Providing Spatial Control in Personal Sound Zones Using Graph Signal Processing

Vicent Molés-Cases, Gema Piñero, Alberto Gonzalez, and Maria de Diego

Institute of Telecommunications and Multimedia Applications (iTEAM)

Universitat Politècnica de València, Spain

Abstract—Personal audio systems aim to create listening (or bright) and quiet (or dark) zones in a room using an array of loudspeakers. For this purpose, many algorithms have been presented in the literature, being Weighted Pressure Matching (wPM) one of the most versatile. The main strength of wPM is that it can render a target soundfield in the listening zone while having control over the mean acoustic potential energy in the quiet zone. In this paper, we propose a variation of wPM such that it can provide control not only over the mean energy, but also over the spatial energy differences, obtaining a more uniform soundfield in the dark zone. The new algorithm is called wPM with Total Variation (wPM-TV), where TV is a tool used in the field of Graph Signal Processing (GSP). Firstly, we propose a graph representation of the control microphones of the dark zone and secondly, we use the wPM-TV algorithm to provide spatial control over that zone. Simulations show the good performance of the proposed algorithm and its versatility to obtain a more uniform distribution of the acoustic potential energy in the dark zone at the cost of slightly increasing the mean square reproduction error in the bright zone.

Index Terms—Sound zones, pressure matching, spatial control, graph total variation.

I. INTRODUCTION

Personal audio systems (PAS) aim to create sound zones [1] with minimum leakage to listeners in other regions by using an array of loudspeakers [2]. In the literature, the term bright zone denotes the spatial region where an audio signal will be rendered, and dark zone the region where the sound level will be minimized [3]. Several techniques can be applied to create sound zones, such as beamforming [4], [5], soundfield synthesis [6], [7], energy cancellation approaches [3], [8], or hybrid approaches [9] [10].

Weighted Pressure Matching (wPM) is an hybrid technique presented in [9]. It is considered as an hybrid approach because it offers a compromise between the performance obtained with two different techniques, namely, Pressure Matching (PM) [7] and Acoustic Contrast Control (ACC) [11]. On the one hand, PM is a soundfield synthesis approach that aims to render a target soundfield in the bright zone, but does not offer control over the acoustic potential energy in the dark zone [7]. On the other hand, ACC is an energy cancellation approach whose goal is to minimize the mean acoustic potential energy in the dark zone, but it can not synthesize a desired target soundfield

in the bright zone [11]. wPM offers the possibility to both render a target soundfield in the bright zone while having control over the mean acoustic potential energy in the dark zone. To do so, the authors in [9] propose a novel cost function in which a tuning parameter is used to balance the components of the cost functions for PM and ACC. With wPM, the optimal source strength vector is computed for a set of frequencies, and then, a time-domain filter for each loudspeaker is obtained using the inverse discrete fourier transform.

To the best of our knowledge, the spatial uniformity of the acoustic potential energy in the dark zone is an issue that has not been considered in the literature. Usually, PAS algorithms try to minimize the mean acoustic potential energy in the control points of the dark zone. However, this strategy does not necessarily lead to scenarios with low energy in all the control points. In some cases, when the soundfield is not uniform, the energy in some control points of the dark zone can be high even if the mean energy is low. This effect can be disturbing for a user located in the dark zone because of the sudden energy changes, which can also affect the privacy in the dark zone. Authors in [12] propose to use a masker signal on the residual interference in the dark zone in order to reduce the intelligibility, and consequently, improve the privacy between zones. However, if the energy distribution in the dark zone is not uniform, the masking signal is only able to reduce the intelligibility in the spatial points where the energy is close to the mean or lower. Thus, not only a low mean acoustic potential energy is desired for the dark zone, but also a uniform energy spatial distribution.

In this paper we propose a novel algorithm that takes into account the spatial uniformity of the acoustic potential energy in the dark zone, based on Graph Signal Processing (GSP). GSP is a framework that aims to process data which resides in irregular domains, that are represented by a graph [13]. This framework has been extensively used in areas such as image processing [14], social networks data processing [15], and many more [16]. However, GSP has not been previously applied to the problem of providing sound zones. In this work, we define a graph to characterize the relations among the acoustic potential energy in different spatial points. The Total Variation of the graph [17] is used to improve the spatial uniformity of the soundfield in the dark zone.

