PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

The EMPAC high-resolution modular loudspeaker array for wave field synthesis

Johannes GOEBEL[1]

Curtis R. Priem Experimental Media and Performing Arts Center (EMPAC)

Rensselaer Polytechnic Institute, USA

ABSTRACT

In order to explore the perceptual potential and limitations of Wave Field Synthesis (WFS) for artistic productions, a WFS array was built with loudspeakers placed as closely together as possible. Knowing from experience, that the high-frequency noisy components found in consonants of speech and transients of instrumental sounds are important for sound localization, the question arose if their coherent presentation through WFS with a higher spatial aliasing frequency affects the localization. The system was built with 5.8 cm between each loudspeaker resulting in an aliasing frequency of 2.9 kHz. Experimental configurations allow us to evaluate aliasing frequencies from below 700 Hz up to 5.8 kHz. Critical listening shows that increasing the number of speakers while decreasing their distance improves the perceived overall quality, including timbre. New modules with loudspeaker spaced 2.9 cm apart are under construction. The current system's 558 individually addressable loudspeakers are housed in 18 modules, each with 31 channels and an external subwoofer. The modular design allows for flexibility in deployment, including vertical and overhead configurations, enabling the simulation of complex acoustic environments and a diversity of artistic and research-driven projects. This paper is an invitation to use the system for research.

Keywords: Wave Field Synthesis, Spatial Aliasing Frequency, Focused Sound Sources

1. INTRODUCTION

The framework for Wave Field Synthesis (WFS) was developed in the 1980s by Berkhout, de Vries, and Vogel (1). A fundamental design criterion is that the distance between the system's loudspeakers should be as small as possible to synthesize a wave field that is acoustically accurate and true to a naturally occurring wave field. Spatial aliasing, a function of the distance between the centers of each loudspeaker in use, results in distortion imposed on the synthesized wave field in the higher frequencies of the virtual sound sources. By reducing the distance between the speakers, one extends the boundary of frequencies that are not impacted by spatial aliasing. This increased range results in a greater coherency of projected sound components. All frequencies above this spatial aliasing frequency are reproduced and audible to the listener; however, they inflict audible artifacts that counteract the construction of a coherent wave field.

With a few exceptions, the majority of WFS systems – especially portable systems and those used in performance environments – space their drivers between 10 and 15 cm apart, with each speaker in the array assigned a specific signal from a WFS rendering application. Systems with smaller distances are usually part of permanent installations. A system with, for instance, 4-inch drivers spaced at 12 cm has a spatial aliasing frequency around 1400 Hz. In a system with 8-inch drivers spaced at 24 cm, the spatial aliasing frequency is half that, at around 700 Hz.

The motivation for the EMPAC high-resolution modular loudspeaker array for wave field synthesis stems from artistic practice and creating new musical works that can only be realized with this technology. The discrepancy between what many existing WFS systems perceptually promise and the actual aural experience of an acute listener encouraged us to design a system that could verify the potential of WFS based on human perception as a point of departure. We believe scientific research could also benefit from this system. Its modularity, physical flexibility, and high spatial aliasing

[1] jeg@rpi.edu

frequency can allow for a real-world listening situation in a spatio-physical environment, where different spatial aliasing frequencies can be compared in situ on the same system.

2. CONTEXT OF THE EMPAC MODULAR WAVE FIELD SYNTHESIS ARRAY

Having experienced WFS speaker arrays that did not perceptually align with the effect described by the presenter or artist, and noting that critical frequencies for localization cues and timbral integrity were well above the aliasing points these systems could deliver, we decided to build a system that would allow us to find answers. The goal was to verify how the perception of a wave field was affected through a closer spacing of drivers and the resulting higher spatial aliasing frequency. For this, we used a technical framework that could provide a high channel count while allowing, at the same time, flexibility in the physical placement of the speaker arrays.

