KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Authors: Jesper Brunnström, Toon van Waterschoot, Marc Moonen Title: Signal-to-interference-plus-noise ratio based optimization for sound zone control Journal/Conference/Book: IEEE Open Journal of Signal Processing Vol/Book Chapter: Pages: ISBN:
Archived version	■ Submitted version ☐ Accepted (Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher) ☐ Published ☐ Published (Open Access)
Acknowledgements	In this manuscript, the following funding is acknowledged: ■ EU H2020 ☐ EU Horizon Europe ☐ FWO ■ KULeuven Internal Funds ☐ VLAIO ☐ AI ☐ Other
Published version	
Copyright statement	
Author contact	jesper.brunnstroem@kuleuven.be



IEEE Open Journal of Signal Processing

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/OJSP.2022.1234567

Signal-to-interference-plus-noise ratio based optimization for sound zone control

JESPER BRUNNSTRÖM (Graduate student member, IEEE), TOON VAN WATERSCHOOT (Member, IEEE) AND MARC MOONEN (Fellow, IEEE)

STADIUS Center for Dynamical Systems, Department of Electrical Engineering (ESAT), KU Leuven, Leuven 3001, Belgium

CORRESPONDING AUTHOR: JESPER BRUNNSTRÖM (e-mail: jesper.brunnstroem@kuleuven.be).

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of Research Council KU Leuven C14-21-0075 "A holistic approach to the design of integrated and distributed digital signal processing algorithms for audio and speech communication devices". This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 956369: 'Service-Oriented Ubiquitous Network-Driven Sound — SOUNDS'. The scientific responsibility is assumed by the authors. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program / ERC Consolidator Grant: SONORA (no. 773268). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information.

ABSTRACT Sound zone control is posed as an optimization problem where finite impulse response control filters are jointly optimized for low transmit power, while maintaining a sufficiently high signal-to-interference-plus-noise ratio (SINR) in all zones. This problem statement in particular allows the consideration of sound zone control under the additional influence of external noise, which is rarely taken into account and indeed necessitates an alternative problem statement. In addition, the problem statement allows for the spectral characteristics of the audio signals to be taken into account, both for the power minimization and the SINR specification. The optimization problem is shown to be solved optimally by semidefinite relaxation, but the computational cost is high. An alternate method is proposed, inspired by a duality between transmit and receive beamforming commonly exploited in digital communications. A virtual receive optimization problem for sound zone control is derived, and it is shown to have the same optimal solution as the original transmit optimization problem. The receive optimization problem is solved efficiently through a fixed point iteration algorithm. The two proposed methods are compared against acoustic contrast control in a simulated reverberant environment, where it is shown that the proposed methods require less transmit power to produce similar SINRs.

INDEX TERMS Sound zone control, array signal processing, duality, convex optimization, semi-definite relaxation, signal-to-interference-plus-noise ratio

I. INTRODUCTION

IN sound zone control different audio content is reproduced in different zones of the same space using a loudspeaker array [1]. Popular methods include acoustic contrast control (ACC) [2]–[5] and pressure matching [6]–[8], both of which can be described as special cases of the variable span trade-off filter [9], [10]. The task is to find a control filter for each loudspeaker and audio signal, such that when the filtered signals are transmitted by the loudspeakers, only the desired signal is heard in each zone.

Existing sound zone control methods generally solve the problem using superposition, a simplified problem formulation where only one zone has a desired signal that should be reproduced, whereas silence should be maintained in the other zones. By solving such a problem for each zone individually and superimposing the obtained solutions, the joint sound zone control problem is solved. In practice however, there is likely an additional non-negligible level of noise from external sources, which can be beneficial to take into account. In that case, the superposition strategy is

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

no longer well suited, necessitating an alternative problem statement where all zones are considered jointly.

The sound zone control problem without superposition has a close resemblance to the transmit beamforming problem in digital communications [11], [12]. There is a duality between transmit and receive beamforming commonly exploited in digital communications, also referred to as uplink-downlink duality [13]–[16]. It is particularly useful for solving transmit beamforming problems, which are in general more difficult than the corresponding receive beamforming problems [17]. An analogous duality exists for sound field reproduction, but is seldom used. It is used in [18] for sound zone control, however the method places significant restrictions on the loudspeaker and microphone arrays.

A common simplifying assumption in sound zone control is that the audio signals are stationary and spectrally white, which avoids having the control filter being dependent on the likely non-stationary audio signals [19], [20]. This is suboptimal in practice, and sound zone control performance can be improved by taking the spectral characteristics of the audio signals into account.

The first contribution of this paper is the formulation of the sound zone control problem without superposition, taking both external noise and the spectral characteristics of the audio signals into account. Two methods are then proposed to solve the problem, the first based on semi-definite relaxation (SDR), which is optimal but computationally costly. For the second method a virtual receive optimization problem for sound zone control is derived, and it is shown to have the same optimal solution as the original transmit optimization problem. It is solved efficiently through a modified version of the fixed point iteration algorithm from [15].

The paper is organized as follows. In section II the signal model and the optimization problem are formulated. In section III and section IV the two proposed methods are described. In section V experiments are designed and their results presented. Conclusions are stated in section VI. Finally, in appendix A and appendix B the virtual receive optimization problem is derived and its relationship to the transmit optimization problem is demonstrated.

A. NOTATION

Throughout the paper, vectors are denoted by lowercase bold letters (such as $\boldsymbol{a}$), matrices by uppercase bold letters (such as $\boldsymbol{A}$ or $\mathbb{A}$), and scalar variables by non-bold letters (such as a or A). The identity matrix is denoted by $\boldsymbol{I}$, and its size should be obvious from context. Time-dependent quantities are functions of the discrete time index n. The set of all real values is denoted by $\mathbb{R}$ and non-negative real values by $\mathbb{R}_{\geq 0}$. The inequality of a vector (such as $\boldsymbol{a} \geq 0$) should be interpreted element-wise, and $\succcurlyeq$ denotes a generalized inequality over the cone of positive semi-definite matrices. The expectation operator is denoted by $\mathbb{E}[\cdot]$, and the trace operator by $\mathrm{tr}(\cdot)$. The col operator creates a column vector from its arguments, starting from the top of the vector, (such as $\mathrm{col}\{a(i)\}_{i=0}^{I-1} = \begin{bmatrix} a(0) & a(1) & \dots & a(I-1) \end{bmatrix}^\top$). The diag and blkdiag operators create a diagonal and block diagonal matrix respectively from their arguments with the analogous ordering.

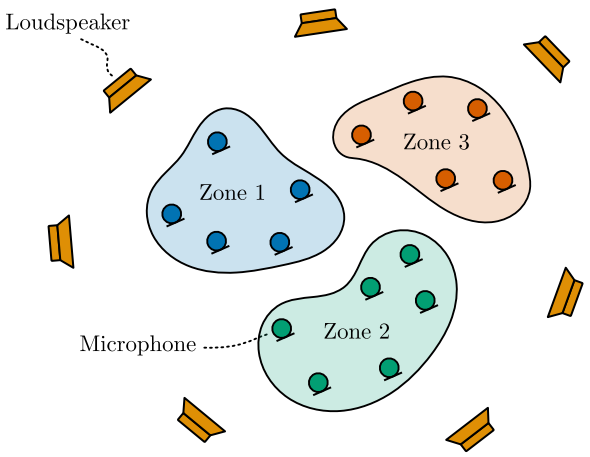

FIGURE 1. Depiction of the constituent parts of a sound zone control system with 3 zones.