The outline of the paper is as follows. Section 2 presents the model used and Section 3 reviews the wPM algorithm. Section 4 presents a novel algorithm based on GSP that takes

into account the uniformity of the energy in the dark zone. Section 5 shows simulations results that prove the efficacy of the proposed algorithm. Finally, Section 6 summarizes the main conclusions.

Notation: Throughout this paper matrices and vectors are represented by upper and lower case boldface letters respectively, $(\cdot)^H$ stands for conjugate transpose, $\|\cdot\|$ for vector 2-norm, $\mathbb{C}$ for the set of complex numbers, and $\mathbb{R}^+$ for the set of non-negative real numbers.

II. MODEL AND DESCRIPTION

Consider a system formed by an array of L loudspeakers and two spatial regions delimited by M_b (bright zone) and M_d (dark zone) control points, respectively. The main goal of such system is to render an audio signal in the bright zone while producing low interference in the dark zone. Next, for the considered system, we derive the expressions for the acoustic pressure in the control points at the single frequency ω, however, index ω has been omitted for the sake of simplicity.

Let us define $h_{qm,l} \in \mathbb{C}$ as the Acoustic Transfer Function (ATF) between the l-th loudspeaker and the m-th control point in zone q, $q = \{b, d\}$. Then, $\mathbf{h}_{qm} = [h_{qm,1}, \ldots, h_{qm,L}]$ is the ATF between the L loudspeakers and the m-th control point in zone q. Let us define the ATF for zone q as

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{h}_{q1} \\ \vdots \\ \mathbf{h}_{qM_q} \end{bmatrix} = \begin{bmatrix} h_{q1,1} & \ldots & h_{q1,L} \\ \vdots & \vdots & \vdots \\ h_{qM_q,1} & \ldots & h_{qM_q,L} \end{bmatrix} . \quad (1)$$

Then, we can define the acosutic pressure in the m-th control point of zone q as [3]

$$x_{qm} = \sum_{l=1}^{L} h_{qm,l} g_l = \mathbf{h}_{qm} \mathbf{g} , \quad (2)$$

where $g_l \in \mathbb{C}$ is the source strength for loudspeaker l at single frequency ω, and

$$\mathbf{g} = \begin{bmatrix} g_1 & \cdots & g_L \end{bmatrix}^T . \quad (3)$$

Using (1)-(3) we can define vector $\mathbf{x}_q$ containing the acoustic pressure at the M_q control points as

$$\mathbf{x}_q = \begin{bmatrix} x_{q1} & \cdots & x_{qM_q} \end{bmatrix}^T = \mathbf{H}_q \mathbf{g} . \quad (4)$$

Finally, we show the expressions of two metrics that will be used in the following sections. First, the mean acoustic potential energy for zone q is defined as [3]

$$E_q = \frac{1}{M_q} \|\mathbf{x}_q\|^2 = \frac{1}{M_q} \mathbf{x}_q^H \mathbf{x}_q = \frac{1}{M_q} \mathbf{g}^H \mathbf{H}_q^H \mathbf{H}_q \mathbf{g} . \quad (5)$$

Second, the Mean Square Error (MSE) of the acoustic pressure in zone q with respect to a target soundfield is defined as [9]

$$\varepsilon_q = \frac{1}{M_q} \|\mathbf{H}_q \mathbf{g} - \mathbf{d}_q\|^2 = \frac{1}{M_q} (\mathbf{H}_q \mathbf{g} - \mathbf{d}_q)^H (\mathbf{H}_q \mathbf{g} - \mathbf{d}_q) , \quad (6)$$

where $\mathbf{d}_q = [d_{q1}, \ldots, d_{qM_q}]^T$, and $d_{qm} \in \mathbb{C}$ is the target soundfield at the m-th control point of zone q.

III. WEIGHTED PRESSURE MATCHING

For each control frequency, wPM is an algorithm that aims to find the source strength vector ($\mathbf{g}$) that minimizes the following cost function

$$J_{\text{wpm}}(\mathbf{g}) = \kappa \, \varepsilon_b(\mathbf{g}) + (1 - \kappa) \, E_d(\mathbf{g}) + \lambda \|\mathbf{g}\|^2 , \quad (7)$$

where $\kappa \in [0, 1]$, and λ is a regularization parameter. The tuning parameter κ offers the possibility to balance the effort used to minimize the MSE in the bright zone and the acoustic potential energy in the dark zone. Thus, values of κ close to 1 lead to solutions where more effort is used to minimize the MSE in the bright zone. Alternatively, values of κ close to 0 offer a solution where more effort is invested in minimizing the mean energy in the dark zone [9]. Regularization parameter λ is used to limit the energy of $\mathbf{g}$ [11].

