Improving Speech Intelligibility using Numerical Sound System Optimization

Wolfgang Ahnert & Stefan Feistel
AFMG Technologies GmbH, Arkonastr. 45-49, 13189 Berlin, Germany

ABSTRACT

Compared to conventional loudspeakers, steered columns and loudspeaker arrays offer measurably increased speech intelligibility in room acoustically difficult spaces. Their configurability allows focusing sound on the audience and avoiding unwanted reflections to a certain degree. Output SPL and sound field uniformity can be improved even further by numerically optimizing the array configuration using an acoustic computer model. Especially suited for this new optimization approach are FIR filters, since nowadays FIR processing is often part of the DSP functions of loudspeaker systems. Several real-world case studies are presented based on systematic measurements and simulations. It is shown how numerical sound system optimization can substantially improve direct sound coverage, signal/noise ratio as well as speech intelligibility in reverberant spaces. Listening tests support these findings and the feasibility of this new technology.

1. INTRODUCTION

Line array systems as well as column loudspeakers have found widespread use in installed sound systems due to their ability to radiate sound in a much more controlled way than available with conventional sources. This provides, e.g., increased speech intelligibility in acoustically difficult spaces and higher SPL in open-air scenarios where long distances have to be covered. DSP-based computing power can be integrated for the purpose of advanced signal processing and provides even better performance and an unseen degree of user configurability.

Due to the challenges of this new technology, the use of acoustic simulation software [1] has to be considered a standard approach in planning and commissioning line arrays as well as digitally steered column loudspeakers. Increasingly, acoustic modeling tools also offer optimization functions in order to shape the radiated sound field based on the requirements given by the user ([2], [3], [4]) instead of a time-consuming trial-and-error approach.

Modern line arrays mostly use substantial signal processing, may it be in an external DSP controller or internal for powered systems. It seems therefore natural to determine the best DSP settings regarding delay, gain and EQ using a computer model for the venue. In particular, FIR filters can be computed automatically in a numerical optimization process [5] that uses venue geometry, user requirements and loudspeaker modeling data as its input. By optimizing the superposition of the elemental sound waves along the audience zones, calculated FIR filters can establish a more uniform sound field and higher SPL figures than can be achieved by manual tuning in the conventional way.

In this article the theoretical background of numerical optimization in the context of sound system tuning is outlined. Two case studies are presented in order to demonstrate the feasibility and effectiveness of this concept. Based on a number of extensive field measurements in a medium-size hall and in a small stadium, the predicted performance for different types of loudspeaker arrays is confirmed and validated. It is shown that numerically computed FIR filters can be used to establish a very uniform SPL distribution as well as very homogeneous speech intelligibility throughout the venue. The level of control achieved with such optimized FIR filters was previously unknown and cannot be reached by conventional means. Additionally, certain zones of the room can be excluded from the irradiation of direct sound. These and other goals have to be considered as different, potentially conflicting options in the optimization process for which the balance must be defined by the user.

2. BACKGROUND

Computer-based design of sound systems is typically performed in an iterative manner:

i. Loudspeaker system and room model are entered or changed.

ii. Simulation runs are conducted based on the new configuration.

iii. Results are compared against design goals and performance criteria.

These steps are repeated numerous times during a professional design process. Therefore it seems desirable that at least a subset of the iterations is executed automatically. This is what we will call *sound system design using numerical optimization.* We emphasize that the term "optimization" relates to the mathematical concept of optimization and not to the subjectively experienced quality of the sound system. The user still has to make design decisions and weigh different trade-offs against each other. But numerical optimization offers increased flexibility and control capabilities to achieve the desired results.