From artistic works using acoustic and electro-acoustic sounds in spaces of different sizes and qualities, we know through critical listening experiences that consonants in human speech, transients at the onset of instrumental sounds, and the carefully shaped attacks in electronic music provide localization cues. The timbral quality of sustained sounds is also greatly shaped by these higher frequencies. In order to determine how sound sources projected through wave field synthesis relate to listeners in complex environments, generation of a wave field that coherently integrates frequencies in a system suitable for performances and installations, with an aliasing frequency as high as possible was deemed essential.

WFS was initially developed for the reproduction of instrumental recordings to create an auditory field in which the localization of individual instruments is stable independent of the listeners' position. This included the auditory envelopment as if players and listeners were in a performance space. We see such "natural" reproduction, or rather re-synthesis, as a subset of any auditory environment that can be created with WFS, where the localization of static or moving sounds can be perceived independent of the listener's position within the physical space of the loudspeaker array. The understanding of the spatial and timbral conditions established by the physical parameters of a loudspeaker array is especially important when using sounds that are not imitating, modeling, or modifying acoustic instruments or timbres known to a listener, but which are "unheard of." With familiar vocal or instrumental sounds, the knowledge and memory of such sounds may make up for "non-fidelity" introduced by a wave field synthesis loudspeaker array. However, when creating sounds without a point of reference to generally known sounds, a musician may want to know what the resulting auditory quality will be if the music is played on arrays with wider or closer spacing of loudspeakers than the one she created the sounds on. This enables the composer to shape the details of sounds in such a way that their core characteristics can be either perceived through systems with different spacing or restrict the piece to a defined minimum loudspeaker spacing.

It was not our intention to limit ourselves to the creation of continuous wave fields in front of or around listeners, ones that simulate enveloping natural acoustic environments. Instead, we sought a method that allows the user, artist, or scientist to configure physically the arrays in many different ways. For example, by creating non-contiguous areas of speakers anywhere in a venue; combining a horizontally oriented array with another arranged vertically; placing arrays above an audience projecting downwards; or even positioning several "windows of sound" in different locations, on the walls and ceiling of a space.

Working with a variety of artists and researchers, we have been able to test different configurations using the EMPAC WFS array. Composers and researchers with extensive WFS programming and content creation on other systems, have noted that the linear array set-up of the EMPAC system provides a perceivable gain in quality over systems with larger distances between the drivers. Figure 1 shows a section of a linear configuration.

We also noted that halving the distance between the drivers and pushing the spatial aliasing frequency to 5.9 kHz results in a clearly perceivable increase of sonic quality and engagement versus the 5.9 cm distance at half the spatial aliasing frequency. We achieved 5.9 kHz by positioning two arrays on top of each other, offsetting half the on-center distance between drivers as shown in Figure 2, resulting in 2.9 cm between speakers. Whether this perceived increase of quality is due to the raised aliasing frequency, or by adding speakers in general (while maintaining the same level of loudness), still needs to be investigated.

Figure 1 – Linear configuration with 5.8 cm distance between drivers.

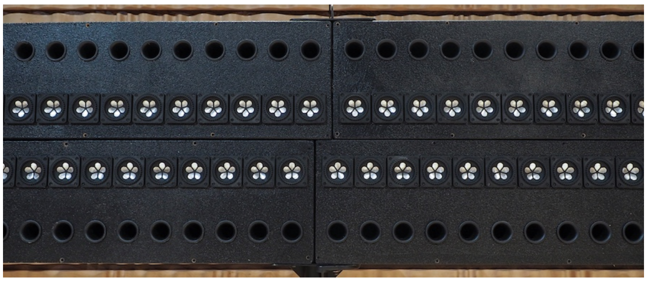

Figure 2 – Experimental set-up, halving the horizontal distance between drivers.

