inter·noise 2021
1–5 August | Washington, DC

Demonstration of a unified approach to beamforming

Christof Puhle[1]
Society for the Advancement of Applied Computer Science, GFaI e.V.
Volmerstraße 3, 12489 Berlin, Germany

ABSTRACT

In this paper, we discuss a unification of several well-known frequency domain beamforming methods into one working principle. The methods under consideration include functional beamforming, asymptotic beamforming, adaptive beamforming and conventional beamforming. Common to most of these methods is the underlying eigenvalue decomposition of the cross-spectral matrix. Introducing a weighted Hölder mean in terms of these eigenvalues for every map point, each of the above methods is represented by a certain real power value p. Going from the case p=1 of conventional beamforming to lower power values results in the attenuation of side lobes and sharpening of the main lobes in the corresponding beamforming map. We demonstrate this and other effects using real-world measurements.

Keywords: Signal processing, beamforming, acoustic maps
I-INCE Classification of Subject Number: 74–76
(see http://i-ince.org/files/data/classification.pdf)

1. INTRODUCTION

For many years now, beamforming has been a widely accepted and robust acoustic source mapping technique when using microphone arrays with a limited number of sensors. In very broad terms, it can be described to work as a spatial sound receiver that smoothly either filters out or favors sound emanating from the map locations. In the frequency domain, this so-called conventional beamforming produces maps that exhibit only low spatial resolution together with a limited dynamic range, it is therefore often subject to further post-processing procedures. Incorporating conventional beamforming in a family of beamforming algorithms that depend on an exponent parameter, Dougherty [1] dramatically increased both resolution and dynamic range, and the requirements for this so-called functional beamforming are not extraordinary. This paper aims at motivating a completion of this point view and demonstrates the effects of leaving the range of conventional functional beamforming, reaching adaptive beamforming and presents even examples beyond the latter in this extended family of beamforming methods.

[1]puhle@gfai.de

2. THEORY

Suppose there are M microphones in a phased array with measurable complex sound pressures $p_1, \dots, p_M \in \mathbb{C}$. The array cross spectral matrix $C \in \mathbb{C}^{M\times M}$ of these signals is composed by the elements

$$C_{ij} = E\left[p_i p_j^*\right], \tag{1}$$

where $\cdot^*$ denotes complex conjugation, and $E[\cdot]$ is expectation. In principle, the methods of this paper are still applicable, if C would be replaced by something derived from it that adheres to its properties of self-adjointness

$$C_{ji} = C_{ij}^* \tag{2}$$

and non-negativity

$$z^H C z \geq 0 \quad \forall z \in \mathbb{C}^M, \tag{3}$$

here z^H denotes the conjugate transpose of z. For example, C could be the result of a diagonal optimization procedure (see [2,3]), an orthogonal beamforming filter (see [4]), or simply a correlation to a reference signal.

Let

$$C = \sum_{i=1}^{M} \lambda_i v_i v_i^H \tag{4}$$

be the decomposition of C with respect to its eigenvalues, i.e.

$$Cv_i = \lambda_i v_i, \quad v_i \in \mathbb{C}^M \backslash \{0\}, \quad v_i^H v_j = \delta_{ij}. \tag{5}$$

Since C is self-adjoint and non-negative, all eigenvalues are non-negative real numbers,

$$\lambda_i \geq 0. \tag{6}$$

Now, given an arbitrary steering vector g associated to a map point, the value $b_1(g)$ of the conventional beamforming map at this point is

$$b_1(g) = g^H C g = g^H \left(\sum_{i=1}^{M} \lambda_i v_i v_i^H \right) g = \sum_{i=1}^{M} \lambda_i \left| v_i^H g \right|^2 . \tag{7}$$

Again, we emphasize the universality of the formulation: there are many definitions of steering vectors in the literature, all of which have their strengths and weaknesses (see [5]). Even formulations that allow for a compensation of shear-layer effects are feasible (see [6–9]).