II. PROBLEM STATEMENT

A. SIGNAL MODEL

An array of L loudspeakers transmits audio signals to Z zones, each with M_z microphones. The problem setting is depicted in Figure 1. The set of indices for the zones is $\mathcal{Z} = \{1, \dots, Z\}$ and for the loudspeakers $\mathcal{L} = \{1, \dots, L\}$. The monophonic audio signal $x_z(n)$ associated with zone z should be reproduced only in that zone by filtering it with the I-tap finite impulse response (FIR) filter $\boldsymbol{w}_{zl} \in \mathbb{R}^I$ before driving loudspeaker l. The audio propagates from loudspeaker l to microphone m through an acoustic path modelled by the time-invariant J-tap room impulse response (RIR) $\boldsymbol{h}_{lm} \in \mathbb{R}^J$. The signal received at microphone m is

$$y_m(n) = \eta_m(n) + \sum_{l \in \mathcal{L}} \sum_{z \in \mathcal{Z}} \boldsymbol{h}_{lm}^\top \boldsymbol{X}_z(n) \boldsymbol{w}_{zl}, \tag{1}$$

where $\eta_m(n)$ is additive external noise, and $\boldsymbol{X}_z(n) \in \mathbb{R}^{J \times I}$ is a Toeplitz matrix whose jth row and ith column is $x_z(n-i-j+2)$. The signal model can be expressed more compactly by introducing the RIR matrix

$$\boldsymbol{H}_z = \begin{bmatrix} \boldsymbol{h}_{11} & \dots & \boldsymbol{h}_{1M_z} \\ \vdots & \ddots & \vdots \\ \boldsymbol{h}_{L1} & \dots & \boldsymbol{h}_{LM_z} \end{bmatrix} \in \mathbb{R}^{JL \times M_z} \tag{2}$$

for all RIRs associated with zone z, the duplicated audio signal matrix $\mathbb{X}_z(n) = \mathrm{blkdiag}\{\boldsymbol{X}_z(n)\}_{l=1}^L$, the control filter $\boldsymbol{w}_z = \mathrm{col}\{\boldsymbol{w}_{zl}\}_{l \in \mathcal{L}}$, and the noise signal $\boldsymbol{\eta}_z(n) = \mathrm{col}\{\eta_m(n)\}_{m=1}^{M_z}$. The desired signal $\boldsymbol{y}_z^D(n)$ and interference-plus-noise $\boldsymbol{y}_z^I(n)$ for zone z are then obtained as

$$\begin{aligned} \boldsymbol{y}_z^D(n) &= \boldsymbol{H}_z^\top \mathbb{X}_z(n) \boldsymbol{w}_z \\ \boldsymbol{y}_z^I(n) &= \boldsymbol{\eta}_z(n) + \sum_{\substack{i \in \mathcal{Z} \\ i \neq z}} \boldsymbol{H}_z^\top \mathbb{X}_i(n) \boldsymbol{w}_i. \end{aligned} \tag{3}$$

B. SIGNAL-TO-INTERFERENCE-PLUS-NOISE RATIO

With the signal model (3), a natural metric for the quality of the sound zone control is the ratio of desired to undesired sound in each zone, which is the signal-to-interference-plus-noise ratio (SINR). The SINR can be written as

$$\mathrm{SINR}_z = \frac{\mathbb{E}\left[\|\boldsymbol{y}_z^D(n)\|_2^2\right]}{\mathbb{E}\left[\|\boldsymbol{y}_z^I(n)\|_2^2\right]} = \frac{\boldsymbol{w}_z^\top \boldsymbol{R}_{zz}\boldsymbol{w}_z}{\sum_{\substack{i\in\mathcal{Z}\\ i\neq z}} \boldsymbol{w}_i^\top \boldsymbol{R}_{zi}\boldsymbol{w}_i + \sigma_z^2}, \quad (4)$$

where the mean noise power of zone z is $\sigma_z^2 = \mathbb{E}[\|\boldsymbol{\eta}_z(n)\|_2^2]$, and the spatial covariance matrices are defined as

$$\boldsymbol{R}_{zi} = \mathbb{E}\left[\mathbb{X}_i^\top(n)\boldsymbol{H}_z\boldsymbol{H}_z^\top\mathbb{X}_i(n)\right] \in \mathbb{R}^{LI\times LI}. \quad (5)$$

The definition of the SINR used here to characterize a zone is using the summed power of the signals at the microphones, which is likely satisfactory if the microphones are spaced densely and evenly. Otherwise, a method such as kernel interpolation [21] could be applied according to [22]. Defining the SINR for each microphone individually instead of the zone as a whole is also possible, but leads to a non-convex NP-hard problem [23].

C. OPTIMIZATION PROBLEM

The considered optimization problem is a SINR constrained transmit power minimization problem, which allows for the desired level of audio separation to be selected via the SINR constraints. Transmit power in this context refers to the expected sum power of the signals emitted by the loudspeakers. With the loudspeaker signals denoted by $\boldsymbol{s}(n) \in \mathbb{R}^L$, the optimization problem is

$$\begin{aligned} \underset{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_Z}{\text{minimize}} \quad & \mathcal{P} = \mathbb{E}[\|\boldsymbol{s}(n)\|_2^2] + \alpha\sum_{z\in\mathcal{Z}}\|\boldsymbol{w}_z\|_2^2 \\ \text{subject to} \quad & \mathrm{SINR}_z \geq \gamma_z, \end{aligned} \quad (6)$$

where the parameter $\gamma_z \in \mathbb{R}_{\geq 0}$ sets the required SINR for the zone, and $\alpha \in \mathbb{R}_{\geq 0}$ is a parameter controlling the amount of regularization. The regularization can help mitigate problems in the case of imperfectly estimated covariance matrices. Introducing an audio signal vector $\boldsymbol{x}_z(n) = \mathrm{col}\{x_z(n-i)\}_{i=0}^{I-1}$, the audio signal covariance for zone z is $\boldsymbol{A}_z = \mathrm{blkdiag}\{\mathbb{E}[\boldsymbol{x}_z(n)\boldsymbol{x}_z^\top(n)]\}_{l=1}^L$. Under the assumption that the audio signals for each zone are mutually independent and zero mean, the objective function in (6) can be rewritten as

$$\mathcal{P} = \sum_{z\in\mathcal{Z}} \boldsymbol{w}_z^\top \tilde{\boldsymbol{A}}_z \boldsymbol{w}_z, \quad (7)$$

where $\tilde{\boldsymbol{A}}_z = \boldsymbol{A}_z + \alpha\boldsymbol{I}$.