3. TECHNICAL SPECIFICATIONS

The array is comprised of 3 primary systems:

- 18 speaker modules with 31 drivers for WFS, and two 16-channel integrated DANTE power amplifiers (Attero Tech) that provide a 31.1 amplification platform, adding up to a total of 576 individually addressable channels, 558 for WFS plus 18 for the subwoofers.
- An uncompressed, low-latency, multi-channel audio network built on the DANTE protocol.
- IRCAM SPAT software [2] suite running on 2 Apple Mac Pros.

A small ancillary subwoofer array (Figure 3) is used to fill out the lower octaves below 140Hz, since this range is not realistically attainable with the WFS array, given the size of the driver and the geometry of its enclosure. There is one subwoofer associated with each WFS module. The signals for the subwoofers are sourced from the 32nd channel of the DANTE amplifiers residing on the back of the WFS enclosures (Figure 4).

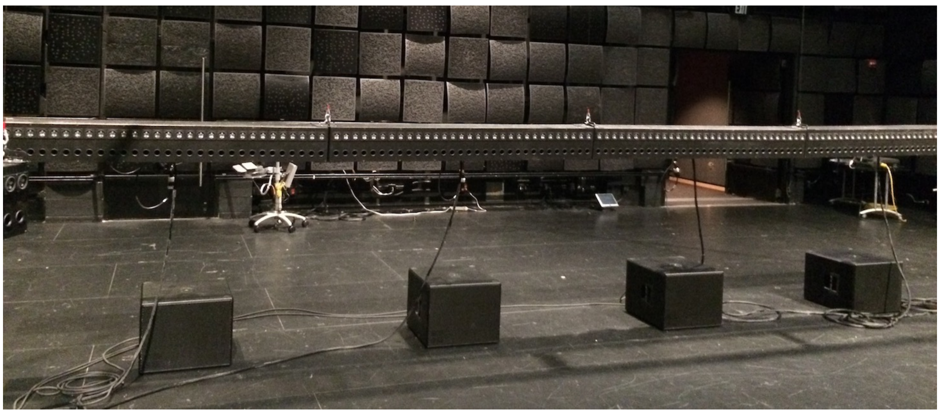

Figure 3 – Four modules of the EMPAC array with associated subwoofer per module.

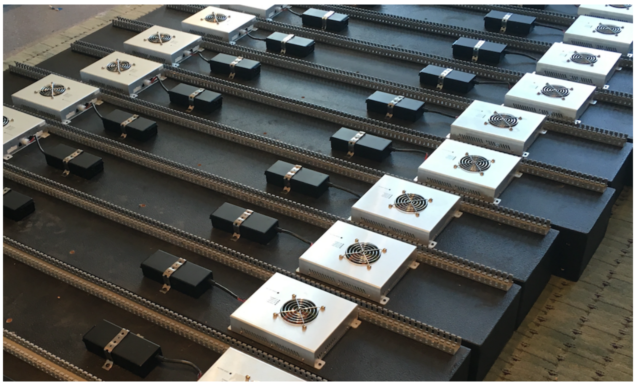

Figure 4 – Backside of array modules, with two 16-channel DANTE amplifiers and power supplies.

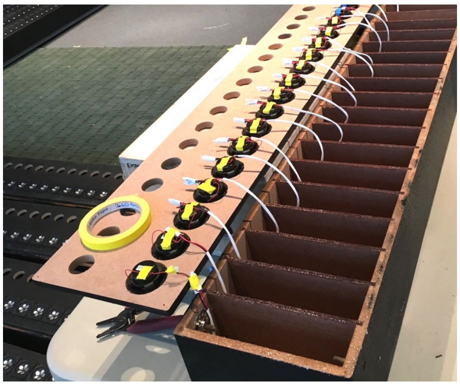

Figure 5 – Speaker installation, with openings for port tubes and internal subdivisions.

