24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

INVESTIGATION OF THE GROUP LASSO ALGORITHM FOR SOUND FIELD REPRODUCTION: COMPARISON WITH THE LASSO AND ELASTIC-NET ALGORITHMS

Philippe-Aubert Gauthier, Pierre Grandjean, Alain Berry
Groupe d'Acoustique de l'Université de Sherbrooke, Université de Sherbrooke, Sherbrooke, Canada
Centre for Interdisciplinary Research in Music, Media, and Technology, McGill University, Montréal, Canada
email: philippe_aubert_gauthier@hotmail.com

The reproduction of a sound field measured using a microphone array is an active topic of research. To this end, loudspeaker and microphone arrays are used. Classical methods rely on spatial transforms (such as spherical Fourier transform for Ambisonics) or pressure matching using least-mean-square formulation. For both methods, all the reproduction sources (i.e. loudspeakers) will typically be activated. Although this can provide a reduced or minimized reproduction error evaluated at the microphone array, it is not necessarily the most useful solution for listening purposes. Indeed, the fact that all reproduction sources are concurrently active can potentially lead to a blurry spatial image (an example would be the common front-back confusion). To solve that potential limitation, the lasso and elastic-net algorithms were recently studied in order to favor the sparsity of the reproduction source signals. In these studies, it was shown that sparsity can indeed lead to a sharper source distribution at the cost of reduced physical accuracy of the reproduced sound field. In this paper, the group lasso is investigated to alleviate such potential limitations of the lasso, where the "group" refer to groups of reproduction sources. The aim of the group lasso is to provide sparsity at the group level and continuous smooth solution inside groups. In the recent literature, many simple or detailed algorithms have been proposed for the real-valued group lasso without a consensual position from the community. For sound field reproduction, a simple algorithmic implementation is proposed as an interpretation of the group lasso for complex quantities. Simulation results in free field show that several of the limitations of the lasso and elastic-net algorithms are solved. Potentials and current limitations of the group lasso are discussed. Based on the reported investigation, future research openings include: the possibility of overlapping groups and the algorithmic implementation.
Keywords: sound field reproduction, group lasso

1. Introduction

In spatial audio, Sound Field Reproduction (SFR) is aimed at the extended reproduction of a sound pressure field within a listening area. In this paper, we are concerned with the scenario where a microphone array is used to capture an unknown target sound field to be reproduced in a subsequent step using an array of sources. Classical methods to convert a microphone array recording (regular or irregular) into suitable inputs of reproduction sources include spatial transforms (such as spherical Fourier transform for Ambisonics) or pressure matching using least-square formulations. For both methods, all the reproduction sources will typically be active in order to minimize the reproduction error at the microphone array. However, this might not be ideal for listening purposes: The fact that all sources are active can lead to a blurry spatial image or confusion of source localization. One possible

approach to solve that issue is the use of a least-square formulation while looking for solution sparsity, i.e. a limited amount of concurrently active reproduction sources to increase the solution contrast. For linear regression and least-square problems, 1-norm regularization can induce solution sparsity [1]. Common methods are the lasso [1] and elastic-net [2] cost functions.

In the context of SFR, the lasso [1] and the elastic-net [2] introduced interesting features: 1) reduced amount of concurrently active source, and 2) automatic snapping to the closest reproduction source [3, 4]. One of the obvious potential limitation is that in the case of extremely sparse solution, the automatic snapping feature can, for sufficiently dense reproduction source array, typically provides the correct incoming direction but with an erroneous reproduced wave front curvature [4]. This paper is a continuation or our previous work [4], but the group lasso [5] is investigated. The group lasso also combines 1-norm and 2-norm regularization [5]. However, the group lasso, when adapted to SFR, imposes 1-norm regularization between groups of reproduction sources and 2-norm regularization within groups of reproduction sources. Therefore, the group lasso favors sparsity between group but provides non-sparse solution within active groups. This has foreseeable advantages over the 2-norm regularization, the lasso and the elastic-net. First, for the group lasso, once a group is activated, all sources within this group will be active and a correct wave front curvature might therefore be recreated by the group of closely spaced-reproduction sources. Second, if groups are logically defined (e.g. front, left, right, top, and so on), the potential issue of localization blur or front-back confusion might be solved. Indeed, these issues are caused by active reproduction sources that are not between the sources that created the target sound field and the listening area.