D. ESTIMATION OF EXPECTED VALUES

The optimization objective in (6) includes the expectation over the audio signal in the definition of $\tilde{\boldsymbol{A}}_z$, and the SINR constraint includes the expectation over both the audio signal and the RIRs, as well as the expected external noise power. Only in exceptional cases would any of these expected values be known a priori, so in most cases they must be estimated from the available data.

The audio signals are assumed to be locally stationary, such that the covariances are approximately constant over a fixed segment length. With a segment of N samples starting at time index n_0, the audio signal covariances can be estimated as

$$\tilde{\boldsymbol{A}}_z \approx \sum_{n=n_0}^{n_0+N-1} \mathrm{blkdiag}\{\boldsymbol{x}_z(n)\boldsymbol{x}_z^\top(n)\}_{l=1}^L. \quad (8)$$

The spatial covariance matrices as in (5) are estimated in an analogous way, by filtering the audio signals through the estimated RIRs, and then calculating the sample covariance matrix as in (8).

The RIRs are slowly time-varying even if loudspeakers and zones are stationary, as temperature changes alone can cause a significant change [24]. The time-variations of the RIRs are likely slower than for the audio signals. The RIRs are therefore modelled as constant, and assumed to be measured to sufficient precision.

To estimate the power of the external noise, at least one physical microphone per zone is required. If the noise power is stationary, it could be directly estimated from the microphone signals before transmission of the audio signals is begun. In the general case, the transmitted audio signals at the microphones can be calculated according to (1) and subtracted from the recorded microphone signals. In the case of perfectly estimated RIRs, the true noise signal is obtained, otherwise a combination of the external noise and artifacts from RIR estimation errors is obtained.

E. ACOUSTIC CONTRAST CONTROL

For comparison, the popular sound zone control method ACC can be obtained from the optimization problem

$$\begin{aligned} \underset{\boldsymbol{w}_z}{\text{maximize}} \quad & \boldsymbol{w}_z^\top \boldsymbol{R}_{zz}\boldsymbol{w}_z \\ \text{subject to} \quad & \boldsymbol{w}_z^\top \boldsymbol{R}_{-z}\boldsymbol{w}_z = \gamma_z, \end{aligned} \quad (9)$$

where $\boldsymbol{R}_{-z} = \alpha\boldsymbol{I} + \frac{1}{Z-1}\sum_{\substack{z\in\mathcal{Z}\\ i\neq z}} \boldsymbol{R}_{iz}$. The parameter $\alpha \in \mathbb{R}_{\geq 0}$ controls the regularization. The optimal control filter $\boldsymbol{w}_z$ is the principal generalized eigenvector of the pencil $(\boldsymbol{R}_{zz}, \boldsymbol{R}_{-z})$. The optimal control filter associated with each zone is computed independently, and then the obtained solutions are superimposed.

III. METHOD 1: SEMI-DEFINITE RELAXATION

The problem in (6) can be written in the form of a separable quadratically constrained quadratic program (QCQP). QCQPs are non-convex and NP-hard in general [25], however, some can be solved by SDR. Introducing the matrix $\boldsymbol{W}_z = \boldsymbol{w}_z\boldsymbol{w}_z^\top$ and using the fact that $\boldsymbol{w}_i^\top\boldsymbol{R}_{zi}\boldsymbol{w}_i = \mathrm{tr}(\boldsymbol{W}_i\boldsymbol{R}_{zi})$, the

SDR of (6) leads to the optimization problem

$$\begin{aligned} \underset{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_Z}{\text{minimize}} \quad & \sum_{z\in\mathcal{Z}} \text{tr}(\boldsymbol{W}_z\tilde{\boldsymbol{A}}_z) \\ \text{subject to} \quad & \text{tr}(\boldsymbol{W}_z\boldsymbol{R}_{zz}) - \gamma_z \sum_{\substack{i\in\mathcal{Z}\\ i\neq z}} \text{tr}(\boldsymbol{W}_i\boldsymbol{R}_{zi}) \geq \sigma_z^2\gamma_z, \\ & \boldsymbol{W}_z \geq 0. \end{aligned} \tag{10}$$

The semi-definite program (10) is equivalent to (6) if the rank constraint $\text{rank}(\boldsymbol{W}_z) = 1$ is added for all z. However, the rank constraint is impractical as it is not convex. Without the rank constraint, the problem in (10) is a relaxation of the original problem (6), and therefore provides a lower bound for the optimal value. If a rank-1 solution is optimal for the relaxed problem, it is feasible for the original problem, and since the relaxed problem provides a lower bound, the solution is optimal for the original problem. The optimization problem (10) can also be identified as the Lagrangian bidual of (6).

It is shown in [25]–[27] that when SDR is used to solve a real-valued separable QCQP with K constraints and Z optimization variables, there is at least one optimal solution satisfying

$$\sum_{z=1}^{Z} \frac{\text{rank}(\boldsymbol{W}_z)(\text{rank}(\boldsymbol{W}_z)+1)}{2} \leq K. \tag{11}$$

As long as no optimal matrix is the zero matrix (a rank-0 matrix), a rank-1 solution is then guaranteed to exist if $K \leq Z+1$. The optimization problem (6) corresponds to a QCQP satisfying $K = Z$, and the zero matrix cannot be feasible for any $\gamma_z > 0$, so the problem (10) is equivalent to the original problem (6). When the optimal solution has been found, the control filter can be extracted as $\boldsymbol{w}_z = \sqrt{\lambda_z}\boldsymbol{u}_z$ where λ_z and $\boldsymbol{u}_z$ are the principal eigenvalue and eigenvector of the matrix $\boldsymbol{W}_z$ respectively.

IV. METHOD 2: VIRTUAL RECEIVE OPTIMIZATION PROBLEM

It can be shown that the optimal transmit control filters of (6) are the same up to a scaling as the optimal control filters for the receive optimization problem

$$\begin{aligned} \underset{\boldsymbol{q},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}} \quad & \boldsymbol{\sigma}^\top\boldsymbol{q} \\ \text{subject to} \quad & \text{SINR}_z^{\text{rec}} \geq \gamma_z, \\ & \boldsymbol{q} \geq 0, \\ & \|\bar{\boldsymbol{w}}_z\|_2 = 1, \end{aligned} \tag{12}$$

where the receive SINR is defined as

$$\text{SINR}_z^{\text{rec}} = \frac{q_z\bar{\boldsymbol{w}}_z^\top\boldsymbol{R}_{zz}\bar{\boldsymbol{w}}_z}{\bar{\boldsymbol{w}}_z^\top\left(\tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z}\\ i\neq z}} q_i\boldsymbol{R}_{iz}\right)\bar{\boldsymbol{w}}_z}. \tag{13}$$

The power allocations associated with the different zones are collected in a vector $\boldsymbol{q} = \text{col}\{q_z\}_{z\in\mathcal{Z}}$, as are the noise powers $\boldsymbol{\sigma} = \text{col}\{\sigma_z^2\}_{z\in\mathcal{Z}}$. Two alternative derivations of the receive optimization problem (12) can be found in the appendices A and B, as well as proof that (12) and (6) are solved by the same control filters and with the same objective value.