The array is comprised of eighteen 187 cm-long units, totaling around 33 meters in linear length. Each module is 22.8 cm high with a depth of 15.7 cm, not including the rear-mounted amplifiers and power supplies. Each unit is internally subdivided into 31 loudspeaker enclosures housing a ported 5.5 cm "full range" electrically and mechanically decoupled driver, 5.8 cm on center from its neighbor (Figure 5). These loudspeakers (Tang Band W2-2136S) act as individual 20-watt cells. Mounting and rigging assemblies at the ends of each module allow for a high degree of flexibility in their configuration and placement (Figure 6 and 7).

The modular approach allows for arbitrary placement of modules. So far, each array set-up contiguously along a straight line has been calculated independent of array sections positioned at a differing angle. Synthesizing virtual focused sound sources or a unified reverberation environment within a space circumscribed by rectangular or circular array lines, where all jointly contribute to the synthesis of a unified sound field, will require further development.

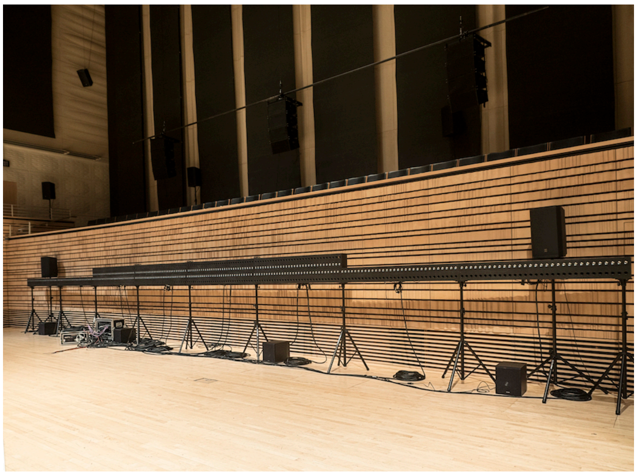

Figure 6 – Configuration as linear array with added inverted array in center portion to reduce distance between speakers to 2.9 cm.

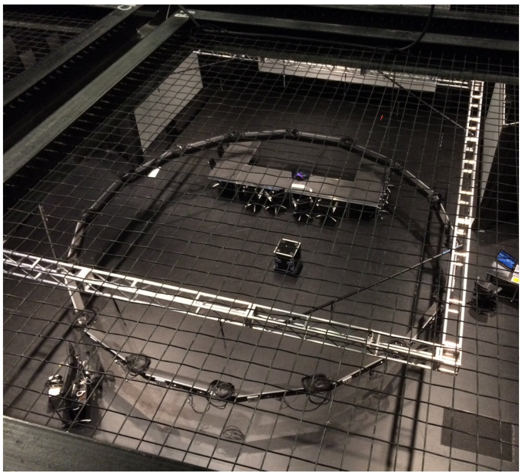

Figure 7 – Circular configuration, suspended from trusses (viewed from above).

Along with generous support in the development, adaptation and testing from IRCAM, Paris, IRCAM's SPAT real-time processor toolset computes EMPAC's WFS render within Cycling 74's MAX/MSP on several Apple MacPros. The number of computers used expands and contracts with the number of discrete arrays rendered. Currently, DANTE's PCIe driver architecture imposes a 256 channel hardware limit per machine. There is a dedicated OSC network that provides external access and control of SPAT. MAX/MSP, IRCAM's TOSCA and Hexlers's Touch OSC are the dominant external controllers used at this time.

4. EVALUATION THROUGH LISTENING IN MUSICAL CONTEXT

EMPAC built the WFS array as an instrument for artistic, acoustic and perceptional inquiries and projects aiming for the highest possible quality hearing can detect. A scientific validation with appropriate tests and statistical evaluation is welcome and we invite researchers to use this physical system to also verify experimental results they have deduced from binaural renderings of computational models of loudspeaker arrays for wave field synthesis.

The EMPAC system was built to understand and potentially prove how the closer spacing of speakers and resulting higher spatial aliasing frequency would affect the perception and use of such systems. In comparing different sounds and music with different array configurations, two aspects could be experienced and confirmed in dramatic ways.