2. Theoretical background: lasso, elastic-net, group lasso

The direct problem at angular frequency ω (rad/s) relates the reproduced sound pressures $\mathbf{p}$ to the reproduction source signals $\mathbf{s}$: $\mathbf{p} = \mathbf{G}\mathbf{s}$ with $\mathbf{p} = [p_1 \ \cdots \ p_M]^T \in \mathbb{C}^{M\times 1}$, $\mathbf{G} \in \mathbb{C}^{M\times L}$ is a known transfer matrix, and reproduction source signals vector $\mathbf{s} = [s_1 \ \cdots \ s_L]^T \in \mathbb{C}^{L\times 1}$. The arrays include M microphones and L reproduction sources (loudspeakers). In this paper, simulations are provided for the free-field case, therefore spherical waves are used for the reproduction sources: $G_{ml} = e^{jkr_{ml}}/r_{ml}$ where r_{ml} is the Euclidean distance between microphone m and source l, k is the wavenumber. The target sound field is denoted $\mathbf{d} \in \mathbb{C}^{M\times 1}$. The p-norm is $\|\mathbf{s}\|_p = (\sum_{l=1}^{L} |s_l|^p)^{1/p}$ with the 1-norm: $\|\mathbf{s}\|_1 = \sum_{l=1}^{L} |s_l|$ and the 0-norm $\|\mathbf{s}\|_0$: the number of non-zero elements.

2.1 The lasso and the elastic-net

Both the lasso [1, 7] and the elastic-net [2] were recently compared by the authors [4] for SFR. The corresponding cost-function is recalled [4]:

$$J_{\lambda,\alpha} = \frac{1}{2}\|\mathbf{d} - \mathbf{G}\mathbf{s}\|_2^2 + \lambda\alpha\|\mathbf{s}\|_1 + \frac{(1-\alpha)}{2}\lambda\|\mathbf{s}\|_2^2 \quad \text{and} \quad \mathbf{s}_{\lambda,\alpha} = \operatorname{argmin}(J_{\lambda,\alpha}) \tag{1}$$

The amount of regularization is controlled by λ. The α parameter controls the transition between the lasso $\alpha = 1$ and the elastic-net $1 > \alpha \geq 0$. The first term of Eq. (1) introduces reproduction error minimization. The presence of the solution 1-norm introduces selection of a limited amount of active sources in the solution $\mathbf{s}_{\lambda,\alpha}$. A larger λ reduces the number of active sources. One advantage of the lasso is that the solution path (i.e. source coefficients s_l as function of penalization λ) activates one (or very few) sources at a time while decreasing λ. It is also known that for $\lambda \gtrsim \|\mathbf{G}^H\mathbf{d}\|_\infty$, there is no active source [3]. Therefore, λ was introduced as a control of reproduction source sparsity in [3, 4].

Since the cost function includes the 1-norm, the optimal solution $\mathbf{s}_{\lambda,\alpha}$ cannot be obtained by analytical means. Because of its simplicity and widespread use, the coordinate descent algorithm is used (see Section 3) to find the solution iteratively. With respect to the selection of the penalization amount λ, if one sets a parameter $L_{\max}$ defined as the maximum number of concurrently active sources, a decreasing λ search is performed from $\|\mathbf{G}^H\mathbf{d}\|_\infty$ until $\|\mathbf{s}\|_0 \approx L_{\max}$.

2.2 The classical least-square solution

The classical least-square (LS) solution with Tikhonov regularization will be used for comparison purpose. It is derived from the cost function given by Eq. (1) with $\alpha = 0$. In that case, the complex gradient can be derived and set to zero to obtain the following optimal solution $\mathbf{s}$ [8]:

$$\mathbf{s}_{\lambda,0} = (\mathbf{G}^H\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^H\mathbf{d} \tag{2}$$

In this paper, λ is set so that it corresponds to the group lasso penalization amount.

2.3 The group lasso

The group lasso was initially introduced by Yuan and Lin in 2006 [5]. An implementation based on the coordinate and block coordinate descent algorithms was later proposed for the sparse group lasso [6]. In this paper, the original version and algorithmic approach [6] to the group lasso are adapted to the problem at hand. It is reminded that the original implementation of the group lasso was provided for real-valued quantities. In this paper, the group lasso is adapted to complex quantities. The corresponding algorithm, which is an interpretation of the group lasso, is discussed in Sec. 3.

