ICA 2013 Montreal  
Montreal, Canada  
2 - 7 June 2013  

Architectural Acoustics  
Session 4pAAa: Room Acoustics Computer Simulation II  

4pAAa7. Speech intelligibility prediction in very large sacral venues  
Wolfgang Ahnert*, Stefan Feistel and Tobias Behrens  

*Corresponding author's address: AFMG Technologies GmbH, Arkonastr. 45-49, Berlin, 13189, Berlin, Germany, wahnert@ada-amc.eu

In very large sacral venues like cathedrals or mosques the intelligibility of spoken words is very important especially during preaching. For such venues with volumes of up to and more than one million m3 special routines are needed for simulation to obtain predicted STI values based on more than 1000 sound sources. Special cloud computing has been developed which allows to do the calculation by providing the needed memory size and by cutting the calculation time from days or weeks to hours. Here also modern binaural or ambisonic B-format impulse responses are derived. In order to have a room acoustically correct model for simulation, the absorption behavior of typical floor materials in such venues has to be known like worshipers in church pews or sitting or kneeling on carpets in mosques. This absorption provided by the floor area is often the only one in sacral venues to reduce the reverberation time. For mosque projects absorption coefficients of persons in typical postures are unknown, so measurements have been done according to the reverberation room method. Persons have been tested on a carpet while standing, sitting or being in Muslim specific praying posture on a carpet.

Published by the Acoustical Society of America through the American Institute of Physics
SPEECH INTELLIGIBILITY PREDICTION IN VERY LARGE SACRAL VENUES

1. Introduction

The computer simulation in large sacral venues is always a problem, especially if a large number of loudspeakers and here especially of modern line arrays are in use. The following items are here to consider:

- Large volume of more than 30,000m$^3$
- Very large number of active sound sources
- Unknown acoustic properties of wall, floor and ceiling materials

In the past only rough estimations have been done or the whole project, the whole volume has been subdivided in sub-volumes to allow simulation runs, to receive values for coverage or achievable intelligibility. Quite often simple calculation routines have been done by neglecting the right reflection pattern and only the direct sound coverage has been used to estimate the speech intelligibility to be expected.

In European large sacral buildings like cathedrals this method was more or less usable because the installation points of the arrays or sound columns and the corresponding audience areas are not too far from each other. Here a lot of construction columns allow placing arrays very close to the audience. A different situation we have in mosques or also in baroque churches with large open worship areas without any construction details to attach or better hide sound sources.

Here only very directed sources in high numbers may solve the problem.

Another issue is the unknown sound absorption of the surface materials. In most of all sacral buildings the ceiling and wall materials are sound reflecting, in cathedrals also the floor. But the used pews in churches do absorb enough sound to decrease the reverberation [1]. In mosques we don’t have pews but carpets. Here we need to know the acoustic properties of typical carpets and this with and without worshipers.

2. Influence of large volumes

In the following we use some examples to show the range of volumes we have to work with:
FIGURE 2.1: Computer models of different sacral buildings with different volumes

In all these buildings a speaker design has been done to make sure that on all worshippers’ places the speech intelligibility is perfect. It must be mentioned, that these models only show spaces which are connected to the same volume. By doing the speaker design one has had to consider, that not only the direct sound from the nearest arrays or speakers will cover the corresponding prayer area but also the diffuse radiated sound from all the other sources in the connected volume. The calculation in the models A and B with altogether around 50 line arrays was still possible on normal PC’s. Depending on the computer power it took a time but the calculation with Raytracing and by using scattering was finished within hours.

In contrast sample C could only be handled in parts on PC’s. Calculation time issues (weeks of calculation in the demanded resolution) but especially memory limitations did not allow to make the calculation in the full model. For calculations in model D even the splitting of the whole volume in sub-volumes did not solve the problem. Even with sub-models calculation times of days have been needed on normal PC’s.

Five years ago we did start with so-called cloud-computing, the basics will be explained in the next chapter.