A. FIXED POINT ITERATION

The problem in (12) can be solved using a fixed point iteration algorithm as in [15]. With an initial value of $\boldsymbol{q} = \boldsymbol{0}$, the following two steps are repeated until the control filters and power allocation have converged to a fixed value.

1) Calculate control filters $\bar{\boldsymbol{w}}_z$ according to section IV-B.
2) Calculate power allocation $\boldsymbol{q}$ according to section IV-C if (17) is satisfied, otherwise according to section IV-D.

Once converged, the transmit power allocation for the original problem (6) can be calculated according to section IV-E.

If the chosen SINR constraints can not be supported by any set of control filters, the problem will stay infeasible, meaning (17) will remain false. A maximum number of iterations can be set, after which the problem is considered infeasible. Because the algorithm generally only requires a few iteration to converge, the maximum number of iterations can be set to a moderately low number.

B. OPTIMAL CONTROL FILTER FOR GIVEN POWER ALLOCATION

For a given power allocation $\boldsymbol{q}$, the optimal control filter should maximize the SINR. Only the control filter $\bar{\boldsymbol{w}}_z$ affects $\text{SINR}_z^{\text{rec}}$ in (13), so the control filters can be found independently. Maximization of (13) is the maximization of a generalized Rayleigh quotient, which is solved by the principal generalized eigenvector of the pencil

$$\left(\boldsymbol{R}_{zz},\ \tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z}\\ i\neq z}} q_i\boldsymbol{R}_{iz}\right). \tag{14}$$

The principal generalized eigenvector is normalized to ensure $\|\bar{\boldsymbol{w}}_z\|_2 = 1$.

C. RECEIVE POWER ALLOCATION FOR FEASIBLE CONTROL FILTER

The optimal power allocation of (12) for a given feasible control filter will satisfy the SINR constraints with equality [15]. The optimal power allocation can therefore be directly solved for. Introducing the diagonal matrix $\boldsymbol{D} = \text{diag}\{\frac{\gamma_z}{\bar{\boldsymbol{w}}_z^\top\boldsymbol{R}_{zz}\bar{\boldsymbol{w}}_z}\}_{z\in\mathcal{Z}}$, the transmit power vector $\boldsymbol{s} = \text{col}\{\bar{\boldsymbol{w}}_z^\top\tilde{\boldsymbol{A}}_z\bar{\boldsymbol{w}}_z\}_{z\in\mathcal{Z}}$ and the interference matrix

$$\boldsymbol{\Psi}_{zi} = \begin{cases} \bar{\boldsymbol{w}}_i^\top\boldsymbol{R}_{zi}\bar{\boldsymbol{w}}_i & \text{if } z \neq i \\ 0 & \text{if } z = i \end{cases}, \tag{15}$$

the optimal power allocation for (12) is given by

$$\boldsymbol{q} = (\boldsymbol{I} - \boldsymbol{D}\boldsymbol{\Psi}^\top)^{-1}\boldsymbol{D}\boldsymbol{s}. \tag{16}$$

The power allocation problem is feasible if there is a power allocation $\boldsymbol{q} > 0$ that supports the requirement

$\text{SINR}_z^{\text{rec}} \geq \gamma_z$ for all z. A sufficient and necessary criterion for feasibility is

$$\rho(\boldsymbol{D}\boldsymbol{\Psi}^\top) < 1, \tag{17}$$

where $\rho(\cdot)$ is the spectral radius [14]. The transmit power allocation is feasible if and only if the receive power allocation is feasible, because $\rho(\boldsymbol{D}\boldsymbol{\Psi}) = \rho(\boldsymbol{D}\boldsymbol{\Psi}^\top)$ [14].

D. RECEIVE POWER ALLOCATION FOR INFEASIBLE CONTROL FILTER

If the given control filter is not feasible according to (17), then the power allocation scheme in (16) cannot be used. The following power allocation scheme can then be used instead, which if iterated according to the algorithm in section IV-A will find a feasible control filter if it exists.

While the control filter is infeasible, the power constrained SINR maximization problem will be solved instead, which can be stated as

$$\begin{aligned} C = \underset{\boldsymbol{q}}{\text{maximize}} \quad & \underset{z\in\mathcal{Z}}{\text{minimize}} \quad \frac{\text{SINR}_z^{\text{rec}}}{\gamma_z} \\ \text{subject to} \quad & \|\boldsymbol{q}\|_1 \leq P_{\max} \end{aligned} \tag{18}$$

for a given fixed control filter. The optimum of (18) will satisfy $\frac{\text{SINR}_z^{\text{rec}}}{\gamma_z} = \frac{\text{SINR}_i^{\text{rec}}}{\gamma_i}$ for all i, z [15]. The optimum objective can thereby be written in matrix form as

$$\boldsymbol{q}\frac{1}{C} = \boldsymbol{D}\boldsymbol{\Psi}^\top\boldsymbol{q} + \boldsymbol{D}\boldsymbol{s}. \tag{19}$$

The power constraint will also be satisfied exactly at the optimum, meaning that $\mathbf{1}^\top\boldsymbol{q} = P_{\max}$. Equation (19) can be multiplied by $\mathbf{1}^\top$ from the left, giving

$$\frac{1}{C} = \frac{1}{P_{\max}}\mathbf{1}^\top\boldsymbol{D}\boldsymbol{\Psi}^\top\boldsymbol{q} + \frac{1}{P_{\max}}\mathbf{1}^\top\boldsymbol{D}\boldsymbol{s}. \tag{20}$$

Together, (19) and (20) produces the eigenvalue problem

$$\boldsymbol{\Lambda}\begin{bmatrix}\boldsymbol{q}\\ 1\end{bmatrix} = \frac{1}{C}\begin{bmatrix}\boldsymbol{q}\\ 1\end{bmatrix} \tag{21}$$

in terms of the extended coupling matrix

$$\boldsymbol{\Lambda} = \begin{bmatrix}\boldsymbol{D}\boldsymbol{\Psi}^\top & \boldsymbol{D}\boldsymbol{s}\\ \frac{1}{P_{\max}}\mathbf{1}^\top\boldsymbol{D}\boldsymbol{\Psi}^\top & \frac{1}{P_{\max}}\mathbf{1}^\top\boldsymbol{D}\boldsymbol{s}\end{bmatrix}. \tag{22}$$

Only the maximum eigenvalue has a strictly positive eigenvector [15], so the power allocation is given by the first K components of the principal eigenvector of $\boldsymbol{\Lambda}$, scaled such that its last component is one.

If the ratio $\frac{\text{SINR}_z^{\text{rec}}}{\gamma_z} \geq 1$ is satisfied then so is the feasibility criterion (17). Because the ratio $\frac{\text{SINR}_z^{\text{rec}}}{\gamma_z}$ is equal for all z at the optimum, if the criterion can be satisfied for one zone, it can be satisfied for all zones. Therefore, if a feasible control filter exists, and $P_{\max}$ is set to a large enough value, the control filter will eventually be found. The dependence on $P_{\max}$ should not be restrictive in practice, as $P_{\max}$ can be set to a value much larger than a reasonable transmit power.