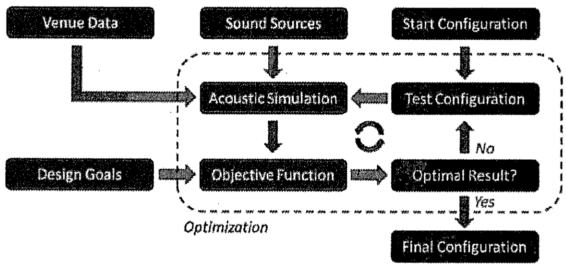

Figure 1: Sound system design using numerical optimization.

As shown in Fig. 1, numerical optimization based on acoustic simulation requires input data such as the venue geometry, the loudspeaker data, and the design goals. Using these data, the optimization algorithm will then perform acoustic simulations for a large number of different sound system configurations or sets of loudspeaker parameters. Comparing the results for each set with the design goals allows the algorithm to find the best configuration given that an appropriate metric has been defined. In order to establish such a metric the user has to assign priorities to different optimization goals, such as SPL uniformity, sound radiation directed towards the audience area, available head room, or the importance of excluding specific zones from sound irradiation.

Mathematically, this definition leads to an objective function [5] that uses the predicted sound field quantities in order to automatically compute a rank for each tested configuration. In a simple case, the objective function F may consist only of the standard deviation of SPL across all listening locations:

$$F(\{X_i\}) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(L_k - \langle L\rangle\right)^2}\,, \quad (1)$$

where $\{X_i\}$ is the set of adjustable parameters, L_k denotes the SPL at location k and $<L>$ is the average SPL over all N locations, that is, over the entire audience zone. The actual values for the sound pressure level are calculated based on the current configuration parameters X_i. For the case above, the numerical algorithm will iterate through different sets X_i in order to minimize the objective function and thereby maximize the uniformity of SPL in the venue.

Our focus in this work will be on using FIR filters as variable parameters X_i for the optimization process, Fig. 2. The results presented in the following sections have been obtained using the AFMG FIRmaker technology [6] in combination with the actual version of EASE Focus 2 [1]. Future versions of EASE are also expected to support FIRmaker optimization to an increasing degree.

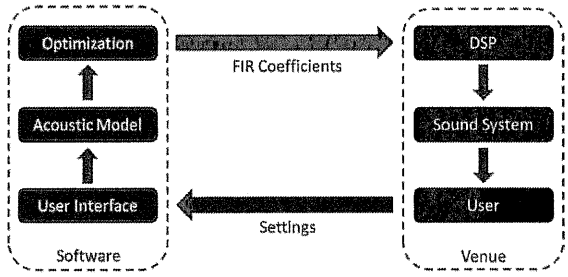

Figure 2: Optimization of FIR filters.

3. CASE STUDIES

3.1. SPL Uniformity

A field test was performed in a former movie-recording hall of about 40 m length and 25 m width. A medium-size state-of-the-art line array system was used for the test. The array consisted of 8 two-way cabinets, each of them driven by one FIR channel using a Four-Audio HD2 DSP controller with 1300 taps per channel at 48 kHz. The line array was flown at a height of about 5 m and mechanically aimed at the intended coverage area, see Figs. 3 and 4. A total of 106 impulse response measurements were taken in 25 cm spacing along a distance of 26 m on the axis of the array. In order to extract the direct sound transfer functions the impulse responses were windowed accordingly.

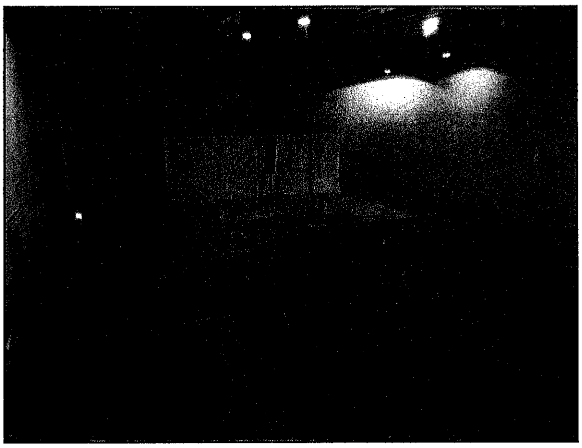

Figure 3: Setup in medium-size hall.