Using the notation

$$\rho_i(g) := \left| v_i^H g \right|^2 , \tag{8}$$

we readily have

$$b_1(g) = \sum_{i=1}^{M} \lambda_i \rho_i(g). \tag{9}$$

Moreover, since $v_1, \dots, v_M$ form an orthonormal basis of $\mathbb{C}^M$, we observe

$$\rho(g) := \sum_{i=1}^{M} \rho_i(g) = \|g\|_2^2 > 0. \tag{10}$$

This leads to the validity of

$$b_1(g) = \rho(g) \sum_{i=1}^{M} \frac{\rho_i(g)}{\rho(g)} \lambda_i. \tag{11}$$

Inspecting this equation, conventional beamforming with respect to a steering vector g can – at least up to the factor $\rho(g)$ – be interpreted as a weighted arithmetic mean of the eigenvalues λ_i with weights

$$w_i(g) := \frac{\rho_i(g)}{\rho(g)}, \quad \sum_{i=1}^{M} w_i(g) = 1. \tag{12}$$

We now assume

$$\lambda_i > 0 \quad \forall i \in \{1, \dots, M\}, \tag{13}$$

and generalize (11) using the weighted Hölder mean (or weighted power mean), that is, for each $p \in \mathbb{R}$, $p \neq 0$, we introduce

$$b_p(g) := \rho(g) \left(\sum_{i=1}^{M} w_i(g) \lambda_i^p \right)^{1/p}. \tag{14}$$

Given properties (12) and (13), it can be shown that for every fixed g the function

$$\mathbb{R}\backslash\{0\} \ni p \mapsto b_p(g) \in \mathbb{R} \tag{15}$$

is continuous, even its discontinuity at zero is removable:

$$b_0(g) := \lim_{p \to 0} b_p(g) = \rho(g) \prod_{i=1}^{M} \lambda_i^{w_i(g)} = \rho(g) \exp\left(\sum_{i=1}^{M} w_i(g) \ln(\lambda_i) \right). \tag{16}$$

Another interesting property of the weighted power mean leads to

$$b_p(g) \leq b_q(g) \tag{17}$$

if $p < q$, which results in

$$\min(\lambda_1, \dots, \lambda_M) \leq \frac{b_p(g)}{\|g\|_2^2} \leq \max(\lambda_1, \dots, \lambda_M), \tag{18}$$

since

$$\lim_{p \to -\infty} b_p(g) = \rho(g) \min(\lambda_1, \dots, \lambda_M), \quad \lim_{p \to \infty} b_p(g) = \rho(g) \max(\lambda_1, \dots, \lambda_M). \tag{19}$$

As a result, when using the type I formulation of steering vectors (see [5]), the dynamic range of any map b_p in this one-parameter family of beamforming maps has a theoretical upper limit which is solely determined by the largest and smallest eigenvalue of C.

Let us discuss the family b_p, $p \in \mathbb{R}$ in more detail. Trivially, conventional beamforming coincides with the case $p = 1$, and, whenever we are interested in suppressing side-lobes of this particular map, equation (17) suggests to consider maps b_p with $p < 1$. By choosing $p = 1/\nu$ with $\nu \geq 1$, the map b_p corresponds to functional beamforming with ν (see [1]). Moreover, adaptive beamforming and asymptotic beamforming can be identified as the special cases b_{-1} and b_0, respectively (cf. [10]). Also, it is worth noticing, that every map value is expected to change continuously with the power p, and we will call this unified approach *power beamforming* throughout this paper.

