Digital Loudspeaker Arrays driven by 1-bit signals

Nicolas – Alexander Tatlas and John Mourjopoulos

Audiogroup, Electrical Engineering and Computer Engineering Department, University of Patras, Patras, 26500 Greece
ntatlas@wcl.ee.upatras.gr, mourjop@wcl.ee.upatras.gr

ABSTRACT
Loudspeaker Arrays driven by digital bitstreams are direct digital-signal to acoustic transducers, usually comprising of a digital signal processing module driving actuators. Current research efforts are focusing on topologies directly driven by multi-bit digital bitstreams. In this work, the above investigations are extended to the case of using 1-bit signals such as Sigma-Delta for driving such topologies, using time and frequency domain analysis. Simulation results will be presented for idealized actuators. Finally, an optimized architecture for such a loudspeaker will be proposed, based on this analysis.

0. INTRODUCTION

A “Digital Loudspeaker Array” (DLA) [1] is a direct digital-signal to acoustic transducer, typically comprising of a digital signal processing module driving self-powered miniature elements, strategically positioned in order to optimally reproduce the digital audio data stream directly.

As seen in Figure 1, generally a DLA [4] consists of three stages:
- digital signal processing (DSP)
- digital audio amplification (DAMP)
- digital acoustic emission (DAE)

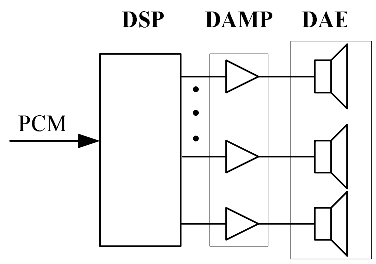


Figure 1 Digital Loudspeaker Array structure

The main advantage of such a loudspeaker is that the signal remains in the digital domain and is converted to analog through the element-to-air coupling. Hence, the complete audio reproduction chain can be easily integrated into compact and efficient components. Other advantages of such digital arrays, relate to the flexible control of their directivity [3], a feature which has been recently utilized for multichannel reproduction via a single array [5], though such aspects are not addressed here.

This work introduces novel results related to the potential implementation of digital loudspeaker arrays, fed by one-bit signals, typically based on Sigma-Delta modulation (SDM), as employed in the DSD format. Hence, results previously presented for PCM signals [2] are extended here, allowing useful conclusions to be drawn on the respective merits and disadvantages of multi bit and single bit formats for such applications.

Simulation results for the 1-bit digital loudspeaker arrays will be presented for two topologies: (a) linear array and (b) two-dimensional scheme, in both cases using idealized actuator elements.

The paper is organized as following:

In section 1.1, the signal pre-processing stage necessary for 1-bit loudspeakers is analyzed, where a novel scheme for grouping Sigma-Delta modulated signals into frames is proposed. In section 1.2, the time-domain and frequency-domain representation algorithms for bit-grouped signals are presented. In section 2 the simulation parameters, such as array topology, are considered, while in section 3 the simulation results for different factors are presented. Finally, in section 4 a structure for implementing a DLA is proposed and the conclusions of this work are summarized.

1. THEORY

1.1. Signal Pre-processing

The 1-bit Digital Loudspeaker Array can receive PCM digital audio samples and pre-process them to an appropriate 1-bit signal for the direct transduction. Alternatively, a DSD stream may be fed in the DLA.

The input PCM signal s is represented here as a $N \times M$ matrix, where N is the number of bits per sample and M is the total number of samples (in theory infinitely large), e.g. N=16 and M=88200.

$$s = \begin{bmatrix} b_{1,1} & . & . & b_{1,m} \\ b_{2,1} & . & . & b_{2,m} \\ . & . & . & . \\ b_{n,1} & . & . & b_{n,m} \end{bmatrix} \quad (1)$$

The signal s, through the Sigma-Delta modulation process *SDM{s}* is oversampled R times, noise-shaped and re-quantized, thus providing the bitstream

$$s_D = SDM\{s\} = \begin{bmatrix} b_{1,1} & . & . & b_{n',1} & b_{1,2} & . & . & b_{n',m} \end{bmatrix} \quad (2)$$

represented as a single-row matrix, with M' columns where $M' = R \cdot M$.