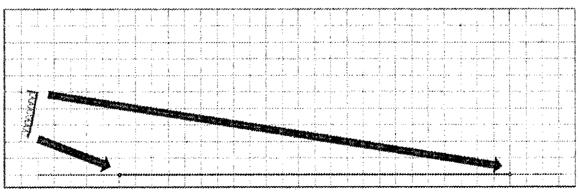

Figure 4: Software model.

The measurement results for the conventional configuration without FIR optimization are shown in Figs. 5 and 6. Fig. 5 shows the direct sound frequency response (left to right) from the front of the hall (bottom) to the back of the hall (top). The level at each position was smoothed to 1/12 octave bands and displayed using a color code relative to the average. The system exhibits the typical line array performance: The level decays smoothly over distance for low frequencies where the array is not large compared to the wavelength. In the mid-range the system shows typical beaming spots in the middle of the hall. In the high-frequency range side lobe structures appear, particularly notable close to the array. Fig. 6 displays the standard deviation of SPL across the entire audience zone.

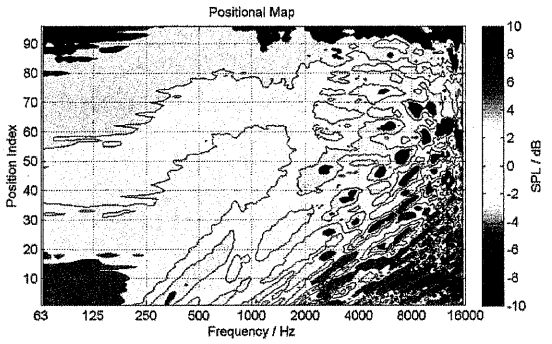

Figure 5: Positional map for conventional configuration.

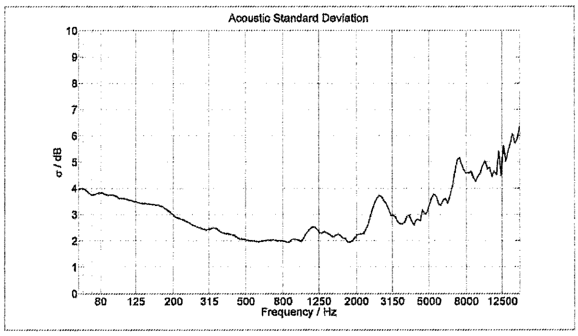

Figure 6: Standard deviation for conventional configuration.

FIR filters were calculated with the goal of optimizing the radiated sound field for good SPL coverage and maximum homogeneity. As Fig. 7 shows clearly, after applying these FIR filters the line array exhibits a very smooth frequency response over the largest part of the audience area. Only at very low frequencies, below 100 Hz, and at high frequencies, above 5 kHz, there is still measurable variation of SPL. In the LF range, the modeling accuracy and windowing effects in the measurements will limit the validity of the results. In the HF range, side lobe suppression especially close to the array is increasingly more difficult due to the spacing of independently controlled sound sources and due to small mechanical positioning deviations of the real system compared to the model. However, as displayed in Fig. 8, the standard deviation is reduced substantially from about 2.5 dB to about 1 dB in the range of 200 Hz to 3 kHz. The spread is notably reduced over the complete frequency bandwidth below 8 kHz.

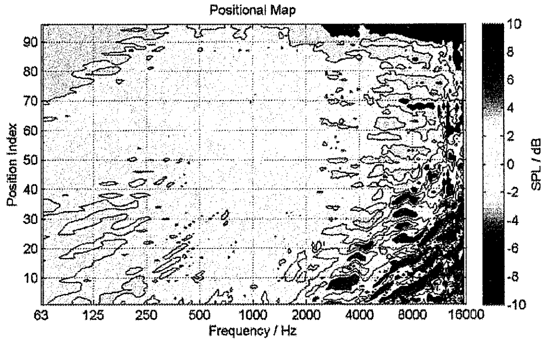

Figure 7: Positional map for FIR-optimized configuration.