There are N predefined groups of reproduction sources so that $\mathbf{G}$ and $\mathbf{s}$ can be partitioned: $\mathbf{G} = [\mathbf{G}_1 \cdots \mathbf{G}_n \cdots \mathbf{G}_N]$ and $\mathbf{s} = [\mathbf{s}_1^T \cdots \mathbf{s}_n^T \cdots \mathbf{s}_N^T]^T$. It is assumed that there is no overlap between groups and that each group includes L_n sources. The cost function of the group lasso is given by:

$$J_\beta = \frac{1}{2}\|\mathbf{d} - \sum_{n=1}^{N}\mathbf{G}_n\mathbf{s}_n\|_2^2 + \beta\sum_{n=1}^{N}\|\mathbf{s}_n\|_2 \quad \text{and} \quad \mathbf{s}_\beta = \operatorname{argmin}(J_\beta) \tag{3}$$

The second term introduces 1-norm regularization at group level and 2-norm regularization within a group [6]. Therefore the group lasso sparsity operates at the group level: depending on β, an entire group may be active or not. Because of the specific nature of the group lasso, a different algorithm must be used for the computation of the optimal solution that will minimize Eq. (3) [6]. These algorithmic implementations are discussed in Sec. 3. In this paper, β is a user-defined parameter.

3. Complex coordinate descent and iterative group lasso solution

3.1 Complex coordinate descent for the lasso and the elastic-net

Within the context of SFR, reference [3] provides a complex version of the coordinate descent for the lasso. The complex coordinate descent algorithm for the lasso and the elastic-net is found in [4]. Since the group lasso requires the knowledge of the sources belonging to which group, the coordinate descent cannot be used for the group lasso. An iterative implementation of current group lasso algorithms is proposed.

3.2 Group lasso algorithmic solution

The algorithmic solution of the sparse group lasso proposed by Simon *et al.* [6], from which our proposal is derived, is briefly recalled for the case of the group lasso [5]. Our algorithm is defined as (details are provided in the appendix):

1. Initialize iteration index $i = 0$, and initially set β (optimal selection of β is left for future research)
2. Start with an initial solution, e.g. a LS solution with all groups active:

$$\mathbf{s}^{(0)} = (\mathbf{G}^H\mathbf{G} + \beta\mathbf{I})^{-1}\mathbf{G}^H\mathbf{d} \tag{4}$$

3. Start the iteration with $i = 1$
4. For iteration index i

(a) Initialize group iteration $n = 1$, copy the solution $\mathbf{s}^{(i)} = \mathbf{s}^{(i-1)}$

(b) For group n, define the following partial residual with group n contribution missing

$$\mathbf{r}_n^{(i)} = \mathbf{d} - \sum_{\substack{k=1 \\ k \neq n}}^{N} \mathbf{G}_k \mathbf{s}_k^{(i)} \tag{5}$$

(c) Perform n-th group-level test based on

$$\text{if } \|\mathbf{G}_n^H \mathbf{r}_n^{(i)}\|_2 \leq \beta, \text{ then } \mathbf{s}_n^{(i)} = \mathbf{0} \tag{6}$$

(d) If $\|\mathbf{G}_n^H \mathbf{r}_n^{(i)}\|_2 > \beta$ then minimize the regularized 2-norm of the error between the partial residual $\mathbf{r}_n^{(i)}$ and the pressure reproduced by group n:

$$J_{\beta,n} = \frac{1}{2}\|\mathbf{r}_n^{(i)} - \mathbf{G}_n \mathbf{s}_n^{(i)}\|_2^2 + \beta \|\mathbf{s}_n^{(i)}\|_2 \quad \text{and} \quad \mathbf{s}_n^{(i)} = \text{argmin}\{J_{\beta,n}\} \tag{7}$$

Since the regularization is based on $\|\mathbf{s}_n\|_2$ and not $\|\mathbf{s}_n\|_2^2$, a Tikhonov-like explicit solution is not available (see Appendix) and an iterative gradient descent (iteration j) is used with step size μ

$$\mathbf{s}_n^{(i,j+1)} = \mathbf{s}_n^{(i,j)} - \mu \frac{\partial J_{\beta,n}(\mathbf{s}_n^{(i,j)})}{\partial \mathbf{s}_n} = \mathbf{s}_n^{(i,j)} - \mu \left(\mathbf{G}_n^H \mathbf{G}_n \mathbf{s}_n^{(i,j)} + \frac{\beta \mathbf{s}_n^{(i,j)}}{\|\mathbf{s}_n^{(i,j)}\|_2} - \mathbf{G}_n^H \mathbf{r}_n^{(i)} \right) \tag{8}$$

(e) Repeat steps 4.(b) to 4.(e) for all groups $n = 1...N$

5. Stop iteration i when a stop criterion on the reproduction error $\|\mathbf{d} - \mathbf{G}\mathbf{s}\|_2$ or on the source input $\mathbf{s}$ is reached. Otherwise, continue.
6. Repeat step 4 for next iteration index $i + 1$.