In practice, however, any error in computing the weights $w_i(g)$ leads to a significant computational error due to the sensitive nature of power beamforming, especially when considering negative p. More precisely, those summands that correspond to smaller eigenvalues - which are usually not considered to contain relevant information on the acoustics sources (cf. [4]) - start dominating the sum given that their weights are non-trivial. Usually, this issue is addressed by adding a small value $\varepsilon > 0$ to each eigenvalue, or equivalently, to consider

$$\bar{C} = C + \varepsilon \cdot \mathrm{diag}(1, \dots, 1). \tag{20}$$

Therefore, this technique is called diagonal loading and was used extensively in the case of adaptive beamforming, $p = -1$ (see [11–14]). Another interesting approach addresses the uncertainty of the steering vectors themselves, by designing a robust adaptive beamformer to maintain an acceptable accuracy within an uncertainty ellipsoid (see [15]).

3. MEASUREMENTS

All examples were produced using *Mikado* (see figure 1), a hand-held acoustic camera which consists of 96 MEMS microphones arranged in two sunflower spirals (see [16–18]) with a relative angular offset of $\pi/3$ around the optical cameras, its diameter is approximately 0.35m. The microphone

Figure 1: *Mikado*, a hand-held acoustic camera which consists of 96 MEMS microphones arranged in two sunflower spirals

positions are visualized in Figure 2. Moreover, we plotted the density function of all microphone distances there.

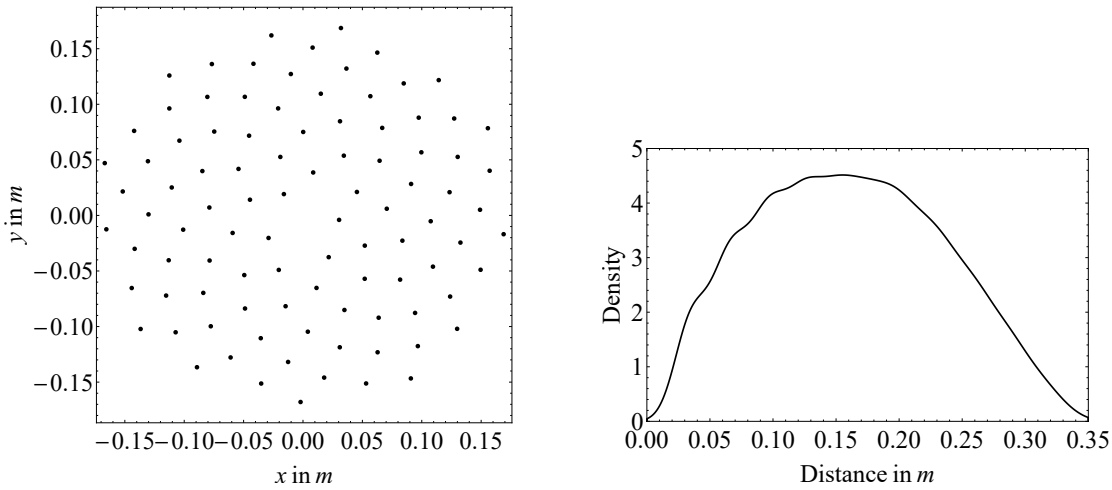

Figure 2: Microphone positions (left) and density function of microphone distances (right) of *Mikado*

The measurements were made in the hemi-anechoic chamber (dimensions: $6.05 \times 3.35 \times 2.5$m) of Technical University of Applied Sciences Wildau, Germany. The measurement object, a vacuum cleaner, was placed at a distance of 0.6m in front of the microphone array. The vacuum cleaner was held in a hanging position by rubber straps, the two orientations chosen are shown in Figure 3. Each measurement took 10s using a sampling rate of 48kHz.

Figure 3: Measurement object, side orientation (left) and front orientation (right)

4. RESULTS AND DISCUSSION

A fast Fourier transform with prior von Hann weighting was applied to every microphone signal using 50% overlapping blocks of 4096 samples. All 96^2 elements of the array cross spectral matrix (1) were computed by averaging over these blocks. Adopting the technique of section 2, we produced several maps $b_p(g)$, $p \in \mathbb{R}$ over a plane of dimensions 0.44×0.44m using steering vector formulation I of [5]. The resolution of all maps is 500×500px. As a reference, we additionally computed CLEAN-SC maps (see [19]) using the same steering vector formulation and resolution.