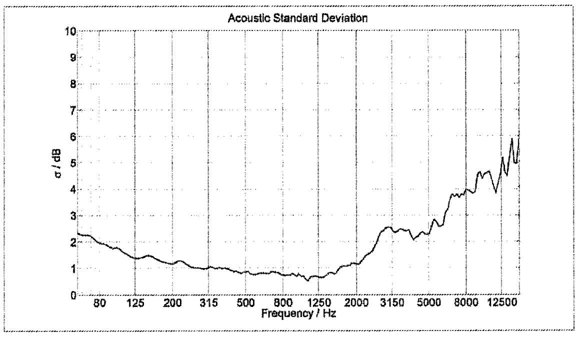

Figure 8: Standard deviation for FIR-optimized configuration.

3.2. Speech Intelligibility

A large-scale field test was conducted in an indoor sports arena, see Figs. 9 and 10. A high-performance line array system consisting of 16 cabinets was flown to cover a throw distance of about 70 m. Each cabinet used one channel of FIR processing with 1024 taps at 48 kHz sample rate. Ground-plane impulse response measurements were acquired in a resolution of 0.5 m along the axis of the array. The direct sound field for the conventional case is depicted in Figs. 11 and 12. The FIR-optimized version is shown in Figs. 13 and 14. Evidently, substantial improvements regarding the uniformity of the sound field could be achieved in the frequency range from 250 Hz to 4 kHz. Over this bandwidth the optimized standard deviation is 1 dB or even less compared to about 2.5 dB without optimization.

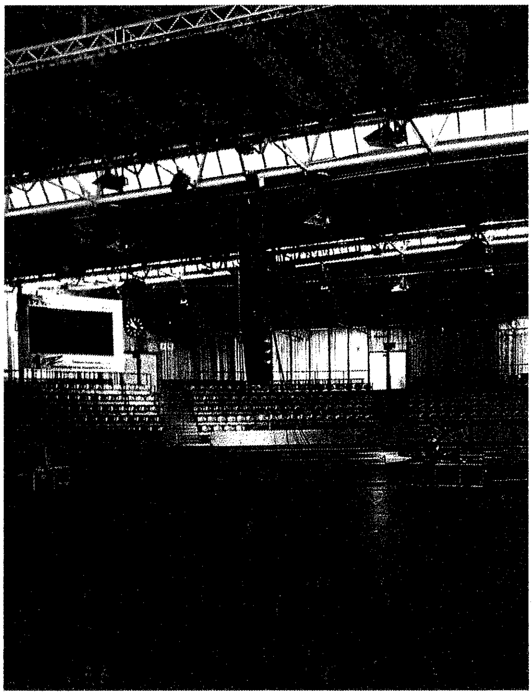

Figure 9: Measurement Setup.

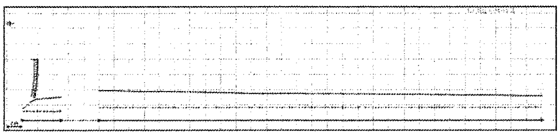

Figure 10: Model in EASE Focus 2.

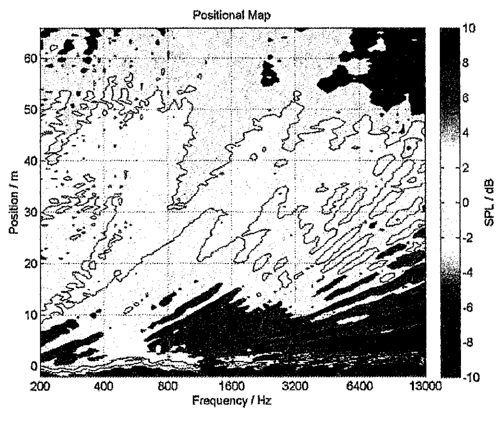

Figure 11: Positional map for conventional configuration, stage area: -2 m to +3 m, audience area: 8 m to 68 m.