Theoretical developments and explanations for Eqs. (6) and (8) are provided in the appendix.

4. Simulation results

4.1 Array configuration

The loudspeaker and microphone arrays used for the simulation are shown in Fig. 1(a). The loudspeaker array circumscribes a 4×4 m horizontal area. There are $L = 96$ loudspeakers arranged in $N = 4$ groups corresponding to straight bars of 24 loudspeakers each. The microphone array includes $M = 9$ microphones arranged in a sparse 1-m-radius circular array with a central microphone.

4.2 Results

Simulation results are first reported for the simple case of a harmonic target sound field caused by a single spherical wave pulsating at 440 Hz with sound speed $c = 343$ m/s. The real part of the target sound field is shown in Fig. 1(a). The sound fields reproduced using the lasso ($\alpha = 1$, $L_{\max} = 24$), the elastic-net ($\alpha = 0.1$, $L_{\max} = 24$), the group lasso ($\beta = 0.5$) and the least-square solutions ($\lambda = 0.5$) are shown in Fig. 1(b) to (e), respectively. The aim of this simulation is to validate the group lasso iterative solution, check for group activity and compare the four solutions. Since each group includes 24 sources, the maximum number of concurrently active sources $L_{\max}$ was set to 24 for the lasso and elastic-net. First, one notes that the lasso and elastic-net solutions effectively reproduce the target field (Figs. 1(b) and (c)) and provides sparse solutions. However, some active reproduction sources are visible on the right, left and bottom sides of the reproduction array. Ideally, they should not be active since they do not correspond to target sound field incoming direction (i.e. from the top of the figures, as in Fig. 1(a)). For both the lasso and the elastic-net, the reproduced sound pressure level is slightly below the target sound pressure level (see Fig. 1(f)). This is an expected behavior of both solutions caused by the shrinkage of the solution coefficients [1, 2]. In practical situations,

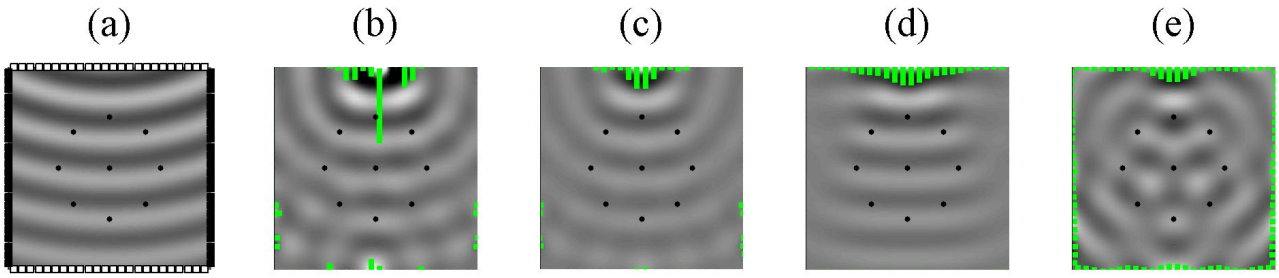

Figure 1: (a) Microphone and loudspeaker arrays in the horizontal plane. The $L = 96$ loudspeakers are shown as square markers. There are four source groups ($N = 4$) where the group alternation corresponds to black and white marker alternation. One group is the bottom bar, the other the right bar, and so on. The $M = 9$ microphones are shown as black bullets. Top views (x_1–x_2 horizontal plane) of the real parts of sound fields (Pa) at 440 Hz. (a) Target sound field, (b) reproduced using the lasso ($\alpha = 1$, $L_{\max} = 24$), (c) reproduced using the elastic-net ($\alpha = 0.1$, $L_{\max} = 24$), (d) reproduced using the group lasso ($\beta = 0.5$), (e) reproduced using the LS solution ($\lambda = 0.5$). Loudspeakers signals $|s_l|$ and active loudspeakers are shown as green bars.

this can be simply compensated using a global gain adjustment where the 2-norm of the reproduced sound pressure at the microphone array is adjusted to fit the 2-norm of the target sound pressure at the microphones. The reproduced sound field using the group lasso also correctly reproduces the target sound field and only one group of reproduction sources is active: the top bar which corresponds to the target sound field incoming direction. Furthermore, as it is theoretically expected for the group lasso, all the sources within the active group are active: only between-group sparsity is observed for the group lasso. It is therefore concluded that the proposed algorithmic implementation and group lasso cost function are both correct.