Here, a novel pre-processing conversion stage is proposed for mapping the SDM stream to the array elements. It is based on grouping Sigma-Delta samples, by creating Sigma-Delta frames of length E (bits). The variable E is directly related to the number of array actuators to be used. So, the grouped *G{}* signal s' is derived, where

$$s' = G\{SDM\{s\}\} = \begin{bmatrix} b_{1,1} & . & . & b_{1,m} \\ b_{2,1} & . & . & b_{2,m} \\ . & . & . & . \\ b_{n',1} & . & . & b_{n',m} \end{bmatrix} \quad (3)$$

is a $N' \times M$ matrix with $N' = E$. Each column of s' is defined as a Sigma-Delta "frame". From this grouping, separate bitstreams, s'_1 to $s'_{N'}$, are created. Each of these bitstream groups is then sent to a specific actuator element, as seen in Figure 2.

1.2. Digital Acoustic Emission Simulation

In order to evaluate the sound pressure waveform derived from the DAE, for a given position in space, two alternative approaches have been employed: (i) time domain and (ii) frequency domain simulation. In principle, both methods are equivalent and generate similar results, under some specific constraints that are analyzed below.

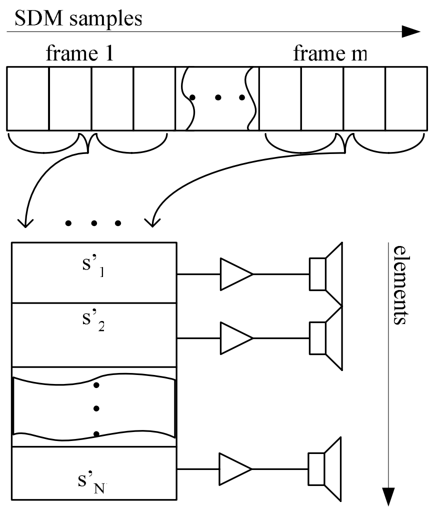


Figure 2 Bitstream grouping and emission

1.2.1. Time Domain Simulation

Considering that each DAE element produces a sound pressure level C (N/m) at 1m distance, then the pressure at a distance r_k is:

$$P_k(n) = \frac{C}{r_k} s'_k(n - n_k) \quad (8)$$

where s'_k is the binary input signal to the actuator and n_k is the time delay in samples given by:

$$n_k = \left\lfloor \frac{cf_r}{r_k} \right\rfloor \quad (9)$$

where c (m/s) is the speed of sound, f_r (Hz) the input signal sampling rate and $\lfloor\ \rfloor$ denotes floor integer truncation.

The total sound pressure produced can be evaluated by adding the contribution of all E elements, so:

$$P(n) = \sum_{k=1}^{E} \frac{C}{r_k} s'_k (n - n_k) \quad (10)$$

For any given input test signal period of M samples, the number n_{total} of samples forming the sound pressure under observation would be equal to:

$$n_{total} = M + n_{k\max} - 1 \quad (11)$$

where $n_{k\max}$ the maximum delay in samples. Since no elements affect the generated pressure in the interval $n = [0, n_{k\min} - 1]$, where $n_{k\min}$ the minimum delay (i.e. prior to the arrival of the sound), P(n) is zero at this interval. Furthermore, the last $(n_{k\max} - n_{k\min})$ samples must not be taken under consideration since the total pressure observed will result from the contribution of the elements farther away, while the elements closer to the observation point will not emit since the input signal to the DLA will be halted.

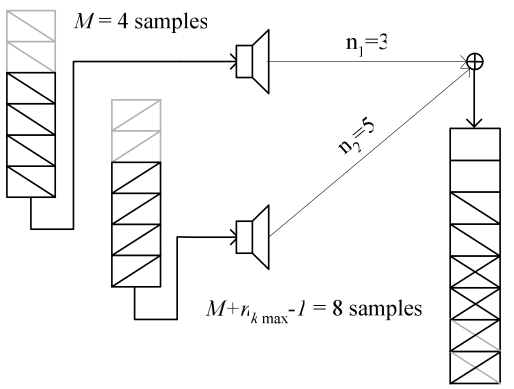


Figure 3 Time domain simulation

For example, in Figure 3, the last two samples in the generated pressure (*5-3=2*) contain only the contribution from element 2.