The long linear array with speakers spaced at 5.8 cm on-center, as shown in Figure 1, provides superbly stable sound localization without any strong spatial aliasing. But once two arrays were set on top of each other (Figure 2 and 6) and the distance on-center was halved to 2.9 cm, the quality increased in immediately perceivable ways. By this comparison, it became clearly apparent how much the extension of the spatial aliasing frequency beyond 1.5 kHz enhanced the quality of the projected wave field and that it perceivably improved the system's fidelity.

The most radical results emanate when localizing sounds in front of the array. It has been demonstrated that WFS is the only system through which sounds can be placed and heard in an area in front of the speakers, as long as listeners do not move into the area between the array and the sound's projected placement. With closer-spaced speakers at 5.8 cm, and even more so at 2.9 cm, front-focused sound tests of prerecorded instrumental music projected at distances up to 8 m off the front of the array rendered a "stunning" experience.

4.1 Recording and Projection of Instrumental Music

W.A. Mozart's String Quartet No.3 in G major, KV156, as well as his Serenade in E-flat major, KV 375, for eight wind instruments were recorded in EMPAC's concert hall. Each piece was recorded as an ensemble, with each instrument closely miked, and with additional room microphones capturing the ambience of the hall. After the initial recording, each instrument was recorded solo, with the full recording serving the instrumentalist as a synchronization reference over headphones. A video recording of the full wind ensemble was played back for the individual instrumentalists so they could pick up the visual cues from the other players. In contrast to playing with a click-track, this allowed for freely flowing agogical deviations from a strict meter, important for music of this genre and period.

The wind octet was rendered through the WFS array on the EMPAC theater stage, with 8 modules (248 channels) creating a linear array of almost 15 m in length along the upstage wall. The virtual sound sources for the instruments were projected out onto the stage, with each instrument placed in the area in front of the array (the area was 8 meters in depth and not quite as wide as the array itself). Figure 8 shows where their locations were marked on the floor, so one could move between those spots and experience being right next to, or even "in" the specific instrument, while still hearing all other instruments in their respective locations under the restriction that precise localization of focused sound sources is not possible when one is closer to the array than the spot of the sound source. For example, in Figure 8, standing in the location of Clarinet 1 will result in the diffuse localization of Oboe 1 and 2.

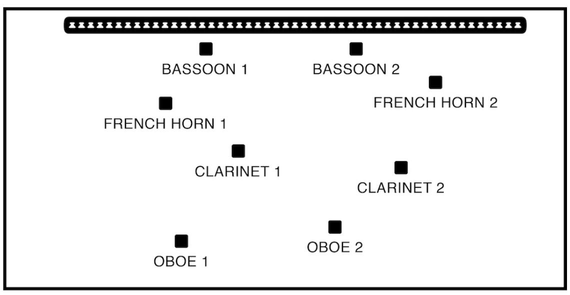

Figure 8 – Stage lay-out for Mozart Octet with 248 channels, 15m long, and virtual locations of instruments marked on floor.

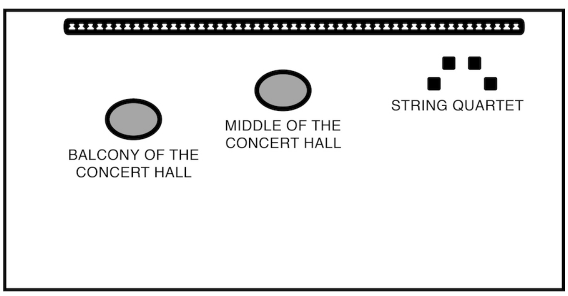

Figure 9 – Mozart Quartet with virtual locations of strings and the two sonic portals.

Sitting in the audience chamber in front of the stage with the instruments placed across the stage, as if musicians were sitting there, yielded an extraordinary, acoustically transparent experience.