For both orientations of the vacuum cleaner, results are assembled in the following manner in Figure 5 and Figure 6: The rows represent different maps (from top to bottom: $p = 1$ (conventional beamforming), $p = 1/4$ (functional beamforming with $\nu = 4$), $p = 0$ (asymptotic beamforming), $p = -1/4$, $p = -1$ (adaptive beamforming) and CLEAN-SC), the columns represent the third-octave bands of the evaluation (from left to right: 4kHz, 6.3kHz, 8kHz, 10kHz, 16kHz).

Following each of the columns in Figure 5 and Figure 6 from higher to lower p, it is seen that power beamforming maintains the smoothness of conventional beamforming while at the same time improving both resolution and dynamic range of the map (cf. inequality (17)). For each of the considered frequencies a dynamic range of at least 30dB is reached in the map, but the value of p at which this can be exhibited differs. Due to the nature of the underlying method, the CLEAN-SC reference maps still have unmatched dynamic range, but they lack in smoothness. More precisely, these maps often exhibit unrealistic "quiet zones" between or around sources. This can be problematic when trying to identify or differentiate sources and is very pronounced in the front orientation results.

From a theoretical standpoint, it is feasible to consider power values $p < -1$, that is to produce beamforming maps with higher resolution and dynamic range compared to adaptive beamforming. However, from a numerical point of view, the problems described section 2 will only intensify with lower p. On the other end of the spectrum, it is possible to produce maps with $p > 1$, but it is expected that these maps will exhibit even lower resolution and dynamic range than conventional beamforming. Exemplifying this, we present the cases $p = -2$ and $p = 2$ in Figure 4.

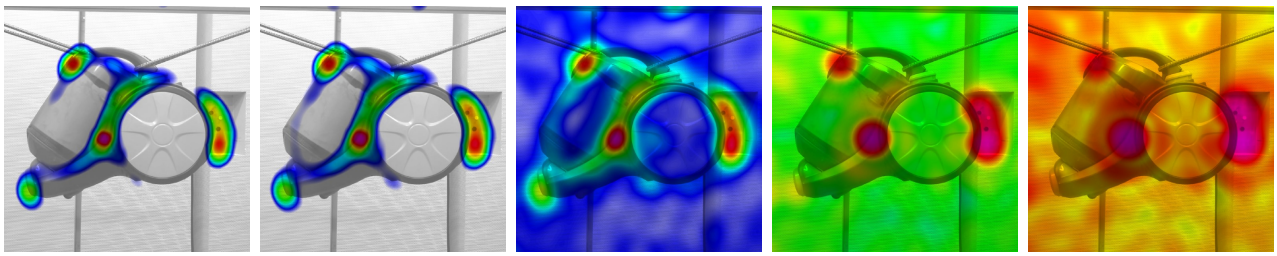

Figure 4: Results beyond adaptive and conventional beamforming, side orientation, 16kHz third-octave band, from left to right: $p = -2$, $p = -1$ (adaptive beamforming), $p = 0$ (asymptotic beamforming), $p = 1$ (conventional beamforming), $p = 2$, map scale 30dB