The global solution to the power constrained SINR maximization problem can be obtained instead of the SINR constrained power minimization problem by continuing to use this power allocation method until convergence.

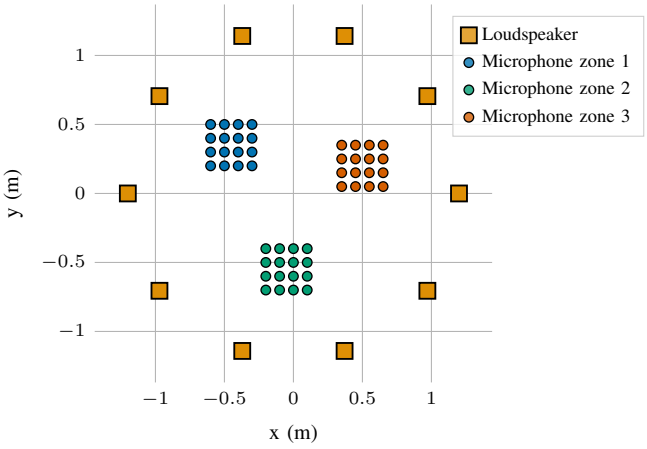

FIGURE 2. The positions of the microphones and loudspeakers.

E. TRANSMIT POWER ALLOCATION FOR OPTIMAL CONTROL FILTER

Once a set of normalized optimal control filters is obtained, the power allocation for (6) can be obtained directly as

$$\boldsymbol{p} = (\boldsymbol{I} - \boldsymbol{D}\boldsymbol{\Psi})^{-1}\boldsymbol{D}\boldsymbol{\sigma}. \tag{23}$$

This is a required last step, as the power allocation differs between the transmit and receive optimization problems. The final control filters are obtained as $\boldsymbol{w}_z = \sqrt{p_z}\bar{\boldsymbol{w}}_z$ for all z.

V. EXPERIMENTS

A. EXPERIMENT DESIGN

To characterize the performance of the proposed methods, experiments in a simulated reverberant environment have been conducted. There were three zones with 16 microphones each, sharing 10 loudspeakers, all placed on the same plane in a room of size 5×4.8×2 m. The geometry is shown in Fig. 2. The reverberation was simulated with the randomized image-source method [28], [29] with a maximum image position shift of 5 cm, using the implementation from [30].

The external noise was simulated as spatio-temporally white Gaussian noise. The same noise power was used for all microphones associated with the same zone. The noise of any two microphones was mutually independent. The noise powers for the three zones were $\boldsymbol{\sigma} = \begin{bmatrix}3 & 1 & 5\end{bmatrix}^\top$, and were assumed to be known. The SINR constraints were set to $\boldsymbol{\gamma} = \begin{bmatrix}10 & 20 & 30\end{bmatrix}^\top$.

The considered methods were the SDR based method of section III, the fixed point iteration of section IV, and ACC of section II-E. These methods will be referred to as SDR, FPI and ACC respectively. To illustrate the impact of regularization, both FPI and ACC were computed with regularization values $\alpha = 10^{-2}$ and $\alpha = 0$. In addition, two versions of FPI with less information about the audio signals were used for comparison. For one modified FPI, the optimization objective was the sum of the square norms of the control filters, equivalent to setting $\tilde{\boldsymbol{A}}_z = \boldsymbol{I}$. For the

TABLE 1. Control filter square norm, computed transmit power, and cost function using synthetic stationary audio signals.

	$\sum_{z\in\mathcal{Z}}\|\boldsymbol{w}_z\|_2^2$	$\sum_{z\in\mathcal{Z}}\boldsymbol{w}_z^\top\boldsymbol{A}_z\boldsymbol{w}_z$	$\mathcal{P}$
SDR $\alpha=10^{-2}$	27.5	38.9	**39.2**
FPI $\alpha=10^{-2}$	27.5	38.9	**39.2**
FPI $\alpha=0$	16 204.9	**36.5**	198.5
FPI $\tilde{\boldsymbol{A}}_z=\boldsymbol{I}$	**11.7**	41.6	41.7
FPI $x_z(n)=\delta(n)$	16.8	62.3	62.5
ACC $\alpha=10^{-2}$	19.5	52.0	52.2
ACC $\alpha=0$	158.4	53.0	54.6

TABLE 2. Mean transmit power and SINR using synthetic stationary audio signals over 1 min in a simulated environment.

	$\mathbb{E}\|\boldsymbol{s}(n)\|_2^2$	SINR_1	SINR_2	SINR_3
SDR $\alpha=10^{-2}$	39.5	9.3	20.3	30.9
FPI $\alpha=10^{-2}$	39.4	9.4	20.3	30.8
FPI $\alpha=0$	**32.5**	11.8	0.1	34.0
FPI $\tilde{\boldsymbol{A}}_z=\boldsymbol{I}$	41.8	9.3	19.9	30.8
FPI $x_z(n)=\delta(n)$	62.5	9.3	38.5	36.2
ACC $\alpha=10^{-2}$	53.2	9.5	20.2	31.1
ACC $\alpha=0$	54.2	9.5	20.3	31.1

second modified FPI, the audio signals were assumed to be spectrally white for both the audio signal covariance and spatial covariance matrices, consistent with the assumption $x_z(n)=\delta(n)$.

To solve the optimization problem (10) for the SDR, the splitting conic solver [31] was used, available in the Python package CVXPY [32], [33]. The maximum number of iterations for the solver was set to 10000. The FPI method was considered converged when the entry-wise mean square difference between both subsequent control filters and subsequent power allocations was less than 10^{-12}.

The control filters obtained from all methods were normalized and then scaled according to the power allocation method (23). The SINR constraints were selected to be relaxed enough such that all methods generated feasible control filters, so that the power allocation method could be used. Because ACC solves an entirely different optimization problem, it is not guaranteed to find a feasible control filter even if it exists.

B. SYNTHETIC STATIONARY SIGNALS

The first experiment considered stationary audio signals. Each audio signal was generated as white Gaussian noise filtered through a sharp bandpass filter, with bandwidths 40–450 Hz, 40–250 Hz, and 80–450 Hz, for zone 1 to 3 respectively. The signals were scaled after the bandpass filter to have mean power $\mathbb{E}[x_1(n)^2]=4$, $\mathbb{E}[x_2(n)^2]=3$, and $\mathbb{E}[x_3(n)^2]=2$. The spatial covariance matrices were calculated using 10 s of audio. The sample rate was 1000 Hz, the control filter length $I=32$, and the reverberation time $\text{RT}_{60}=0.18$ s.

The square norm of the control filters is shown in Table 1, as well as the computed transmit power, i.e. the transmit power that would be generated if the estimated covariance $\boldsymbol{A}_z$ was the true covariance. The cost function $\mathcal{P}$, which is a linear combination of the two using $\alpha=10^{-2}$ is shown as well. It can be seen that for both ACC and FPI, the control filter norm becomes significantly larger when the regularization is set to zero.