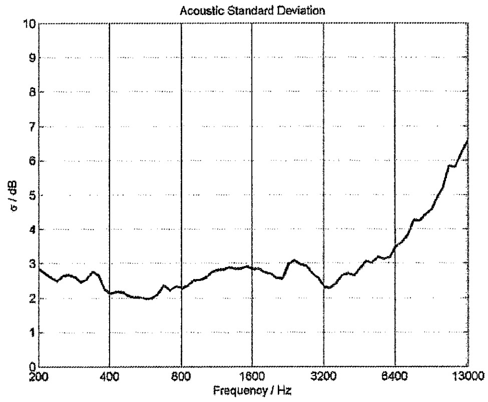

Figure 12: Standard deviation for conventional configuration, audience area.

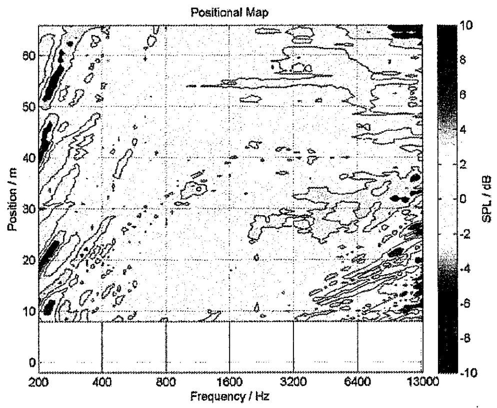

Figure 13: Positional map for FIR-optimized configuration, audience area: 8 m to 68 m.

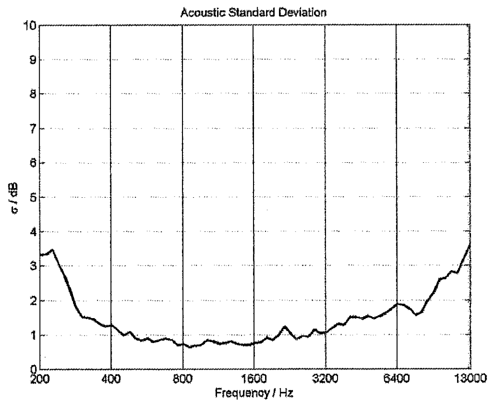

Figure 14: Standard deviation for FIR-optimized configuration, audience area.

The reverberation time of this stadium is about 2.5 s. It is therefore of immediate interest if the optimization can also improve speech intelligibility to a measurable degree. One would expect that better control of the direct sound field will result in improved control of the STI figures, as well. Fig. 15 shows the normalized direct SPL for the octave band of 1 kHz along the length of the hall. As indicated by the previous figures, the FIR filter optimization provides a much more consistent SPL. At the same time the uniformity of the STI is also improved, as shown in Fig. 16. Here a signal-to-noise ratio of 5 dB was assumed to simulate real conditions with crowd noise. Even though the average STI was reduced from about 0.585 to 0.576, the consistency throughout the audience area was improved significantly. The standard deviation dropped from 0.046 to 0.012 and therefore the conventionally used lower limit for the STI distribution increased from 0.538 to 0.564. Provisional subjective listening tests were performed for both settings and confirmed the findings.