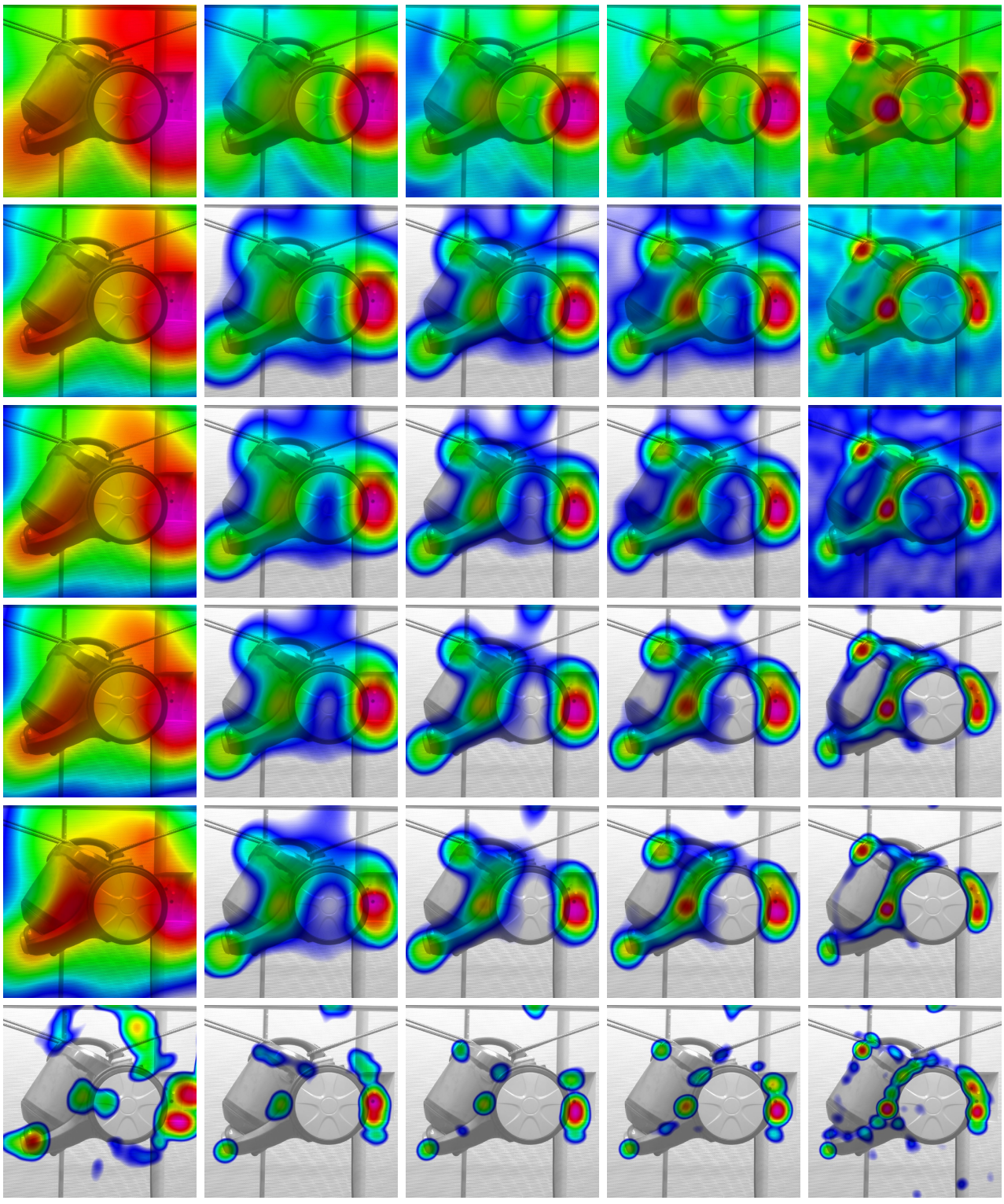

Figure 5: Results for side orientation, rows represent different maps (from top to bottom: $p = 1$ (conventional beamforming), $p = 1/4$ (functional beamforming with $\nu = 4$), $p = 0$ (asymptotic beamforming), $p = -1/4$, $p = -1$ (adaptive beamforming), CLEAN-SC), columns represent third-octave bands (from left to right: 4kHz, 6.3kHz, 8kHz, 10kHz, 16kHz), map scale 30dB

5. ACKNOWLEDGMENTS

This research work has been funded by German Federal Ministry for Economic Affairs and Energy (Bundesministerium für Wirtschaft und Energie BMWi) under project *AdaBeam* registration number 49MF180138.