Using the calculated control filters, the sound in the room was then simulated. The resulting SINR for each zone and total transmit power is shown in Table 2. The values were obtained from the mean power estimated from 1 min of audio. The covariance estimation errors lead to SINR values slightly different from the specified level for all methods. Because FPI with $x_z(n)=\delta(n)$ assumes the audio is spectrally white, the resulting SINR is not close to the desired values. For FPI, the regularization was critical to obtain a reasonable control filter, as Table 2 shows that the SINR values are not close to the targets without regularization. Even in other cases, such as for ACC in this experiment, the regularization can also lower the transmit power in practice. ACC required more transmit power to achieve similar SINRs compared to all proposed methods except FPI with $x_z(n)=\delta(n)$.

1) DUALITY

In Table 1 it can be seen that SDR and FPI obtained the same objective value. However, it is worth noting that the optimal control filters are ambiguous with regards to sign. If $\boldsymbol{w}_z$ is optimal, so is $-\boldsymbol{w}_z$. Therefore, the solutions from SDR and FPI are identical up to a sign difference, which happens independently for each zone z, i.e. the two solutions can be equal for some z while being negative inverses for others. This does not affect the norm or the computed transmit power, as Table 1 confirms. However, in using the filters in practice, there can be some differences in performance depending on the signs, which can be seen in Table 2 if the values for SDR and FPI are compared. This performance difference is entirely accounted for if the sign differences are corrected.

2) COMPUTATIONAL COST

The primary computational cost of FPI is the generalized eigenvalue decomposition (GEVD), which is performed once per iteration. The GEVD is of the same size as for ACC, which means that one iteration of the FPI algorithm has about the same computational cost as ACC. For the experiment presented here, it took 6 to 7 iterations for FPI to converge. The relative computational cost for SDR is variable, but generally is several orders of magnitude higher. For the experiment presented here, SDR took 1752 s, FPI took between 0.233–0.303 s, and ACC took 0.046 s. These numbers only include the time required to calculate the control filter and power allocation, and not the covariance matrices which are similarly calculated for all methods.

TABLE 3. Control filter square norm, computed transmit power, and cost function using music signals.

	$\sum_{z\in\mathcal{Z}}\|\boldsymbol{w}_z\|_2^2$	$\sum_{z\in\mathcal{Z}}\boldsymbol{w}_z^\top\boldsymbol{A}_z\boldsymbol{w}_z$	$\mathcal{P}$
FPI $\alpha=10^{-2}$	13.1	**26.1**	**26.3**
FPI $\alpha=0$	301.0	**26.1**	29.1
FPI $\tilde{\boldsymbol{A}}_z=\boldsymbol{I}$	**5.8**	52.8	52.8
FPI $x_z(n)=\delta(n)$	47.7	163.8	164.3
ACC $\alpha=10^{-2}$	38.3	358.9	359.3
ACC $\alpha=0$	346.3	513.2	516.7

TABLE 4. Mean transmit power and SINR using music signals over 1 min in a simulated environment.

	$\mathbb{E}\|\boldsymbol{s}(n)\|_2^2$	SINR_1	SINR_2	SINR_3
FPI $\alpha=10^{-2}$	**26.5**	10.1	20.0	30.8
FPI $\alpha=0$	**26.5**	10.1	20.0	30.8
FPI $\tilde{\boldsymbol{A}}_z=\boldsymbol{I}$	53.1	10.2	19.8	30.9
FPI $x_z(n)=\delta(n)$	165.3	20.2	3.0	88.6
ACC $\alpha=10^{-2}$	375.2	10.1	20.1	31.1
ACC $\alpha=0$	532.8	10.1	20.4	31.1

C. MUSIC SIGNALS

More realistic signals were also considered, specifically music from the dataset [34]. The sample rate was increased to 8000 Hz, the control filter length to $I=256$, and the reverberation time to $\text{RT}_{60}=0.55$ s. Other parameters are identical to the previous experiment. SDR was not included because of its high computational cost and because it gives identical results to FPI. The covariance matrices were computed from 1 min of audio. The computed transmit power and the control filter square norm can be seen in Table 3. The same 1 min of audio used to estimate the covariances was transmitted, for which the resulting transmit power and SINR can be seen in Table 4.

It is clear from Table 4 that when the audio signal characteristics are not taken into account at all, such as for FPI with $x_z(n)=\delta(n)$, the resulting SINR is not close to the desired values. The music signals are further from spectrally white compared to the synthetic signals, so the method performs even worse here. Comparing FPI with $\tilde{\boldsymbol{A}}_z=\boldsymbol{I}$ to FPI with $\alpha=10^{-2}$, by including the audio signal characteristics in the power minimization, the resulting transmit power is approximately halved. Finally, ACC produces more than 10 times the transmit power compared to FPI at similar SINR.

VI. CONCLUSION

A formulation of sound zone control without superposition was proposed, where the control filters are jointly optimized for low transmit power, while maintaining a sufficiently high SINR. The optimization problem was shown to be solved optimally by semidefinite relaxation, but the computational cost was high. An alternate method was proposed, based on a virtual receive optimization problem derived from the original transmit optimization problem. The two problems were shown to have the same optimal control filters, while the former can be solved more efficiently. Through simulations it was shown that the proposed methods are effective, leading to lower transmit power at similar SINR compared to the popular ACC method. The importance of considering the spectral characteristics of the audio signals was demonstrated, as well as including appropriate regularization.

APPENDIX A
EQUIVALENCE OF TRANSMIT AND VIRTUAL RECEIVE OPTIMIZATION PROBLEMS

In this section the virtual receive optimization problem (12) is derived from the original transmit optimization problem (6). It will demonstrate that the two problems are solved by the same control filters. The derivations follows the approach in [35].

The optimal solution of (6) will satisfy the SINR constraints with equality, which means that the problem can be equivalently expressed as

$$\begin{aligned}\underset{\boldsymbol{p},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}}\quad & \boldsymbol{p}^\top\boldsymbol{s}\\ \text{subject to}\quad & (\boldsymbol{I}-\boldsymbol{D\Psi})\boldsymbol{p}=\boldsymbol{D\sigma},\\ & \boldsymbol{p}>0,\\ & \|\bar{\boldsymbol{w}}_z\|=1.\end{aligned}\tag{24}$$

The positivity requirement for $\boldsymbol{p}$ can be replaced by a constraint on the spectral radius of the matrix $\boldsymbol{D\Psi}$. An important result from Perron-Frobenius theory states that an irreducible non-negative square matrix $\boldsymbol{Q}$ has a strictly positive spectral radius $\rho(\boldsymbol{Q})>0$ which is an algebraically simple eigenvalue, and there is a unique vector that satisfies $\boldsymbol{x}>0$, $\boldsymbol{x}^\top\boldsymbol{1}=1$, and $\boldsymbol{Qx}=\rho(\boldsymbol{Q})\boldsymbol{x}$ [36, Th. 8.4.4]. If the constraints are fulfilled in (24), so that $(\boldsymbol{I}-\boldsymbol{D\Psi})\boldsymbol{p}=\boldsymbol{D\sigma}$, then looking at each row individually, all three of

$$p_z>0,\qquad \frac{(\boldsymbol{D\Psi p})_z}{p_z}<1,\qquad \frac{(\boldsymbol{D\sigma})_z}{p_z}>0\tag{25}$$

must be true. For a non-negative and irreducible matrix such as $\boldsymbol{D\Psi}$, the spectral radius is given by [37, Ch. 2, eq. (2.11)]

$$\rho(\boldsymbol{D\Psi})=\min_{x_z\geq 0}\max_{z\in\mathcal{Z}}\frac{(\boldsymbol{D\Psi x})_z}{x_z}.\tag{26}$$