For the LS solution (Fig. 1(e)), since the microphone array is relatively sparse, the sound field is only correctly reproduced at the microphone positions and all reproduction sources are active. This reduces the performance of the LS solution in comparison with the other solutions.

The group lasso provides several advantages. First, since an entire group of neighboring sources are active, the curvature of the wave front can be reproduced more precisely over an extended area. Second, since the active sources tend to be limited to an area between the virtual target sound and the listening area, there should be less risk of front-back confusion or a resulting blurry spatial image.

The normalized reproduction error fields are shown in Fig. 2 using $E(x_1, x_2) = |d(x_1, x_2) - p(x_1, x_2)|/|d(x_1, x_2)|$ with the target sound pressure field $d(x_1, x_2)$ and reproduced sound field $p(x_1, x_2)$. Clearly, the lasso (Fig. 2(a)), the elastic-net (Fig. 2(b)), and the group lasso (Fig. 2(c)) provide a small reproduction error in the listening area. However, it is clear that the group lasso offers a much smoother reproduction error in the listening area. This is caused by the fact that a single cluster, a group, of reproduction sources is active for the group lasso. For the lasso and the elastic-net, even if the same amount of sources are active, their large separation distances create interference patterns in the resulting error fields. The LS solution provides, as expected, the lowest reproduction error but only at the exact microphone positions. Outside these control points, the LS solution provides normalized reproduction error in the range of (or above) one.

The solution paths (i.e. each coefficient evolution $|s_l|$ as function of β or λ) are shown in Fig. 3.

4.3 Case of multiple sources

This section presents the case of a target sound field created by two virtual sources at 440 Hz. The target and reproduced sound fields are shown in Fig. 4. For the reported reproduced sound fields (Fig. 4(b) to (d)), the shrinkage effect was compensated to illustrate the shrinkage compensation principle. In Fig. 4, this shrinkage is compensated by a simple post-processing gain adjustment.

The LS solution (Fig.4(e)) is not satisfying as the reproduction is only accurate at the microphone

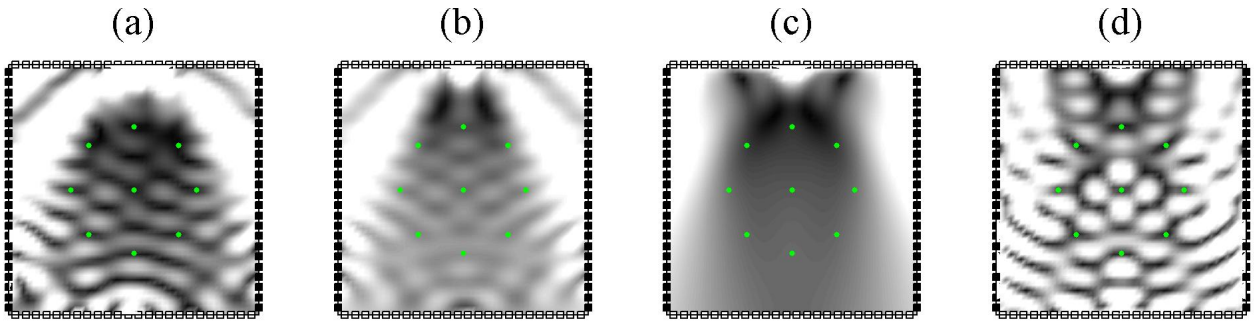

Figure 2: Local normalized reproduction error $E(x_1, x_2) = |d(x_1, x_2) - p(x_1, x_2)|/|d(x_1, x_2)|$ for the cases shown in Fig. 1 for the (a) lasso ($\alpha = 1$, $L_{\max} = 24$), (b) elastic-net ($\alpha = 0.1$, $L_{\max} = 24$), group lasso ($\beta = 0.5$), LS ($\lambda = 0.5$). The color range is the same for each plot with a color range of 0 (black) to 1 (white) (linear scale). Microphones are shown as green dots.