To find the optimal source strength vector for wPM, we calculate the gradient of (7) with respect to $\mathbf{g}$. According to [9], the expression of the gradient is given by

$$\nabla_{\mathbf{g}} J_{\text{wpm}}(\mathbf{g}) =$$
$$2 \left[\frac{\kappa}{M_b} \left(\mathbf{H}_b^H \mathbf{H}_b \mathbf{g} - \mathbf{H}_b^H \mathbf{d}_b \right) + \frac{(1 - \kappa)}{M_d} \mathbf{H}_d^H \mathbf{H}_d \mathbf{g} + \lambda \mathbf{g} \right] , \quad (8)$$

and the optimal solution is obtained when $\nabla_{\mathbf{g}} J_{\text{wpm}}(\mathbf{g}) = \mathbf{0}$:

$$\tilde{\mathbf{g}}_{\text{wpm}} = \left(\frac{\kappa}{M_b} \mathbf{H}_b^H \mathbf{H}_b + \frac{(1 - \kappa)}{M_d} \mathbf{H}_d^H \mathbf{H}_d + \lambda \mathbf{I} \right)^{-1} \frac{\kappa}{M_b} \mathbf{H}_b^H \mathbf{d}_b . \quad (9)$$

As mentioned in Section 1, wPM can minimize the mean acoustic potential energy in the dark zone, but can not guarantee low energy in all control points. As already pointed out in Section I, a solution that aims to improve the uniformity of the energy distribution in the dark zone while minimizing its mean energy would be beneficial. Thus, next we propose an algorithm to improve the uniformity of the energy distribution in the dark zone.

IV. WEIGHTED PRESSURE MATCHING WITH TOTAL VARIATION

In this section we propose a novel algorithm called Weighted Pressure Matching with Total Variation (wPM-TV). This novel method uses concepts from the Graph Signal Processing (GSP) framework [13] to improve the spatial uniformity of the acoustic potential energy in the dark zone. The two fundamental elements of GSP are the graph and the graph signal, which is a signal present in the vertices of the graph [17]. Next, we present the graph and the graph signal that we will use in the wPM-TV algorithm.

A graph is defined as $\mathcal{G} = (\mathcal{V}, \mathbf{A})$, where $\mathcal{V}$ is the set of vertices and $\mathbf{A}$ is the adjacency matrix representing the similarity between vertices [13]. Assume that each of the vertices of the graph is associated with one control point in the dark zone. Then, the set of vertices can be defined as $\mathcal{V} = \{1, \ldots, M_d\}$, and $\mathbf{A} \in \mathbb{R}^{M_d \times M_d}$. We propose a similarity measure between vertices based on the spatial euclidean distance between the control points associated with each vertex. The main motivation is that such a similarity

measure reflects the physical relation between the soundfield at the control points, i.e., for points spatially closer the soundfield is more similar than for points placed further. Then, we can define the element in the i-th row and j-th column of $\mathbf{A}$ as

$$a_{ij} = e^{-\frac{\|\mathbf{z}_i - \mathbf{z}_j\|^2}{2\sigma^2}} \qquad (10)$$

where $\mathbf{z}_i \in \mathbb{R}^3$ contains the spatial coordinates of the i-th control point in the dark zone, and $\sigma = \frac{2}{M_d(M_d-1)} \sum_{i=1}^{M_d} \sum_{j=i+1}^{M_d} \|\mathbf{z}_i - \mathbf{z}_j\|$ is the mean euclidean distance between all control points in the dark zone. The definition in (10) assures the symmetry of matrix $\mathbf{A}$, i.e., $\mathbf{A} = \mathbf{A}^T$.