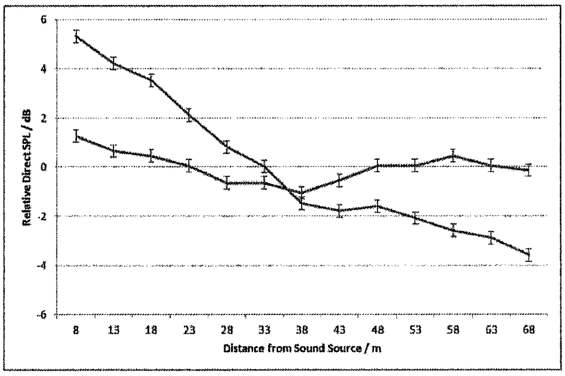

Figure 15: Direct SPL, 1 kHz octave band, conventional setup (blue) and FIR-optimized setup (red).

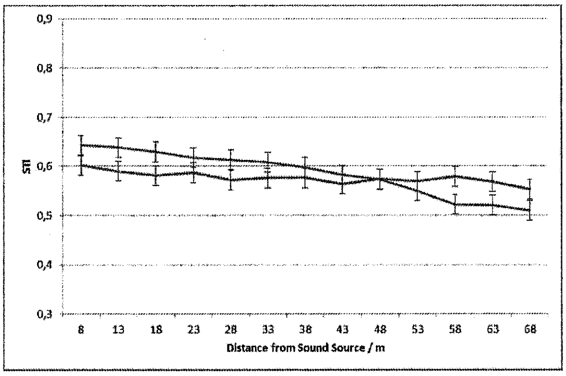

Figure 16: STI, conventional setup (blue) and FIR-optimized setup (red).

3.3. Coverage Control

Another optimization test was performed with the same setup as in section 3.2. A stage area of 5 m length was defined below the array and the optimization routine was configured so that sound radiation into this part of the room was suppressed with high priority, but without completely neglecting the goal to achieve uniform coverage on the audience area. Measurements were performed using a grid size of 0.1 m in the stage area, in addition to audience area measurements like in the previous case. The positional map for this setup is displayed in Fig. 17. Clearly, the SPL in the stage area (from -2 m to +3 m) is significantly reduced compared to the main audience zone. Fig. 18 shows the average frequency response of the stage area with and without FIR filter optimization. The level is reduced by about 10 dB over almost the full frequency range except for the side lobe at about 1.3 kHz which could not be changed that much. One can recognize in Fig. 17 how the optimization algorithm is compressing the bulk of the associated power into the zone between stage and audience (8 m to 68 m) in order to achieve both optimization goals as far as possible, namely, avoidance of irradiation into the stage area and spatially homogeneous coverage in the audience area. This emphasizes that when giving detailed coverage requirements the optimization results will always depend on the exact geometry of the venue as well as on the directional characteristics of the array elements.

These results have been obtained with a high priority on the exclusion of the stage zone. Consequentially, the performance in the listening zone has been compromised to measurable extent. But intermediate results with respect to the weighting between stage attenuation and SPL uniformity can be obtained as well. This example illustrates the trade-offs between different optimization goals. Eventually, the user will have to decide about the most appropriate solution.

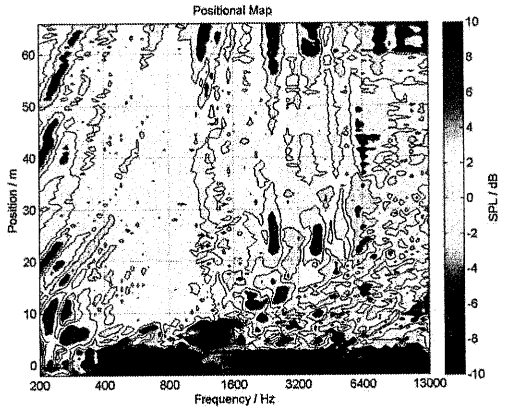

Figure 17: Positional map for FIR-optimized configuration, stage area: -2 m to +3 m, audience area: 8 m to 68 m.