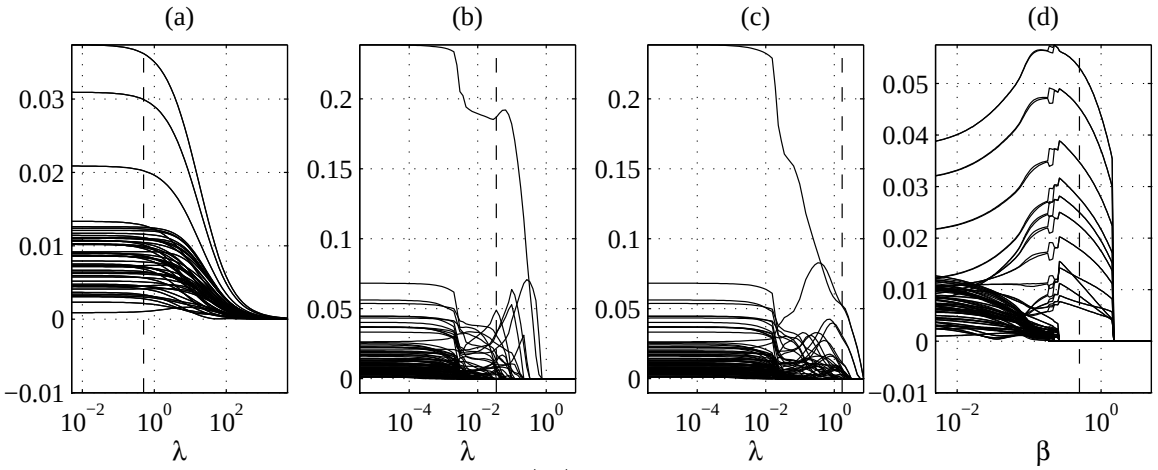

Figure 3: Solution paths, i.e. $L = 96$ traces of $|s_l|$ as function of penalization amount (λ or β). (a) LS. (b) Lasso with $\alpha = 1$. (c) Elastic-net with $\alpha = 0.1$. (d) Group lasso. The λ or β values to obtain the desired sparsity or group sparsity are shown as vertical dashed lines.

locations. Furthermore, all reproduction sources are active even if the target sound field correspond to only two virtual sources. Localization accuracy is therefore expected to be reduced. Both the lasso and the elastic net (Fig. 4(b) and (c)) reproduce the sound field with a sparse solution limited to $L_{\max} = 24$ active sources. However, active sources are noticed in the left bottom corner of the figures. The group lasso provides an interesting solution with only two active groups of reproduction sources. Furthermore, these two active groups correspond to the incoming directions of the target sound field (Fig. 4(a)). This represents an interesting feature of the group lasso that favors sparsity at group level: it avoids the creation of interference patterns caused by active sources in directions not related to the target sound field. Accordingly, the overall performance of the group lasso is demonstrated. Furthermore, even with an incomplete target sound field sampling using a sparse microphone array, the group lasso solution provides an accurate sound field reproduction with correct wavefront curvatures over an extended listening area (Fig. 4(d)). Error fields are shown in Fig. 5.

5. Conclusion

This paper investigated the group lasso cost function applied to SFR. In comparison with the lasso and the elastic-net solutions, the group lasso induces group-level sparsity. As shown in this paper, this has great potential to circumvent some of the limitations of the sparse solutions obtained from the lasso or the elastic-net. Indeed, it was shown that the group lasso is able to recreate a correct