1.2.2.Frequency Domain Simulation

For the frequency domain simulation, the spectrum for each bitstream sent to an element over whole number of periods of M samples is considered. If the Fourier transform of the input signal is

$$S'_k (e^{j\omega}) = \sum_{n=0}^{M} s'_k (n) e^{-j\omega n} \quad (12)$$

then the sound pressure contribution of one DAE element at a distance r_k is derived as:

$$P_k(n) = C \sum_{m=1}^{K} \frac{|S'_k|}{r_k} \cos\left[m\omega(n - n_k) + \arg(S'_k) \right] \quad (13)$$

where $|S'_k|$ and $\arg(S'_k)$ are the magnitudes and phases of $S'_k(e^{j\omega})$ and K is chosen so that all frequencies of the bitstream series are included.

As with the time-domain method, the total sound pressure produced can be found by adding the contribution to sound pressure P_k for all E elements, thus:

$$P(n) = C \sum_{k=1}^{E} \sum_{m=1}^{K} \frac{|S'_k|}{r_k} \cos\left[m\omega(n - n_k) + \arg(S'_k) \right] \quad (14)$$

2. IMPLEMENTATION

For the tests, the 1-bit Sigma-Delta modulated signal was generated using a 5th order Noise Transfer Function (NTF), while the oversampling ratio R was 64 and the original sampling rate was f_s=44.1KHz. Thus, the SDM signal s_D has a sampling rate of aprox. 2.8MHz and conforms to the DSD format, as seen in Figure 4.

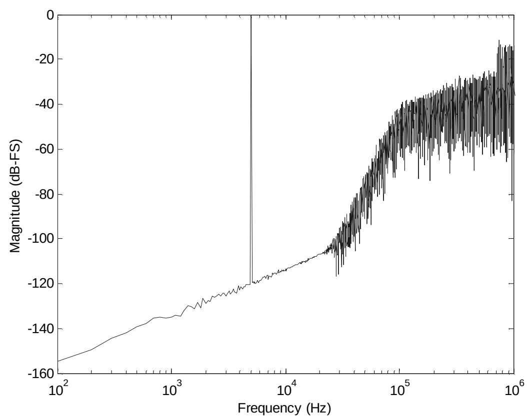


Figure 4 SDM signal spectrum

The acoustic transform elements are assumed here to have an idealized impulse response (e.g. equal to the dirac function) and to be omnidirectional. Each element's diameter is fixed at 1cm. Two topologies have been considered: (a) linear array where the actuator elements are placed uniformly in line as seen in Figure 5 and (b) two-dimensional scheme, where the actuators are mounted on a surface as seen in Figure 6.

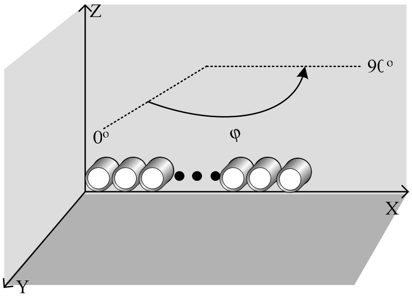


Figure 5 Linear DAE scheme

The elements are able to reproduce either a positive or a negative (out of phase) pulse sound pressure. When a "0" bit is fed to the element a negative pressure is produced, while when a "1" bit is fed, a positive pressure is emitted. Low-pass filtering is employed before the DAE stage, in order to overcome spatial aliasing issues. The filter used is 8th order, with a cut-off frequency of 22KHz.

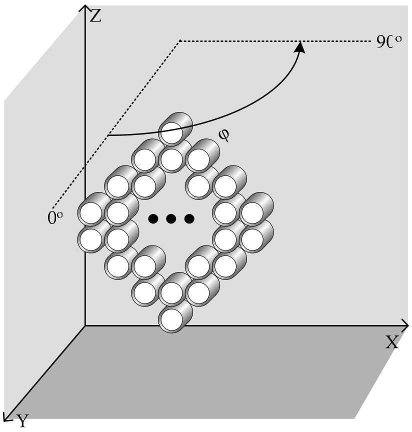


Figure 6 2-D DAE scheme

3. RESULTS

Given that both frequency and time domain simulation methods produce identical results, the frequency domain approach was utilized for the results presented. The audio performance of the system was measured in terms of the DLA Sound Pressure Level for different angles of observation (i.e. its polar response) as well as the resulting Total Harmonic Distortion plus Noise (THD+N) for different input frequencies.

The simulation parameters for both topologies are:

- input digital sine waves frequency f_{in}, varying from 250Hz to 10000Hz, 0dB-FS amplitude, 20 whole periods each
- distance from the center array element l, varying from 2m to 5m
- angle of observation φ, varying from 0^o to 90^o

Additionally, for both arrays implemented, as seen in Figures 5 and 6, the height *(Z)* for the observation point was fixed at 0,5 m.