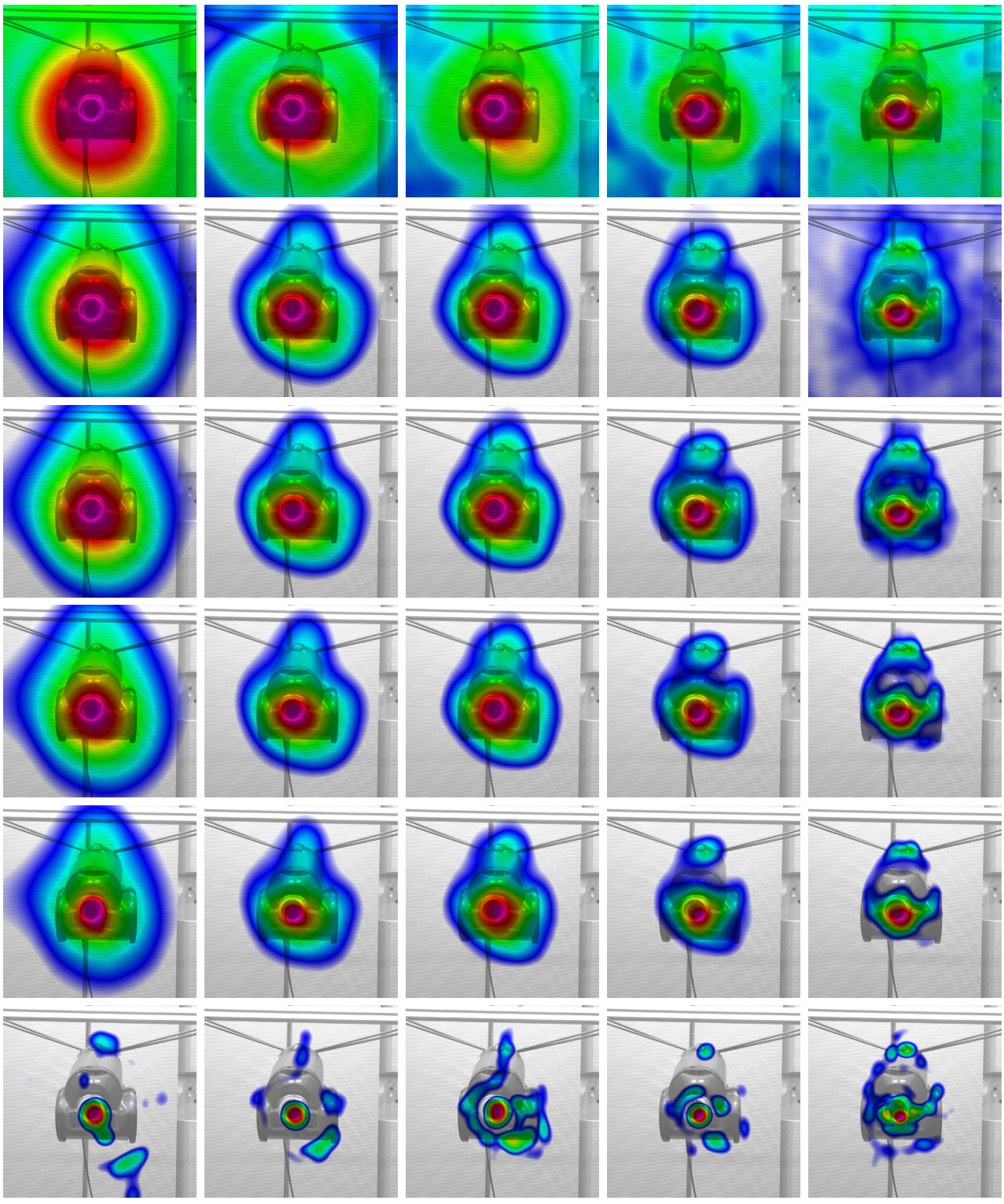

Figure 6: Results for front orientation, rows represent different maps (from top to bottom: $p = 1$ (conventional beamforming), $p = 1/4$ (functional beamforming with $\nu = 4$), $p = 0$ (asymptotic beamforming), $p = -1/4$, $p = -1$ (adaptive beamforming), CLEAN-SC), columns represent third-octave bands (from left to right: 4kHz, 6.3kHz, 8kHz, 10kHz, 16kHz), map scale 30dB