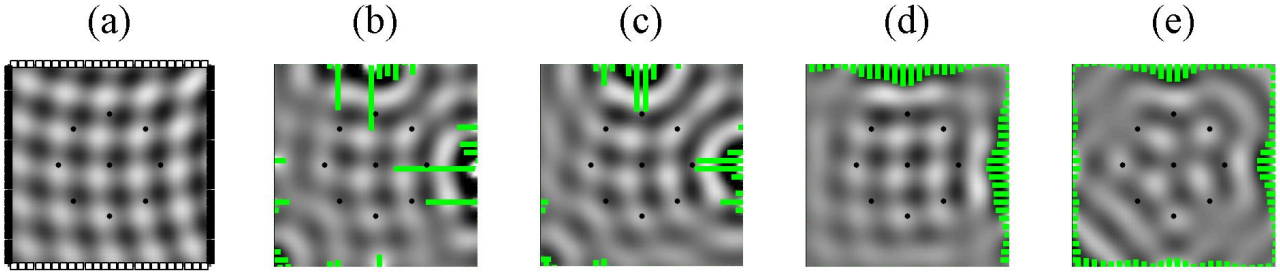

Figure 4: Top views (x_1–x_2 horizontal plane) of the real parts of sound fields (Pa). (a) Target sound field, (b) reproduced using the lasso ($\alpha = 1$, $L_{\max} = 24$, with shrinkage compensation), (c) reproduced using the elastic-net ($\alpha = 0.1$, $L_{\max} = 24$, with shrinkage compensation), (d) reproduced using the group lasso $\beta = 1.5$ (with shrinkage compensation), (e) reproduced using the LS solution ($\lambda = 1.5$). Loudspeakers signals $|s_l|$ and active loudspeakers are shown as green bars.

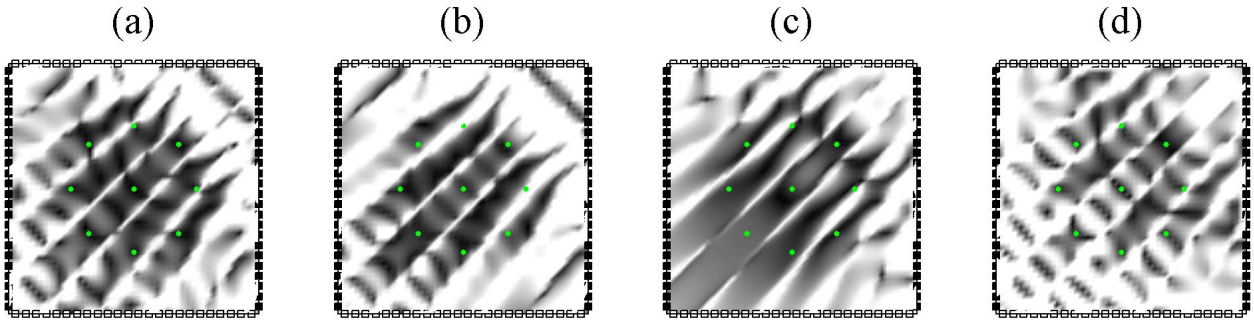

Figure 5: Local normalized reproduction error for the cases shown in Fig. 4. (a) Lasso ($\alpha = 1$, $L_{\max} = 24$, with shrinkage compensation), (b) elastic-net ($\alpha = 0.1$, $L_{\max} = 24$, with shrinkage compensation), (c) group lasso ($\beta = 1.5$, with shrinkage compensation), (d) LS ($\lambda = 1.5$). The color range is the same for each plot with a color range of 0 (black) to 1 (white) (linear scale). Microphones are shown as green dots.

wave front curvature by using clusters of adjacent active reproduction sources. By marked contrast, the lasso and the elastic-net often distort the reproduced wave front curvature. Furthermore, it has been shown that the group lasso has the potential to only activate groups of sources that correspond to incoming directions in the target sound field, hence potentially reducing the risk of front-back confusion or blurry sound localization.

REFERENCES

1. R. Tibshirani. Regression Selection and Shrinkage via the Lasso. *Journal of the Royal Statistical Society B*, **58**:267–288, 1996.

2. H Zou and T Hastie. Regularization and Variable Selection via the Elastic-net. *Journal of the Royal Statistical Society*, **67**:301–320, 2005.

3. G. N. Lilis, D. Angelosante, and G. B. Giannakis. Sound Field Reproduction Using the Lasso. *IEEE Transactions on Audio, Speech and Language Processing*, **18**(8):1902–1912, 2010.

4. P.-A. Gauthier and Berry A. Sound Field Reproduction of Microphone Array Recordings Using the Lasso and the Elastic-Net: Theory, Application Examples and Artistic Potentials. In *inSONIC2015*, Karlsruhe, 2015.

5. M. Yuan and Y. Lin. Model Selection and Estimation in Regression with Grouped Variables. *Journal of the Royal Statistical Society B*, **68**:49–67, 2006.

6. N. Simon, J. Friedman, T. Hastie, and R. Tibshirani. A Sparse-Group Lasso. *Journal of Computational and Graphical Statistics*, **22**(2):231–245, 2013.