3.1. Bit Grouping

The bit grouping method provides "frames" of Sigma – Delta bitstreams to the DAMP/DAE module for direct emission. The spectrum of the resulting grouped signal, for a 5 KHz sine wave input can be seen in Figure 7. The THD+N for all frequencies considered is approximately 0.10%. As will be discussed later, this inhibits the DLA's performance. Hence, the DSD bitstream-to-element allocation scheme may be further optimized.

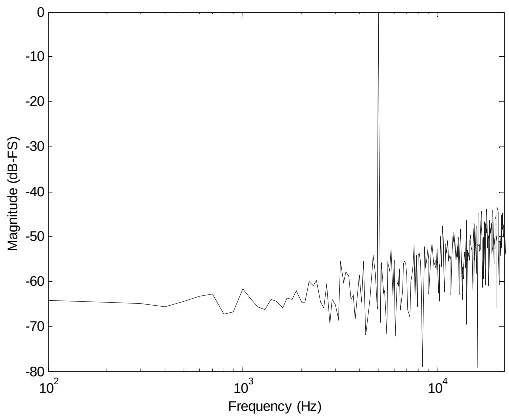


Figure 7 SDM grouped signal spectrum, $f_{in}=5KHz$

3.2. Pressure waveform

An example for the waveform produced for both DAE schemes, for a 5KHz sine wave input at a distance of l=2m and an angle of $\varphi=30^o$ can be seen in Figures 8 and 9. Clearly, the sound pressure produced is an approximation of the original DSD sinewave, while the distortion induced by time-delay errors is evident especially for the case of the Linear DLA.

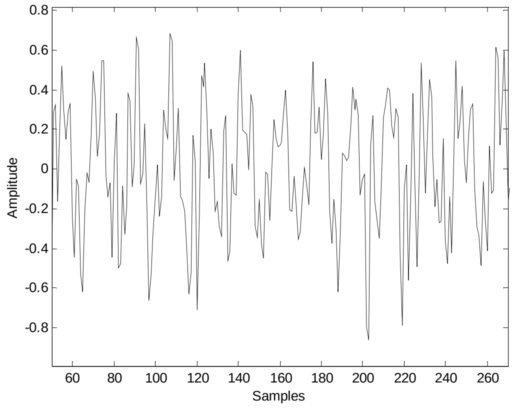


Figure 8 Pressure waveform, Linear DLA
f_{in}=5KHz, *l=2m*, φ=30°

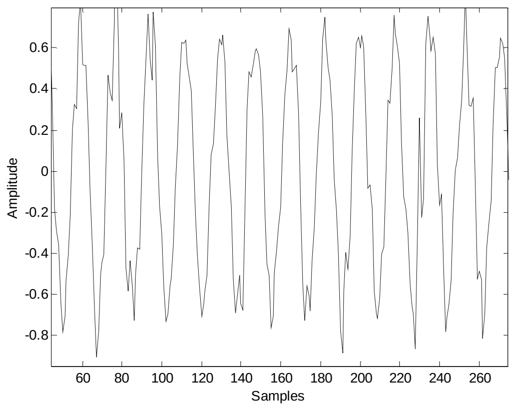


Figure 9 Pressure waveform, 2-D DLA
f_{in}=5KHz, *l=2m*, φ=30°

3.3. Pressure spectrum

The spectrum corresponding to the above two waveforms can be seen in Figures 10 and 11. The THD+N for the Linear DLA is evaluated at 47.0%, while for the 2-D DLA 4.3%. Since the position of sound pressure observation is identical for both DLA setups, the above THD+N results lead to the conclusion that the DAEs geometry is critical in order to achieve proper sound reconstruction.