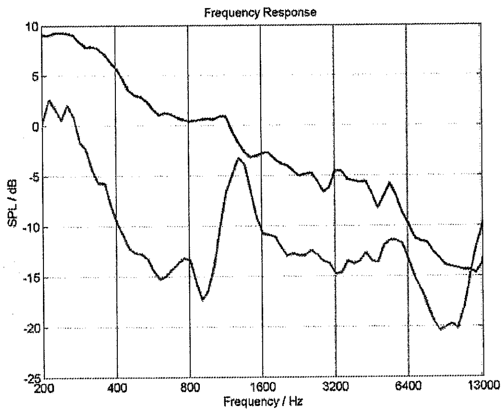

Figure 18: Average frequency response in stage area, for conventional (blue) and FIR-optimized (red) configurations, 1/3rd octave bandwidth.

4. IMPLEMENTATION ASPECTS

A number of different aspects must be considered when implementing FIR filter optimization in practice:

- Frequency limits are established by the modeling accuracy. At very low frequencies (approx. below 100 Hz) the point-source based approach employed here will not be as accurate due to wave-based interaction between array elements. At high frequencies (approx. 8 kHz and above) very small differences in mechanical positioning and aiming between model and actual setup will limit the modeling accuracy [7], [8].

- Direct limits for the steering ability are given by the length of the FIR filters. The tap count limits the control in the low-frequency range and also defines the maximum steering angle as a function of the length of the array.

- It is of high importance that the loudspeaker data is accurate and given in high resolution. Since the optimization approach relies on the precision of the modeling results, inadequate loudspeaker data will invalidate the numerical optimization [7], [8].

- Evidently, FIR filter processing has to be implemented accurately, as well. The optimal filter transfer function has to be reproduced by the implemented FIR with only very small error.

5. CONCLUSIONS

It was demonstrated that the radiation of sound by line arrays can be controlled accurately by numerical optimization of FIR filters. This powerful new technology offers a previously unknown level of design flexibility, coverage control, and system performance. Going beyond simple SPL improvements also speech intelligibility figures can be increased measurably. Nonetheless the results are still limited by the underlying laws of physics. This implies that the user has to weigh different design goals against each other. When optimizing FIR filters for sound systems one should be aware that the quality of the available modeling data and DSP hardware capabilities can influence the results.

Futures studies will include consideration of horizontal and 3D loudspeaker arrays as well as the combination of multiple arrays and their interaction. It will also be important to develop a better understanding of the interaction between loudspeaker directional characteristics, venue geometry and achievable FIR-optimized results.

6. ACKNOWLEDGEMENTS

We are grateful for the help provided by Anselm Goertz and Oliver Strauch from IFAA, Aachen, Germany. We would also like to thank d&b audiotechnik for providing measurement support.

7. REFERENCES

[1] EASE Software, EASE Focus Software, http://afmg.eu.

[2] G.W.J. van Beuningen, E.W. Start, Optimizing directivity properties of DSP controlled loudspeaker arrays, Reproduced Sound 16, Nov 2000.

[3] HK Audio, CAPS Aiming Software Manual, v3, Aug 2004.

[4] A. Thompson, Improved Methods for Controlling Touring Loudspeaker Arrays, 127th AES Convention, Oct 2009.

[5] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press (7th ed. 2009), http://www.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

[6] AFMG FIRmaker, http://firmaker.afmg.eu.

[7] S. Feistel, A. Thompson, W. Ahnert, Methods and Limitations of Line Source Simulation, JAES Volume 57 Issue 6 pp. 379-402; June 2009.

[8] S. Feistel, W. Ahnert, Modeling of Loudspeaker Systems Using High-Resolution Data, JAES Volume 55 Issue 7/8 pp. 571-597; July 2007.

[9] S. Feistel, R. Feistel, Coherent Superposition of Acoustic Sources as a Function of Environmental Parameters, 129th AES Convention, Nov 2010.

[10] S. Feistel, W. Ahnert, The Effect of Sample Variation Among Cabinets of a Line Array on Simulation Accuracy, 127th AES Convention, Oct. 2009.