7. R. J. Tibshirani. The Lasso Problem and Uniqueness. *Electronic Journal of Statistics*, **7**:1456–1490, 2013.

8. P.A. Nelson. *Active Control of Sound*. Academic Press, 1992.

6. Appendix: Group lasso and subgradient equations

The subgradient is required to find the solution of the group lasso cost function. The subgradient, subderivative, and subdifferential are generalizations of derivative to non-differentiable functions. A subderivative of a function $f(x)$ in x_0 is the slope of any line passing by $(x_0, f(x_0))$ below the graph of $f(x)$ and outside $\mathrm{epi}(f(x))$ (the epigraph). The subgradient equations of Eq. (3) for group n at optimal solution $\hat{\mathbf{s}}$ are given by [6]

$$\mathbf{G}_n^H(\mathbf{d} - \sum_{k=1}^{N} \mathbf{G}_k \hat{\mathbf{s}}_k) = \beta \mathbf{u} \quad \text{with} \tag{9}$$

$$\mathbf{u} \in \mathbb{C}^{L_n \times 1} = \begin{cases} \frac{\hat{\mathbf{s}}_n}{\|\hat{\mathbf{s}}_n\|_2} \; \forall \; \hat{\mathbf{s}}_n \neq \mathbf{0} \\ \in \{\mathbf{u} : \|\mathbf{u}\|_2 \leq 1\} \text{ if } \hat{\mathbf{s}}_n = \mathbf{0} \end{cases} \tag{10}$$

If $\hat{\mathbf{s}}_n = \mathbf{0}$, we have

$$\mathbf{G}_n^H \mathbf{r}_n = \beta \mathbf{u} \quad \text{if } \hat{\mathbf{s}}_n = \mathbf{0} \tag{11}$$

with partial residual given by

$$\mathbf{r}_n = \mathbf{d} - \sum_{\substack{k=1 \\ k \neq n}}^{N} \mathbf{G}_k \mathbf{s}_k \tag{12}$$

Then, in terms of norms, one finds

$$\|\mathbf{G}_n^H \mathbf{r}_n\|_2 = \beta \|\mathbf{u}\|_2 \quad \text{if } \hat{\mathbf{s}}_n = \mathbf{0} \tag{13}$$

However, with Eq. (10) for $\hat{\mathbf{s}}_n = \mathbf{0}$, we have $\|\mathbf{u}\|_2 \leq 1$ so that one can write

$$\|\mathbf{G}_n^H \mathbf{r}_n\|_2 \leq \beta \quad \text{if } \hat{\mathbf{s}}_n = \mathbf{0} \tag{14}$$

Therefore, Eq. (14) is the group-level test as found in Step. 4.(c) of the algorithm. To find a group n to minimize over (with $\hat{\mathbf{s}}_n \neq \mathbf{0}$), a new optimization problem is introduced

$$J_{\beta,n} = \frac{1}{2} \|\mathbf{d} - \sum_{\substack{k=1 \\ k \neq n}}^{N} \mathbf{G}_k \hat{\mathbf{s}}_k - \mathbf{G}_n \hat{\mathbf{s}}_n\|_2^2 + \beta \|\hat{\mathbf{s}}_n\|_2 = \frac{1}{2} \|\mathbf{r}_n - \mathbf{G}_n \hat{\mathbf{s}}_n\|_2^2 + \beta \|\hat{\mathbf{s}}_n\|_2 \tag{15}$$

For $\frac{\partial J_{\beta,n}}{\partial \mathbf{s}_n} = \mathbf{0}$ we have an implicit solution that is not like the explicit straightforward Tikhonov-like regularized solution of Eq. (15):

$$\hat{\mathbf{s}}_n = \left(\mathbf{G}_n^H \mathbf{G}_n + \frac{\beta}{\|\hat{\mathbf{s}}_n\|_2} \mathbf{I} \right)^{-1} \mathbf{G}_n^H \mathbf{r}_n \quad \text{with} \quad \frac{\partial J_{\beta,n}}{\partial \mathbf{s}_n} = \left(\mathbf{G}_n^H \mathbf{G}_n + \frac{\beta}{\|\hat{\mathbf{s}}_n\|_2} \mathbf{I} \right) \hat{\mathbf{s}}_n - \mathbf{G}_n^H \mathbf{r}_n \tag{16}$$

This is so since the regularization is $\|\mathbf{s}_n\|_2 = \sqrt{\mathbf{s}_n^H \mathbf{s}_n}$ and not $\|\mathbf{s}_n\|_2^2 = \mathbf{s}_n^H \mathbf{s}_n$. Then we introduce a new iteration index j, and solve iteratively using the gradient descent as shown in Eq. (8). This corresponds to Step 4.(d) of the proposed algorithm.