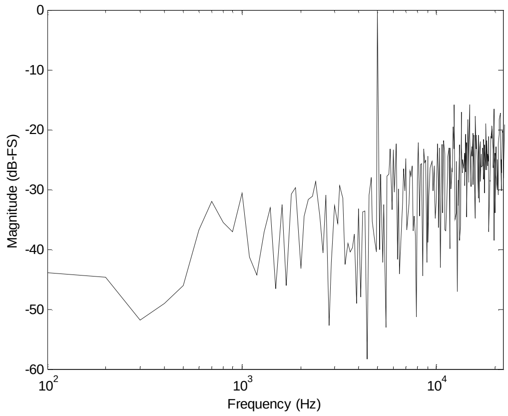


Figure 10 Pressure spectrum, Linear DLA
f_{in}=5KHz, l=2m, φ=30°

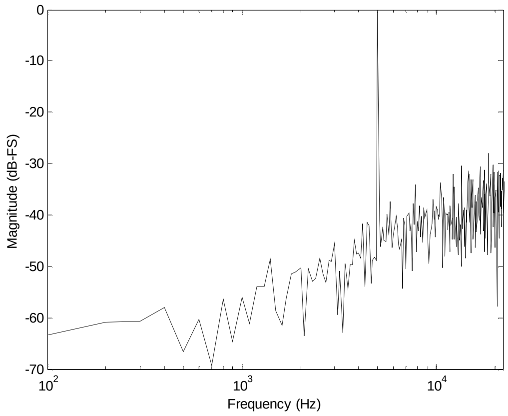


Figure 11 Pressure spectrum, 2-D DLA
f_{in}=5KHz, *l=2m*, φ=30°

3.4. Directivity

As with conventional loudspeaker arrays, the directional effect of a DLA, is primarily caused by the phase differences of the contributions from each source, as a result of different path lengths for different positions [6]. However, the digital nature of the signals emitted directly affects the sound field produced. As expected, the directional characteristics of the linear array are stronger than that of the 2-D array as the DAE elements in the latter case have nearly the same path length for a given position.

The ratio of the sound pressure level observed at each point to the maximum pressure for each frequency is depicted in Figures 12 (a) to (d) where the directivity polar diagrams for the Linear Digital Loudspeaker Array are plotted, while Figures 13 (a) to (d) show corresponding plots for the 2-D DLA.

For both DAE geometries, a strong dependency of the sound field produced to the input frequency is observed. For lower frequencies, approx. up to 500 Hz, the two schemes present almost omnidirectional patterns since the path differences are not significant compared to the sound wavelengths. However, while the 2-D DLA exhibits an omnidirectional pattern for frequencies up to 2 KHz, the linear array from 1 KHz presents patterns that converge to a narrow on-axis lobe.

3.5. Harmonic Distortion

Figure 14 shows the % THD+N vs frequency for the Linear DLA as for $\varphi=60^o$ in (a) and $\varphi=30^o$ in (b), as a function of distance. Figure 15 (a) and (b) show similar results for the case of the 2-D DLA. Figures 16 and 17 display corresponding results for the two array schemes for a distance of l=3m, as a function of angle φ. These results lead to the following conclusions:

- THD+N is largely independent from distance. This verifies that the spatial quantization applied does not introduce a significant error to the approximation of the sound field produced.
- THD+N depends heavily on the receiver's position. Regardless of any other parameter, distortion values are minimal for the on-axis position while an increasing trend is observed for increasing off-axis angles. Acceptable reproduction for the whole frequency band under study can be achieved for the 2-D DLA, around on-axis positions.
- THD+N generally increases with frequency. However in the case of the 2-D DLA and for frequencies below 2 KHz, acceptable reproduction can be achieved regardless of the angle of observation.
- THD+N depends on the DAE architecture. The linear DAE reconstructs sound waveforms with a THD+N of over 20% for almost all off-axis positions and regardless of frequency, as seen in Figure 15. On the other hand, the THD+N for the 2-D DAE is greatly improved, as seen in Figure 17.