By combining (25) and (26) the following can be asserted

$$\rho(\boldsymbol{D\Psi})\leq\max_{z\in\mathcal{Z}}\frac{(\boldsymbol{D\Psi p})_z}{p_z}<1.\tag{27}$$

Any feasible point from (24) will therefore satisfy $\rho(\boldsymbol{D\Psi})<1$, and the reverse is also true, that any $\boldsymbol{D\Psi}$ satisfying $\rho(\boldsymbol{D\Psi})<1$ will give rise to a strictly positive $\boldsymbol{p}$. This means that (24) can be equivalently rewritten as

$$\begin{aligned}\underset{\boldsymbol{p},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}}\quad & \boldsymbol{p}^\top\boldsymbol{s}\\ \text{subject to}\quad & (\boldsymbol{I}-\boldsymbol{D\Psi})\boldsymbol{p}=\boldsymbol{D\sigma},\\ & \rho(\boldsymbol{D\Psi})<1,\\ & \|\bar{\boldsymbol{w}}_z\|=1.\end{aligned}\tag{28}$$

By solving for $\boldsymbol{p}$, the resulting optimization problem is

$$\begin{aligned} \underset{\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{D}(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{\Psi})^{-\top}\boldsymbol{s} \\ \text{subject to} \quad & \rho(\boldsymbol{D}\boldsymbol{\Psi}) < 1, \\ & \|\bar{\boldsymbol{w}}_z\| = 1. \end{aligned} \tag{29}$$

The power allocation $\boldsymbol{p}$ is eliminated from the optimization problem, while the control filters are still required to be of unit norm. At this point, it is possible to introduce an alternative power allocation, $\boldsymbol{q} = \boldsymbol{D}(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{\Psi})^{-\top}\boldsymbol{s}$. It can be equivalently written as $\boldsymbol{q} = (\boldsymbol{I}-\boldsymbol{D}\boldsymbol{\Psi}^\top)^{-1}\boldsymbol{D}\boldsymbol{s}$ [36, 0.7.4]. It can be seen that if $\boldsymbol{q} \geq 0$ and $(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{\Psi}^\top)\boldsymbol{q} = \boldsymbol{D}\boldsymbol{s}$, that implies the spectral radius constraint. The following optimization problem is therefore equivalent

$$\begin{aligned} \underset{\boldsymbol{q},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{q} \\ \text{subject to} \quad & (\boldsymbol{I}-\boldsymbol{D}\boldsymbol{\Psi}^\top)\boldsymbol{q} = \boldsymbol{D}\boldsymbol{s}, \\ & \boldsymbol{q} \geq 0, \\ & \|\bar{\boldsymbol{w}}_z\| = 1. \end{aligned} \tag{30}$$

Expanding the first constraint using the definition of the constituent variables, the same optimization problem becomes

$$\begin{aligned} \underset{\boldsymbol{q},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{q} \\ \text{subject to} \quad & \frac{q_z \bar{\boldsymbol{w}}_z^\top \boldsymbol{R}_{zz} \bar{\boldsymbol{w}}_z}{\bar{\boldsymbol{w}}_z^\top \left(\tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z} \\ i\neq z}} q_i \boldsymbol{R}_{iz}\right)\bar{\boldsymbol{w}}_z} \geq \gamma_z, \\ & \|\bar{\boldsymbol{w}}_z\| = 1, \\ & \boldsymbol{q} \geq 0, \end{aligned} \tag{31}$$

where the equality constraint is relaxed into an inequality constraint, which does not change the optimum. This is the virtual receive optimization problem in a familiar form, and it is thereby shown that the optimal control filters for (31) and (6) are the same.

APPENDIX B
DERIVATION OF VIRTUAL RECEIVE OPTIMIZATION PROBLEM FROM LAGRANGIAN DUAL

The virtual receive optimization problem (12) can also be derived from the Lagrangian dual, which will be done in this section, following the approach of [38]. Because the Lagrangian bidual (10) is shown to have the same optimal objective value as the primal (6) in section III, it means that the Lagrangian dual, and by extension the virtual receive optimization problem, will have the same optimal objective value as well.

The Lagrangian of (6), where the non-negative $\boldsymbol{q} \geq 0$ is the dual variable, is

$$\begin{aligned} &\mathcal{L}(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_Z,\boldsymbol{q}) = \sum_{z\in\mathcal{Z}} \boldsymbol{w}_z^\top \tilde{\boldsymbol{A}}_z \boldsymbol{w}_z - \\ &\sum_{z\in\mathcal{Z}} q_z \left(\frac{\boldsymbol{w}_z^\top \boldsymbol{R}_{zz} \boldsymbol{w}_z}{\gamma_z} - \sum_{\substack{i\in\mathcal{Z} \\ i\neq k}} \boldsymbol{w}_i^\top \boldsymbol{R}_{zi} \boldsymbol{w}_i - \sigma_z^2 \right) \\ &= \boldsymbol{\sigma}^\top \boldsymbol{q} + \sum_{z\in\mathcal{Z}} \boldsymbol{w}_z^\top \boldsymbol{G}_z \boldsymbol{w}_z, \end{aligned} \tag{32}$$

using the following definition

$$\boldsymbol{G}_z = \tilde{\boldsymbol{A}}_z - \frac{q_z}{\gamma_z}\boldsymbol{R}_{zz} + \sum_{\substack{i\in\mathcal{Z} \\ i\neq z}} q_i \boldsymbol{R}_{iz}. \tag{33}$$

The Lagrangian dual function is obtained by minimizing (32) over $\boldsymbol{w}_1,\ldots,\boldsymbol{w}_Z$ [39]. The value of the dual function is $\boldsymbol{\sigma}^\top\boldsymbol{q}$ if it is bounded below. The Lagrangian dual problem is the maximization of the Lagrangian dual function, which is

$$\begin{aligned} \underset{\boldsymbol{q}}{\text{maximize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{q} \\ \text{subject to} \quad & \boldsymbol{G}_z \succcurlyeq 0, \\ & \boldsymbol{q} \geq 0, \end{aligned} \tag{34}$$

where the constraint ensures that the dual problem is feasible, that $\min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_Z} \mathcal{L}(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_Z,\boldsymbol{q}) > -\infty$.