The projection of the string quartet recording used the same set-up of the WFS array on the theater stage, but this time in a less reverberant listening space with curtains blocking the audience chamber and a carpet covering the floor. Figure 9 shows the three areas for focused sound sources. The string quartet was virtually placed to the far right of the system, the four instruments set in a half circle in front of the array, with the distance between the instruments equivalent to a "real" string quartet. A recording taken with stereo room microphones in the middle of the concert hall was focused on an area in front of the center of the array, with the two audio channels as individual virtual sound sources. To the left end of the array, the stereo room microphones from the balcony of the concert hall created another superimposed acoustic field.

A listener could go directly to the string quartet's virtual position, stand in front of the ensemble and listen to the group, then "stick one's head" right into the individual positions of the two violins, the viola, and cello, while still hearing the other instruments in their respective location. While the music played and the focused sound sources of the quartet players stayed in their spots, one could move to the center of the array and enter, through a "sonic portal," the middle of the concert hall. Finally, one could move to the left and find oneself in the further afield ambient space of the balcony. So listeners could move between the three different sound fields, entering different acoustic spaces, and experiencing the very different acoustic environments in relation to the string quartet, which "stayed in its spot."

4.2 Electronic Music, Stereo Recordings, Music for Theater, and Field Recordings

In a third scenario, a piece of electro-acoustic, acousmatic music produced in stereo was played over the WFS array, with each track focused in front of the array. Again, lthe istening experience while sitting in the audience chamber, yielded a spatial, transparent, and stable sound stage, which would have been impossible to obtain with a traditional left-right speaker configuration. A similar positive experience was made when playing the stereo recording of a chamber ensemble through the WFS array set up on stage of EMPAC's concert hall. Further listening experiments were conducted placing stereo channels behind and in front of the WFS array and comparing studio productions of Rock and Blues music with concert recordings of medieval choir music made in a reverberant church. The approach of playing stereo or multichannel recordings through a WFS array needs further investigation into which kind of stereo productions benefit from being played back through WFS systems.

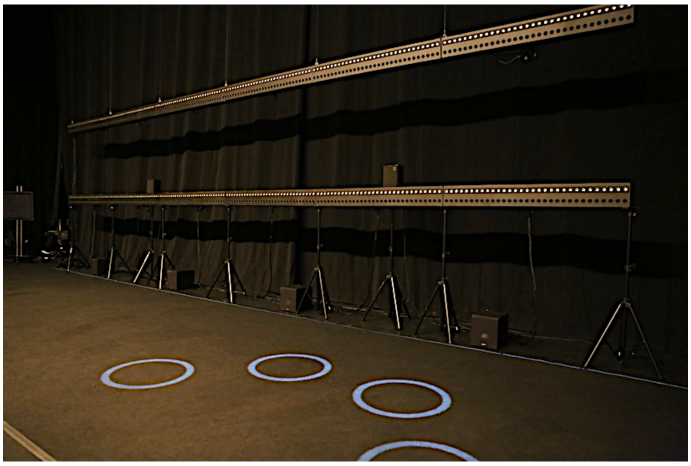

Figure 10 – Two arrays at larger vertical off-set. Placement of focused sound sources the lower array were marked by projected circles on the floor, which also moved as the sound positions were moved.

Figure 11 – Arrays flown above audience.

Experiments were also conducted to gain experience with two arrays positioned at different heights, as shown in Figure 10. Combining two horizontal arrays at a vertical distance yielded a more enveloping perception and allowed for virtual sound sources at different elevations. In this case, each array was treated separately, with the signal processing for one array not being related to the other.

Additionally, flying WFS arrays overhead of listeners has been explored. In 2017, while in residency at EMPAC, the sound designer Bobby McElver collaborated with the performer Andrew Schneider on a new theater piece by using two parallel arrays positioned above the audience to move sounds highly successfully through the audience. He subsequently built a copy of the EMPAC system for touring the work. – In 2019, two arrays in combination with a 7^{th} order Ambisonic system were hung in the EMPAC Concert Hall for presentations by Chris Watson and Tony Myatt (Figure 11), performing underwater field recordings from around the globe and from an elephant habitat (the center stage subwoofer being constructed for sub-audio elephant communication).