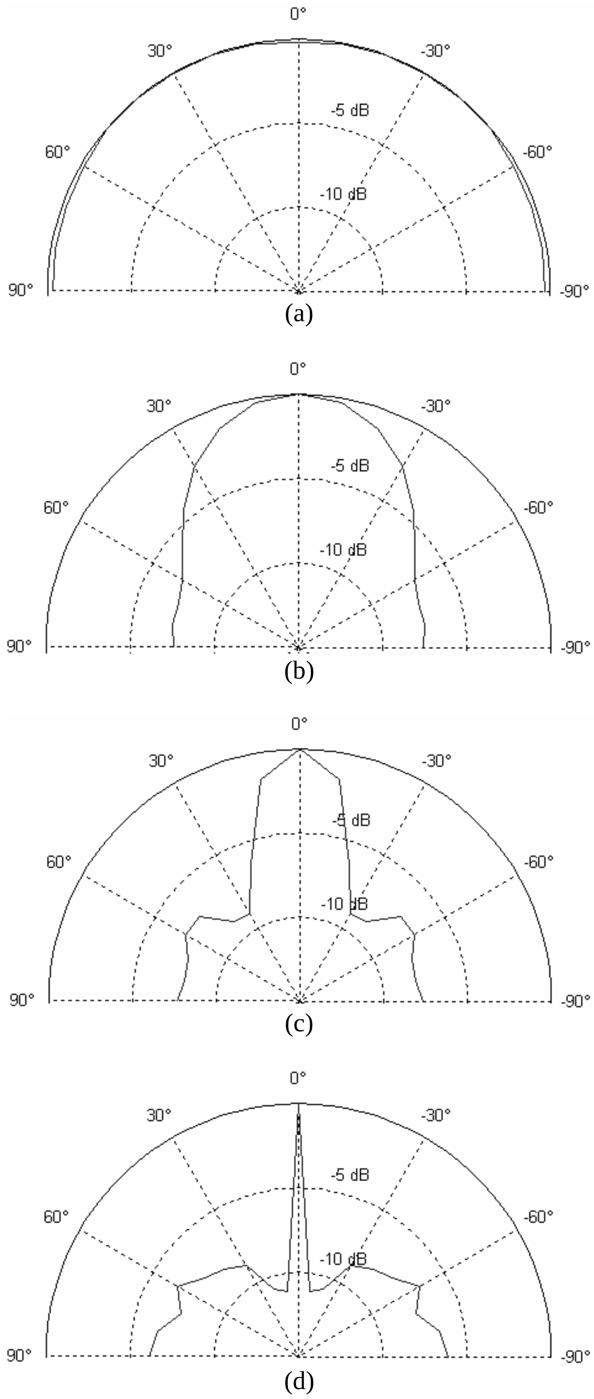
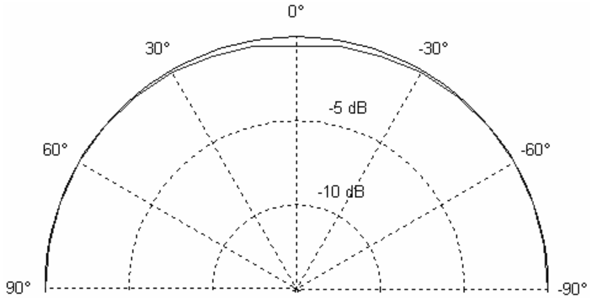
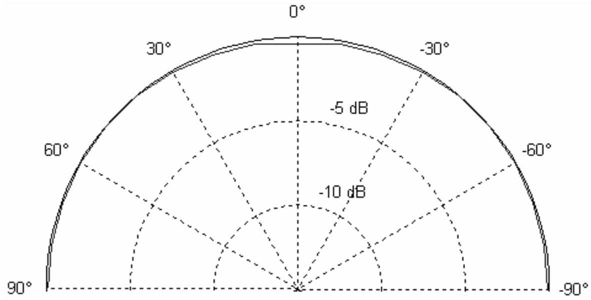


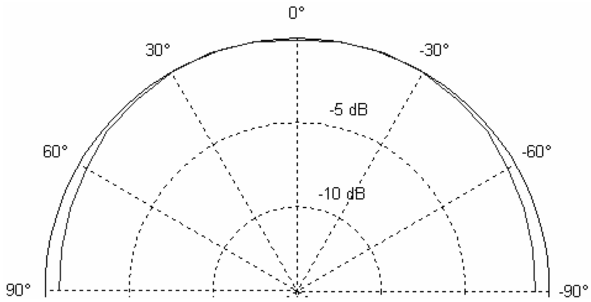
Figure 12 Linear DLA Directivity (a) f_{in}= 250Hz (b) f_{in}= 1KHz, (c) f_{in}= 2KHz, (d) f_{in}= 10KHz, $l=2m$



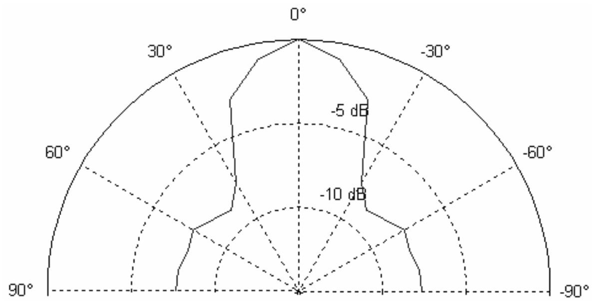
(a)



(b)