A graph signal is defined as a column vector containing the signal in each of the vertices of a graph $\mathcal{G}$. In our case, we define the graph signal $\mathbf{f}_d = [f_{d1}, \ldots, f_{dM_d}]^T$ where

$$f_{dm} = |x_{dm}| = |\mathbf{h_{dm}g}| = \sqrt{\mathbf{g}^H \mathbf{h}_{dm}^H \mathbf{h}_{dm} \mathbf{g}} , \qquad (11)$$

i.e., the graph signal in each vertex m is the magnitude of the acoustic pressure in the m-th control point of the dark zone.

The Total Variation (TV) metric [17] for the graph signal $\mathbf{f}_d$ in a graph $\mathcal{G}$ can be defined as

$$T(\mathbf{f}_d) = \mathbf{f}_d^T \mathbf{L} \mathbf{f}_d = \sum_{i=1}^{M_d} \sum_{j=1}^{M_d} l_{ij} f_{di} f_{dj} = \sum_{i=1}^{M_d} \sum_{j=1}^{M_d} a_{ij}(f_{di} - f_{dj})^2 , \qquad (12)$$

where $\mathbf{L} = \mathbf{D} - \mathbf{A}$ is the Laplacian Matrix [18], $\mathbf{D}$ is a diagonal matrix with diagonal elements $d_{ii} = \sum_{j=1}^{M_d} a_{ij}$, and l_{ij} is the element in the i-th row and j-th column of $\mathbf{L}$. The TV of a graph signal indicates its smoothness in the graph. Thus, the TV of the graph signal $\mathbf{f}_d$ is an indicator of the uniformity of the magnitude of the soundfield at the control points in the dark zone. In this sense, small values of $T(\mathbf{f}_d)$ indicate higher level of uniformity of the soundfield of the dark zone.

Consequently, we define the new cost function for the wPM-TV algorithm as the wPM cost function (7) plus a weighted term related to the TV in the dark zone as follows

$$\begin{aligned} J_{\mathrm{tv}}(\mathbf{g}) &= J_{\mathrm{wpm}}(\mathbf{g}) + \beta\, T(\mathbf{f}_d(\mathbf{g})) \\ &= \kappa\, \varepsilon_b(\mathbf{g}) + (1-\kappa)\, E_d(\mathbf{g}) + \lambda\, \|\mathbf{g}\|^2 + \beta\, T(\mathbf{f}_d(\mathbf{g})) , \end{aligned} \qquad (13)$$

where parameter $\beta \in \mathbb{R}^+$ controls the influence of the TV term in the optimal solution. Cost function (13) offers a balanced solution between MSE in the bright zone, mean acoustic potential energy in the dark zone, total variation in the dark zone, and mean source strength. The influence of each of the terms is determined by κ, β, and λ. For $\beta = 0$, the optimal solution that minimizes (13) is equal to the wPM solution. For high values of β, we achieve higher uniformity in the dark zone at the expense of increasing the MSE in the bright zone or the mean acoustic potential energy in the dark zone.

In order to find the optimal vector $\mathbf{g}$ that minimizes (13) we derive the gradient of $J_{\mathrm{tv}}(\mathbf{g})$ with respect to $\mathbf{g}$. First, we express the gradient of the the graph signal $\mathbf{f}_d(\mathbf{g})$ in vertex m (11) with respect to $\mathbf{g}$ as

$$\nabla_{\mathbf{g}} f_{dm}(\mathbf{g}) = \nabla_{\mathbf{g}} \left(\sqrt{\mathbf{g}^H \mathbf{h}_{dm}^H \mathbf{h}_{dm} \mathbf{g}} \right) = \frac{1}{f_{dm}(\mathbf{g})} \mathbf{h}_{dm}^H \mathbf{h}_{dm} \mathbf{g} , \qquad (14)$$

and the Jacobian matrix of $\mathbf{f}_d(\mathbf{g})$ with respect to $\mathbf{g}$ as

$$\mathbf{J}(\mathbf{f}_d(\mathbf{g})) = \begin{bmatrix} \nabla_{\mathbf{g}} f_{d1}(\mathbf{g}) & \cdots & \nabla_{\mathbf{g}} f_{dM_d}(\mathbf{g}) \end{bmatrix}^T . \qquad (15)$$

