both targets at this frequency (26 dB). This is a desirable performance characteristic, as it results in a similar listening experience in both zones. The non-optimized array attenuates
Figure D.3: Sound pressure level maps at 1 kHz for the optimized array (top) and a chosen non-optimized array (bottom) in the walls-only scenario: (a) target zone A, (b) target zone B. Symbols: □ zone, × source position, ■ surface.

more effectively in the dark zone in this case—this can be attributed to the fact that the dark zone is much further from the most problematic surface.
Appendix D. Source optimization for practical systems: pilot study
Bibliography


Simón-Gálvez, M. F. and Elliott, S. J. (2013). The design of a personal audio superdi-
rective array in a room. In *Proceedings of the 52nd AES International Conference on Sound Field Control, Guildford, UK, 2–4 September*.