REFERENCES

[1] R.P. Dougherty, *Functional beamforming*, Proceedings of the 5th Berlin Beamforming Conference, 19-20 February 2014, 2014. ↑1, 3

[2] R.P. Dougherty, *Cross spectral matrix diagonal optimization*, Proceedings of the 6th Berlin Beamforming Conference, 29 February - 1 March 2016, 2016. ↑2

[3] J. Hald, *Cross-spectral matrix diagonal reconstruction*, Internoise 2016, 2016. 21-24 August 2016, Hamburg, Germany. ↑2

[4] E. Sarradj, *A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements*, Journal of Sound and Vibration **329** (2010), no. 9, 1553–1569. ↑2, 3

[5] E. Sarradj, *Three-dimensional acoustic source mapping with different beamforming steering vector formulations*, Advances in Acoustics and Vibration **2012** (2012), no. 292695, 1–12. ↑2, 3, 5

[6] R.K. Amiet, *Correction of open jet wind tunnel measurements for shear layer refraction*, American Institute of Aeronautics and Astronautics Paper No. 75-532 (1975). ↑2

[7] R.K. Amiet, *Refraction of sound by a shear layer*, Journal of Sound and Vibration **58** (1978), 467–482. ↑2

[8] M. Möser, *Messtechnik der Akustik*, Springer, 2010. ↑2

[9] H.S. Ribner, *Reflection, transmission and amplification of sound by a moving medium*, Journal of the Acoustical Society of America **29** (1957), 435–441. ↑2

[10] D.H. Johnson and D.E. Dudgeon, *Array signal processing: concepts and techniques*, Prentice Hall, 1993. ↑3

[11] R.A. Gramann and J.W. Mocio, *Aeroacoustic measurements in wind tunnels using adaptive beamforming methods*, Journal of the Acoustical Society of America **97** (1995), no. 6, 3694–3701. ↑3

[12] J. Li, P. Stoica, and Z. Wang, *Reflection, transmission and amplification of sound by a moving medium*, IEEE Transactions on Signal Processing **51** (2003), no. 7, 1702–1715. ↑3

[13] Y.T. Cho and M.J. Roan, *Aeroacoustic measurements in wind tunnels using adaptive beamforming methods*, Journal of the Acoustical Society of America **125** (2009), no. 2, 944–957. ↑3

[14] X. Huand, L. Bai, I. Vinogradov, and E. Peers, *Adaptive beamforming for array signal processing in aeroacoustic measurements*, Journal of the Acoustical Society of America **131** (2012), no. 3, 2152–2161. ↑3

[15] Z.S. Wang, J. Li, P. Stoica, T. Nishida, and M. Sheplak, *Constant-beamwidth and constant-powerwidth wideband robust capon beamformers for acoustic imaging*, Journal of the Acoustical Society of America **116** (2004), no. 3, 1621–1631. ↑3

[16] H. Vogel, *A better way to construct the sunflower head*, Mathematical Biosciences **44** (1979), no. 3, 179–189. ↑4

[17] J.N. Ridley, *Packing efficiency in sunflower heads*, Mathematical Biosciences **58** (1982), no. 1, 129–139. ↑4

[18] E. Sarradj, *A generic approach to synthesize optimal array microphone arrangements*, Proceedings of the 6th Berlin Beamforming Conference, 29 February - 1 March 2016, 2016. ↑4

[19] P. Sijtsma, *CLEAN Based on Spatial Source Coherence*, 13th AIAA/CEAS Aeroacoustics Conference, May 21-23, 2007. ↑5