Thus, we can express the gradient of $T(\mathbf{f}_d(\mathbf{g}))$ as

$$\begin{aligned} \nabla_{\mathbf{g}} T(\mathbf{f}_d(\mathbf{g})) &= \nabla_{\mathbf{g}} \left(\sum_{i=1}^{M_d} \sum_{j=1}^{M_d} l_{ij} f_{di}(\mathbf{g}) f_{dj}(\mathbf{g}) \right) \\ &\overset{(a)}{=} 2 \sum_{i=1}^{M_d} \sum_{j=1}^{M_d} l_{ij} \nabla_{\mathbf{g}}(f_{di}(\mathbf{g})) f_{dj}(\mathbf{g}) \\ &\overset{(b)}{=} 2\, \mathbf{J}(\mathbf{f}_d(\mathbf{g}))^T\, \mathbf{L}\, \mathbf{f}_d(\mathbf{g}) , \end{aligned} \qquad (16)$$

where (a) follows from the symmetry of the laplacian matrix [18], and (b) follows from (15). Finally, the gradient of (13) can be derived using (16) as

$$\nabla_{\mathbf{g}} J_{\mathrm{tv}}(\mathbf{g}) = \nabla_{\mathbf{g}} J_{\mathrm{wpm}}(\mathbf{g}) + 2\beta\, \mathbf{J}(\mathbf{f}_d(\mathbf{g}))^T\, \mathbf{L}\, \mathbf{f}_d(\mathbf{g}) , \qquad (17)$$

where $\nabla_{\mathbf{g}} J_{\mathrm{wpm}}(\mathbf{g})$ was defined in (8).

From (17), it can be noted that an analytical solution $\mathbf{g}$ that fulfils the equality $\nabla_{\mathbf{g}} J_{\mathrm{tv}}(\mathbf{g}) = \mathbf{0}$ can not be found. Then, search algorithms (as steepest descent [19]) must be used to find the optimal values of $\mathbf{g}$. The update equation for searching the optimal vector $\mathbf{g}$ with steepest descent is

$$\mathbf{g}_{\mathrm{tv}}^{k+1} = \mathbf{g}_{\mathrm{tv}}^{k} - \mu \nabla_{\mathbf{g}} J_{\mathrm{tv}}(\mathbf{g}_{\mathrm{tv}}^{k}) , \qquad (18)$$

where k is the iteration index, and μ is the step size.

An important aspect that must be highlighted is that (13) is not a convex function, since $T(\mathbf{f}_d(\mathbf{g}))$ is not convex. However, if steepest descent (or another search algorithm) is used with an initial selection $\mathbf{g}_{\mathrm{tv}}^0 = \mathbf{g}_{\mathrm{wpm}}$, we can find the local minimum of (13) that is closest to the global minimum of the wPM cost function in (7). Thus, we can find a solution $\mathbf{g}$ that assures that the MSE in the bright zone and the mean acoustic potential energy in the dark zone do not change considerably with respect to the wPM solution, while improving the spatial uniformity in the dark zone.

V. SIMULATION RESULTS

A. Simulation setup

The proposed algorithm is evaluated for a simulated room using the Room Impulse Response (RIR) generator of [20]. The simulated room has dimensions $5.5 \times 5.5 \times 2.25$m, and reverberation time $T_{60} = 0.4$s. Also, a sampling frequency of 8000Hz, and a RIR length of $N = 4800$ samples has been used. A linear array of $L = 24$ elements with inter-element distance of 0.15m, and two square regions of dimensions 0.2×0.2m and 1m separation are considered (Fig. 1). In each region, a grid of 4×4 microphones with inter-element distance of 0.05m is used for control. Also, a grid of 21×21 microphones with inter-element distance of 0.01m is used for validation.
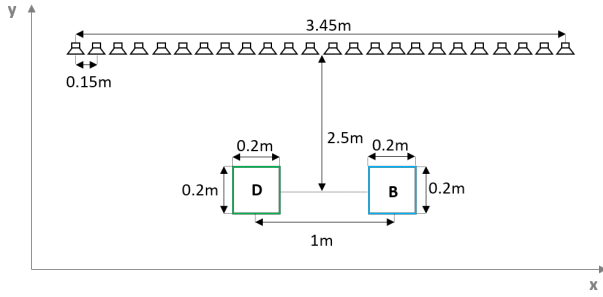
Fig. 1: Layout of the simulated Personal Audio System.