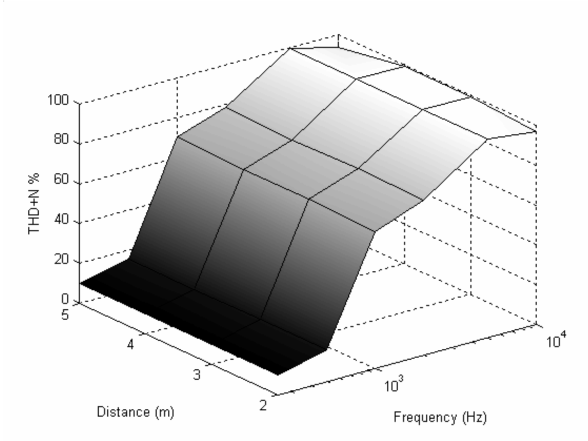


(c)

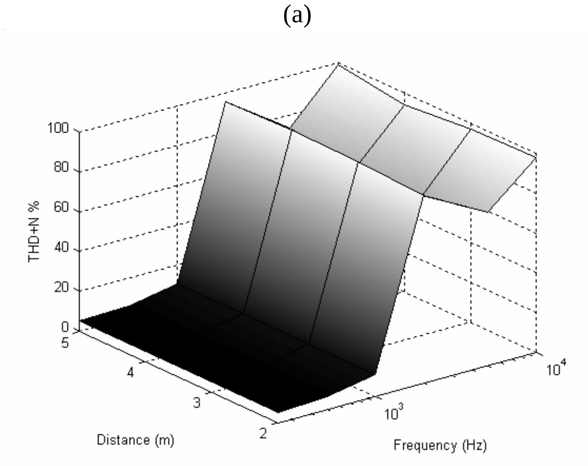


(d)

Figure 13 2-D DLA Directivity, (a) f_{in}= 250Hz (b) f_{in}= 1KHz, (c) f_{in}= 2KHz, (d) f_{in}= 10KHz, *l=2m*



(a)



(b)

Figure 14 Linear DLA THD+N%, (a) φ=60^o, (b) φ=30^o

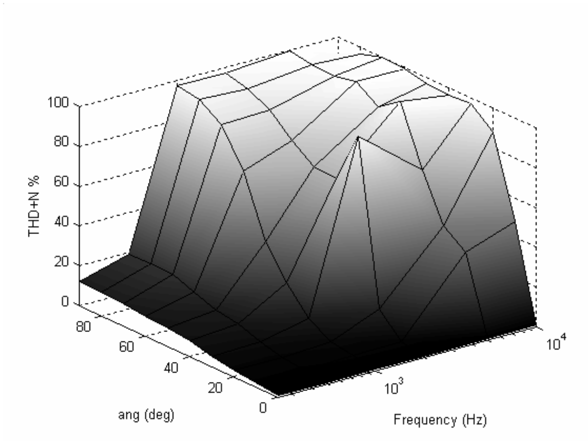
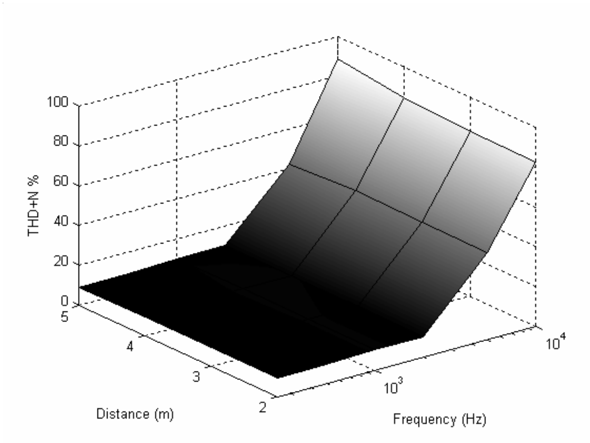
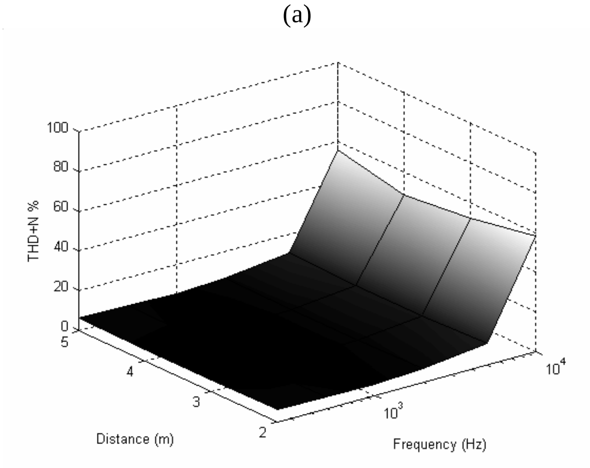