5. FUTURE DEVELOPMENT

As addressed before, our primary goal for using WFS is not to recreate the "as if" experience of a concert hall or to recreate a "natural" acoustic environment, but to use it as an instrument in artistic works. The listening tests with instrumental recordings have allowed the evaluation of the wave field's precision, as we can compare it to our experience with acoustic instruments in performance spaces and thus allow veracity assessment of the rendered wave field in relation to perception.

On the application side, a tracking system for persons on stage will be integrated with the high-resolution WFS system as stage reinforcement option. A WFS array placed along a stage lip or at an upstage wall will map sound to a performer's position. This will allow a horizontal alignment of a performer and the amplified reinforcement independent of the listeners' position. For the spatial alignment of the word spoken by a person on stage with its amplification, a higher aliasing frequency seems to be essential.

An important next step will be to include algorithms to treat distributed sets of arrays as a unified instrument, so each array, wherever placed and however contiguous or not, takes the other arrays into account and is used as a contributing factor to the sound sources created by all arrays.

A second generation of the modules with a higher density of speakers and a spatial aliasing frequency of 5.9 kHz is under development. Figure 12 shows a rendering of the module, without drivers and ports installed. The two inner rows will hold the 64 speakers and the outer rows will hold the ports. The offsets at each end allow the interlocking of modules without affecting the distance between drivers where two modules meet. One module will be 24.7 cm high, 22.8cm deep, and 188.5 cm long, with DANTE integrated power amplifiers on the back.

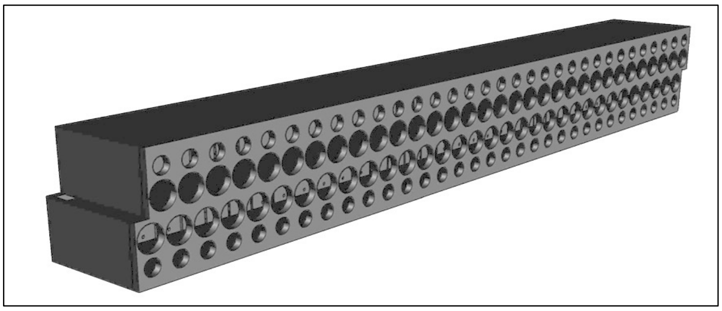

Figure 12 – Rendering of 64-channel module, 2.9 cm distance between drivers.

6. CONCLUSION

The construction of a portable, modular WFS loudspeaker array allows for the deployment of highly flexible configurations, for research and for new artistic works, including flying arrays above listeners. Listening comparisons using complex sounds in musical context made the difference in spatial and timbral fidelity between speakers at 5.8 cm distance and 2.9 cm perceivable; whether this is a result of higher aliasing frequency, needs further investigation. The positioning of focused sound sources was tested with exceptional results up to 8 m and more in front of a linear array. Composers and researchers who had worked with other WFS systems with lower spatial aliasing frequencies testified to the new quality reached on account of the closer placement of the drivers. Researchers are welcome to use the system for further investigations.

ACKNOWLEDGEMENTS

The author would like to thank Markus Noisternig/IRCAM for his continuous illuminations and theoretical input; Todd Vos/EMPAC, for the array's systems definitions, design, construction, and integration; Jeff Svatek/EMPAC for networking and applications services. Argeo Ascani/EMPAC contributed to the programmatic development and presentation of the system.

REFERENCES

1. Berkhout A J, de Vries D, P. Vogel P. Acoustic control by wave field synthesis. The Journal of the Acoustical Society of America 93, 2764 (1993); doi: 10.1121/1.405852
2. http://forumnet.ircam.fr/product/spat-en/