B. Simulation results and discussion

Next, we compare the performance of wPM and wPM-TV. We set $\lambda=10^{-5}$, $\kappa=0.995$, and the soundfield produced by a plane wave in the y-axis negative direction with 76dB SPL as target soundfield. For wPM-TV, β has been selected for each frequency according to the expression $\beta=\beta'T(\mathbf{f}_d(\mathbf{g}_{\mathrm{wpm}}))/E_d(\mathbf{g}_{\mathrm{wpm}})$, with $\beta'=0.1$ and $\beta'=0.8$. The motivation is that for those frequencies where the TV obtained with wPM is already low, high values of β are not beneficial.

In Fig. 2, we show a box-plot for the acoustic potential energy in the control points of the bright and dark zones as a function of frequency for wPM (Fig. 2a), wPM-TV with $\beta' = 0.1$ (Fig. 2b), and wPM-TV with $\beta' = 0.8$ (Fig. 2c). Also, we show in Fig. 3 the MSE in the bright zone as a function of frequency. We can see that wPM-TV with $\beta' = 0.8$ leads to very high uniformity of the soundfield in the dark zone, because the values of acoustic potential energy in all control points are very close to its mean. However, such high uniformity is obtained at expense of increasing the MSE, but the MSE is still lower than $3 \cdot 10^{-5}$ in all frequencies (Fig. 3). For wPM-TV with $\beta' = 0.1$, the MSE is closer to the results obtained with wPM, but the distance between the 25th and 75th percentiles for the dark zone is substantially reduced. Thus, higher uniformity than wPM is achieved at the expense of a slight increase of the MSE. Also, it is important to mention that in all frequencies wPM-TV has converged to a local minimum.

Fig. 4 shows the acoustic potential energy at the validation
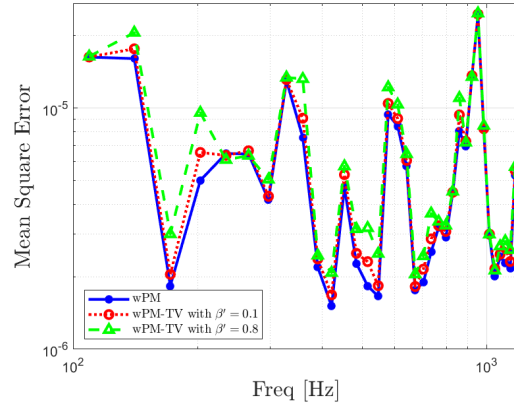
Fig. 3: Mean Square Error in the bright zone for wPM and wPM-TV with $\beta' = 0.1$ and $\beta' = 0.8$.

points (grid of 21×21 microphones) in the bright and dark zones for wPM and wPM-TV at frequencies 200Hz (top row) and 800Hz (bottom row). For 200Hz, wPM provides a mean acoustic potential energy of 64.38 dB SPL in the dark zone. However, in the validation points of the upper-left region in the dark zone the energy reaches a maximum of 72.2 dB SPL (Fig. 4a). This effect would produce that a listener would perceive high sound levels in some locations of the dark zone, even though the mean energy in the dark zone is 11.62 dB lower than in the bright zone. For 200Hz, wPM-TV with $\beta' = 0.1$ (Fig. 4b) mitigates this effect not only because the mean acoustic energy is decreased to 63.9 dB SPL, but also because the maximum acoustic energy in all validation points is 70.2 dB SPL (2 dB lower than with wPM). Consequently, wPM-TV with $\beta' = 0.1$ achieves more uniform acoustic potential energy with lower mean in the dark zone than wPM. For wPM-TV with $\beta' = 0.8$ (Fig. 4c), the uniformity of the energy in the dark zone is very high but its mean value is 65.09 dB SPL (almost 1 dB higher than with wPM). However, the maximum energy in the dark zone is 68.31 dB SPL (3.89 dB lower than with wPM). Thus, the greatest sound level that a listener would perceive with wPM-TV and $\beta' = 0.8$ is substantially lower than with wPM. Similar results are obtained for frequency 800 Hz.