From the definition of semi-positive definiteness, the constraint in (34) is equivalent to $\bar{\boldsymbol{w}}_z^\top \boldsymbol{G}_z \bar{\boldsymbol{w}}_z \geq 0$ for all $\bar{\boldsymbol{w}}_z$. The condition being true for all vectors is equivalent to $\max_{\bar{\boldsymbol{w}}_z} -\bar{\boldsymbol{w}}_z^\top \boldsymbol{G}_z \bar{\boldsymbol{w}}_z \leq 0$. Replacing the constraint in (34), the problem becomes

$$\begin{aligned} \underset{\boldsymbol{q},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{maximize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{q} \\ \text{subject to} \quad & \frac{q_z \bar{\boldsymbol{w}}_z^\top \boldsymbol{R}_{zz} \bar{\boldsymbol{w}}_z}{\bar{\boldsymbol{w}}_z^\top \left(\tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z} \\ i\neq z}} q_i \boldsymbol{R}_{iz}\right)\bar{\boldsymbol{w}}_z} \leq \gamma_z, \\ & \boldsymbol{q} \geq 0. \end{aligned} \tag{35}$$

The constraints will be satisfied with equality at the optimum, so for a given control filter $\bar{\boldsymbol{w}}_z$ the power allocation $\boldsymbol{q}$ has a unique solution determined by the constraint. This means that the solution of (35) is given by the fixed point of

$$\underset{\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{maximize}} \frac{q_z \bar{\boldsymbol{w}}_z^\top \boldsymbol{R}_{zz} \bar{\boldsymbol{w}}_z}{\bar{\boldsymbol{w}}_z^\top \left(\tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z} \\ i\neq z}} q_i \boldsymbol{R}_{iz}\right)\bar{\boldsymbol{w}}_z} = \gamma_z. \tag{36}$$

The same argument can be applied to the following problem

$$\begin{aligned} \underset{\boldsymbol{q},\bar{\boldsymbol{w}}_1,\ldots,\bar{\boldsymbol{w}}_Z}{\text{minimize}} \quad & \boldsymbol{\sigma}^\top \boldsymbol{q} \\ \text{subject to} \quad & \frac{q_z \bar{\boldsymbol{w}}_z^\top \boldsymbol{R}_{zz} \bar{\boldsymbol{w}}_z}{\bar{\boldsymbol{w}}_z^\top \left(\tilde{\boldsymbol{A}}_z + \sum_{\substack{i\in\mathcal{Z} \\ i\neq z}} q_i \boldsymbol{R}_{iz}\right)\bar{\boldsymbol{w}}_z} \geq \gamma_z, \\ & \boldsymbol{q} \geq 0, \end{aligned} \tag{37}$$

so clearly (35) and (37) have the same optimal solution. Note that scaling $\bar{\boldsymbol{w}}_z$ does not change the left-hand side of the constraint in (37). It is therefore possible to introduce the constraint $\|\bar{\boldsymbol{w}}_z\|_2 = 1$, meaning that the derived problem is exactly equal to (12).

REFERENCES

[1] W. F. Druyvesteyn and J. Garas, "Personal sound," *JAES*, vol. 45, no. 9, pp. 685–701, Sept. 1997.

[2] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, Apr. 2002.

[3] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2091–2097, Apr. 2009.

[4] S. J. Elliott, J. Cheer, J. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Trans. Audio Speech Lang. Process.*, vol. 20, no. 7, pp. 2123–2133, Sept. 2012.

[5] Y. Cai, M. Wu, L. Liu, and J. Yang, "Time-domain acoustic contrast control design with response differential constraint in personal audio systems," *J. Acoust. Soc. Am.*, vol. 135, no. 6, pp. EL252–EL257, May 2014.

[6] M. Poletti, "An investigation of 2-D multizone surround sound systems," in *Proc. AES Conv.* Oct. 2008, pp. 1–9.

[7] T. Betlehem and P. D. Teal, "A constrained optimization approach for multi-zone surround sound," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2011, pp. 437–440.

[8] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4518–4525, June 2012.

[9] T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, "A unified approach to generating sound zones using variable span linear filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2018, pp. 491–495.

[10] T. Lee, L. Shi, J. K. Nielsen, and M. G. Christensen, "Fast generation of sound zones using variable span trade-off filters in the DFT-domain," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 363–378, Dec. 2020.

[11] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[12] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. Annu. Allerton Conf. Comm., Control, Comput.*, Sept. 1999, pp. 987–996.

[13] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Vehicular Technol. Conf.*, May 1999, vol. 1, pp. 851–856.

[14] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE Vehicular Technol. Conf.*, Sept. 2002, vol. 1, pp. 87–91.

[15] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[16] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[17] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, July 2014.

[18] G. Piñero, C. Botella, M. De Diego, M. Ferrer, and A. González, "On the feasibility of personal audio systems over a network of distributed loudspeakers," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 2729–2733.

[19] Y. Cai, M. Wu, and J. Yang, "Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2013, pp. 341–345.

[20] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "Time domain optimization of filters used in a loudspeaker array for personal audio," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 11, pp. 1869–1878, Nov. 2015.

[21] N. Ueno, S. Koyama, and H. Saruwatari, "Directionally weighted wave field estimation exploiting prior information on source direction," *IEEE Trans. Signal Process.*, vol. 69, pp. 2383–2395, Apr. 2021.

[22] J. Brunnström, S. Koyama, and M. Moonen, "Variable span trade-off filter for sound zone control with kernel interpolation weighting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2022, pp. 1071–1075.

[23] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, June 2006.

[24] M. Olsen and M. B. Møller, "Sound zones: On the effect of ambient temperature variations in feed-forward systems," in *Proc. AES Conv.* May 2017, pp. 1009–1018.

[25] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[26] G. Pataki, "On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues," *Mathematics of OR*, vol. 23, no. 2, pp. 339–358, May 1998.

[27] A. I. Barvinok, "Problems of distance geometry and convex properties of quadratic maps," *Discrete Comput. Geom.*, vol. 13, no. 2, pp. 189–202, Mar. 1995.

[28] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[29] E. De Sena, N. Antonello, M. Moonen, and T. van Waterschoot, "On the modeling of rectangular geometries in room acoustic simulations," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 4, pp. 774–786, Apr. 2015.

[30] R. Scheibler, E. Bezzam, and I. Dokmanić, "Pyroomacoustics: A Python package for audio room simulations and array processing algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2018, pp. 351–355.

[31] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *J. Optim. Theory Appl.*, vol. 169, no. 3, pp. 1042–1068, June 2016.

[32] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *J. Mach. Learn. Res.*, vol. 17, pp. 1–5, 2016.

[33] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *J. Control Decis.*, vol. 5, no. 1, pp. 42–60, Jan. 2018.

[34] Z. Rafii, A. Liutkus, S. Fabian-Robert, S. I. Mimilakis, and R. Bittner, "MUSDB18-HQ - an uncompressed version of MUSDB18," https://doi.org/10.5281/zenodo.3338373, Dec. 2019.

[35] M. Bengtsson and B. Ottersten, "Optimum and suboptimum transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, Aug. 2001.

[36] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Oct. 2012.

[37] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Society for Industrial and Applied Mathematics, 1994.

[38] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[39] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.