Figure 16 Linear DLA THD+N%, *l*=3m



(a)



(b)

Figure 15 2-D DLA THD+N%, (a) φ=60^o, (b) φ=30^o

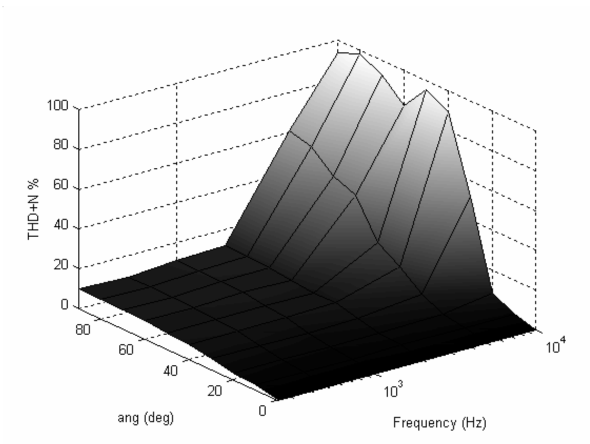


Figure 17 2-D DLA THD+N%, l=3m

4. CONCLUSIONS

Digital Loudspeaker Arrays present a promising and challenging alternative to traditional loudspeaker technologies. Recently, previous studies have demonstrated the feasibility of direct transduction of multibit digital audio data, typically from PCM-based stream. This stuffy demonstrates that direct digital-to-acoustic transduction of one-bit Sigma-Delta modulated audio (as on the DSD format), is feasible, achieving better performance than the comparable multibit scheme. In addition, such one-bit DAE schemes appear to allow for designs of smaller dimensions and complexity.

A novel aspect of this work is related to the necessary mapping of the one-bit stream to frames which are then appropriately distributed to the acoustic transduction elements. The schemes tested here appear to inhibit the system's performance by reducing SNR to approx. 60dB (in audio band). Hence, a conclusion of this work is that such a bit-grouping pre-processing stage needs to be further examined and optimized in order to improve the system's performance prior to transmission.

However, for this study idealized small emission elements were assumed, which for any practical realization would restrict the system's low frequency response. In this sense, the current DLA schemes appear to be appropriate for mid-to-high frequencies range digital reproduction, with a total array size comparable to a traditional converter unit.

The simulation tests indicate that array geometry in conjunction with its dimensions (combination of the required number of elements and the individual element size) is a crucial parameter for DLAs performance. Given the above restriction, the optimum one-bit DLA setup presented here is the 2-D scheme which has in all respects superior performance to a linear array. On-axis positions produce results similar to a conventional loudspeaker and for off-axis positions, for frequencies less than 2KHz, acceptable THD+N ratings are achieved.

As it is obvious, many practical limitations such as element technology and power handling were not examined in this study and will have to be further investigated.

5. ACKNOWLEDGEMENTS

The authors wish to thank Dr. Andreas Floros for his suggestions and contribution to this work.

6. REFERENCES

[1] Y.Huang, S. C. Busbridge, D. S. Gill, "Distortion and Directivity in a Digital Transducer Array Loudspeaker", J. Audio Eng. Soc., Vol. 49, No. 5, 2 May 2001

[2] S. C. Busbridge, P. A. Fryer, Y. Huang, "Digital Loudspeaker Technology: Current State and Future Developments", Audio Eng Soc. 112th Convention, May 2002

[3] M.O.J. Hawksford, "Smart directional and diffuse digital loudspeaker arrays", Audio Eng. Soc.110th Convention, May 2001

[4] N.-A. Tatlas, A. Floros, P. Hatziantoniou, and J. Mourjopoulos, "Towards the All-Digital Audio/Acoustic Chain: Challenges And Solutions", Audio Eng Soc.23th Conference, Copenhagen,, May 2003

[5] B.Fox, "The First Digital Loudspeaker?", Hi-Fi News Rec. Rev., IPC Media (Nov. 1998)

[6] D. G. Meyer, "Computer Simulation of Loudspeaker Directivity"J. Audio Eng Soc, Vol 32, No 5, May 1984