(a) wPM

(b) wPM-TV with $\beta' = 0.1$
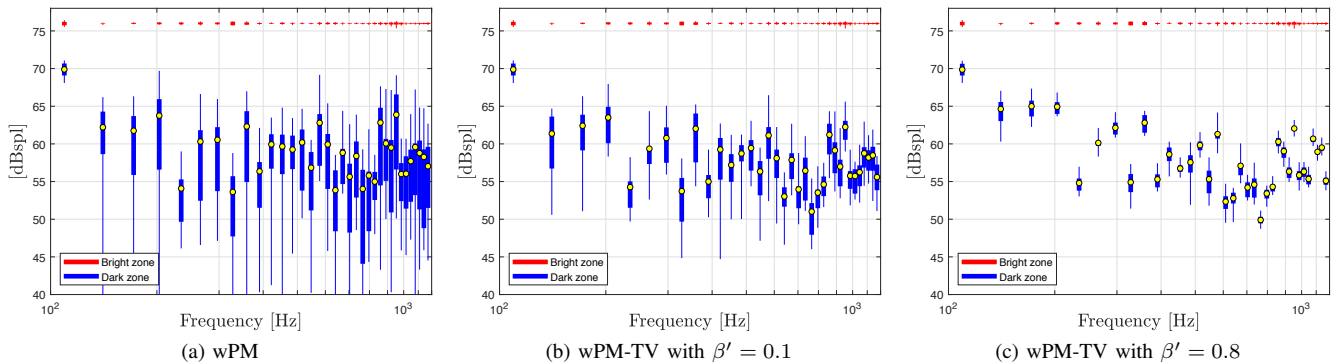
(c) wPM-TV with $\beta' = 0.8$

Fig. 2: Box-plot for the acoustic potencial energy in the control points of the bright and dark zones. The marker '○' indicates the mean, the bottom and top edges of the box indicate the 25th and 75th percentiles, and the whiskers show the lowest and highest value.

$E_d = 64.38$ dBspl $\varepsilon_b = -50.69$ dB

(a) Freq. 200 Hz - wPM

$E_d = 63.90$ dBspl $\varepsilon_b = -49.64$ dB

(b) Freq. 200 Hz - wPM-TV with $\beta' = 0.1$

$E_d = 65.09$ dBspl $\varepsilon_b = -48.16$ dB

(c) Freq. 200 Hz - wPM-TV with $\beta' = 0.8$

$E_d = 56.46$ dBspl $\varepsilon_b = -49.51$ dB

(d) Freq. 800 Hz - wPM

$E_d = 54.04$ dBspl $\varepsilon_b = -49.23$ dB

(e) Freq. 800 Hz - wPM-TV with $\beta' = 0.1$

$E_d = 53.63$ dBspl $\varepsilon_b = -48.95$ dB

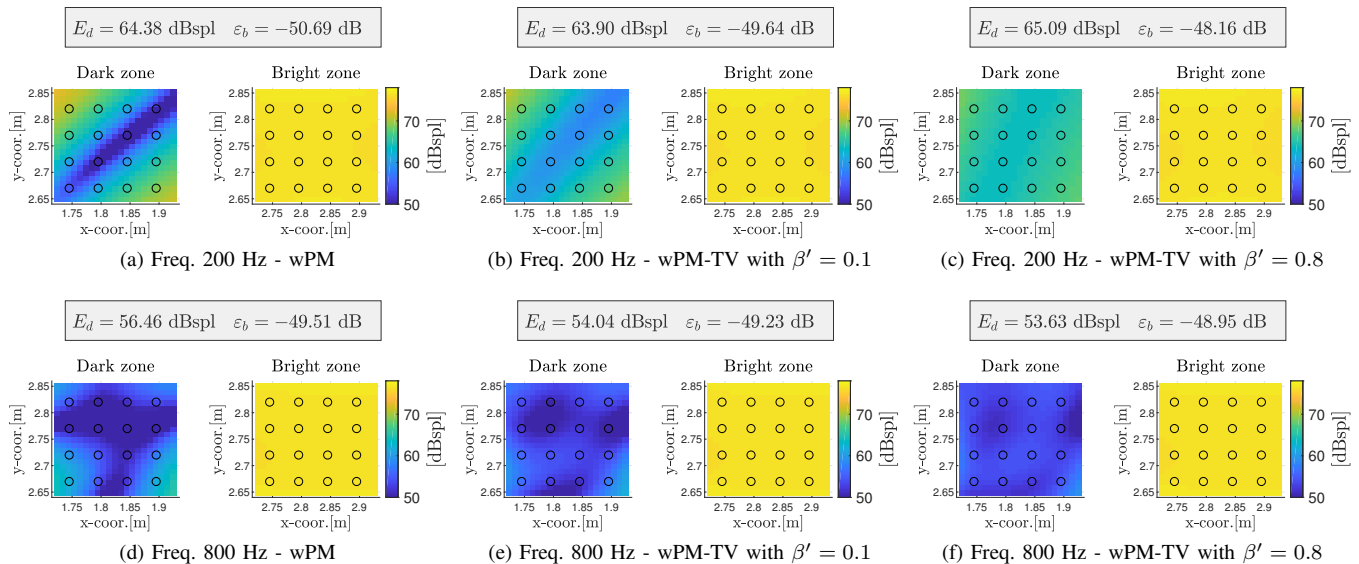(f) Freq. 800 Hz - wPM-TV with $\beta' = 0.8$

Fig. 4: Acoustic potential energy at the validation points of the bright and dark zones for wPM and wPM-TV, and for two different frequencies. In the graphs, the circles indicate the location of the control points.

From the presented results, we can conclude that wPM-TV with an appropriate selection of β' can offer higher spatial uniformity for the acoustic potential energy in the dark zone than wPM, at the expense of a very small increase in the MSE.

VI. CONCLUSIONS

In this paper we proposed the novel algorithm wPM-TV to generate Personal Sound Zones, which uses tools from the Graph Signal Processing framework to increase the spatial uniformity of the acoustic potential energy in the dark zone. It not only provides control over the mean square error in the bright zone and the mean acoustic potential energy in the dark zone (as wPM), but also over the spatial uniformity of the energy in the dark zone. We presented simulation results showing that with an appropriate selection of the tuning parameter β, wPM-TV improves substantially the spatial uniformity of the acoustic potential energy in the dark zone with respect to wPM, at the expense of slightly increasing the mean square error in the bright zone.

REFERENCES

[1] T. Betlehem, W. Zhang *et al.*, "Personal sound zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 81–91, 2015.
[2] W. F. Druyvesteyn and R. M. Aarts, "Personal sound," *The Journal of the Acoustical Society of America*, vol. 45, no. 9, pp. 685–701, 1997.
[3] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1695–1700, 2002.
[4] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.
[5] E. Mabande and W. Kellermann, "Towards Superdirective Beamforming with Loudspeaker Arrays," 2007.
[6] T. Betlehem and P. D. Teal, "A constrained optimization approach for multi-zone surround sound," *Proc. of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.
[7] M. A. Poletti, "An Investigation of 2D Multizone Surround Sound Systems," *Proc. of the 125th Audio Engineering Society Convention*, 2008.

[8] M. Shin, S. Q. Lee *et al.*, "Maximization of acoustic energy difference between two spaces," *The Journal of the Acoustical Society of America*, vol. 128, p. 121, 2010.
[9] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *The Journal of the Acoustical Society of America*, vol. 131, p. 4518, 2012.
[10] A. Canclini, D. Marković *et al.*, "A Weighted Least Squares Beam Shaping Technique for Sound Field Control," *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.
[11] S. J. Elliott, J. Cheer *et al.*, "Robustness and Regularization of Personal Audio Systems," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 7, pp. 2123–2133, 2012.
[12] D. Wallace and J. Cheer, "Combining artificial and natural background noise in personal audio systems," *Proc. of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2018.
[13] A. Sandryhaila and J. M. F. Moura, "Discrete Signal Processing on Graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.
[14] W. Hu, G. Cheung *et al.*, "Multiresolution Graph Fourier Transform for Compression of Piecewise Smooth Images," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 419–433, 2015.
[15] N. Tremblay and P. Borgnat, "Graph Wavelets for Multiscale Community Mining," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5227–5239, 2014.
[16] A. Ortega, P. Frossard *et al.*, "Graph Signal Processing: Overview, Challenges, and Applications," *Proc. of the IEEE*, vol. 106, no. 5, 2018.
[17] D. I. Shuman, S. K. Narang *et al.*, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, pp. 83–98, 2013.
[18] C. D. Godsil and G. Royle, *Algebraic graph theory*, 1st ed. Springer-Verlag New York, 2001.
[19] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*. Prentice Hall, 2000.
[20] E. A. P. Habets, "Room Impulse Response Generator," Technische Universiteit Eindhoven, Tech. Rep. October, 2010.