3. CLOUD-COMPUTING AND PROGRESS IN CALCULATION TIMES

3.1. Overview

Raytracing- and Particle-based Methods

There exist a variety of different approaches based on the ray-like propagation of the sound wave [2]. They all have in common that reflection paths are computed by taking into account the geometry of the room boundaries relative to the source and the receiver. Usually one or several of the following methods are employed:

- A systematic search for important reflections by "scanning" the room's surface. This may be accomplished by sending out rays from the source using a defined angular resolution or by directly computing possible reflection paths utilizing the mirror-image model.
- A Monte Carlo approximation of the specular reflections of the reverberant field by radiating sound particles from a source in random directions and tracing them until a receiver is hit.
- A Monte Carlo approximation of the diffuse part of the reverberant field by generating non-specular or scattered reflections at the boundaries and tracing them to the receiver.

In order to determine the room's transfer function most accurately all three strategies have to be combined in a hybrid approach [3]. Approaches not implementing the last but computationally expensive step have often resorted to estimating artificially a so-called random tail from extrapolating the early reflections.
Parallelization of Algorithms
The type of algorithm described above is well suited for parallelization. This can happen on two levels:
- Computation can be parallelized over the number of rays or particles traced since they can be processed largely independent from each other. For some methods a shared memory is required in order to avoid double-detection etc.
- If the number of sound sources is large compared to the number of parallel processing units the computation can also be parallelized over the number of sound sources, assuming that linear superposition principles hold for the results.

In practice, a calculation cluster is usually composed of several servers with multiple processors each. It has turned out that it is most efficient to incorporate both above levels of parallelization [4]. Particle-based parallelization requires more data exchange between processing threads than loudspeaker-based parallelization. It should therefore preferably take place on the same server. In contrast, raytracing computations for different loudspeakers usually require no or only small data exchange. These calculations can be distributed over different servers.

Cluster Architecture
The calculation cluster has to provide a communication interface that can be used by the end user. A typical installation will thus consist of
- several node servers that perform the calculations,
- a master server that manages the nodes and stores the results,
- a webserver that transfers input and output data between the master server and the client computer via the internet,
- the computer of the end user that hosts a client software, e.g. EASE, in order to configure and control the calculation runs.

Figure 3.1 illustrates this data flow between cluster components.

![Cluster Architecture Diagram](image)

FIGURE 3.1: Cluster Architecture

3.2. Large Volumes
Especially in very large venues, acoustic simulation is mostly necessary as well as difficult to conduct. Rooms with large volumes are usually very reverberant and acoustically challenging. They have long reverberation times and the sound system typically incorporates many loudspeakers. The design difficulties arising from this complexity make acoustic simulations a necessity. At the same time these room properties also lead to exceptionally long computation times and high memory demands.

In order to estimate the late part of the reverberant field particle numbers are required that scale with the room volume for resolution reasons. For example, ray-tracing results of satisfying accuracy may require 10-100 particles per m³ [5]. For volumes of 1 Mio m³ this translates to 10 Mio particles and more. Processing millions of particles, each experiencing 10-20 reflections or more, equates to calculation times of hours and days on a contemporary desktop computer.

Similarly, comparably long reverberation times necessitate large amounts of memory. In rooms with RTs of 5-10 s, the memory required for the simulated time response (echogram) data can be as much as 2 MB per loudspeaker-receiver combination. Evidently this leads to extraordinary needs regarding RAM and HD space when considering spaces with hundreds of loudspeakers and hundreds to thousands of simulated receive locations.

Both aspects, long calculation times as well as high memory demands, can only be approached sensibly by employing cluster computing technologies.

3.3. Practical Realization
In recent works a first implementation of a computing cluster was introduced [4]. The new version is strongly expanded and revised with respect to both software and hardware and allows performing extensive acoustics simulation runs also for very large rooms, e.g. prominent sacral venues.
Hardware Implementation

There are several base requirements for the hardware of the calculation cluster. Cluster nodes have to provide high performance for integer and floating point operations. The RAM of each node has to be sufficiently sized in order to facilitate the most extended computations needed. It also has to have short access times. Hard disk requirements are high, too, since especially for many loudspeakers a large amount of intermediate results are generated before the final data can be assembled.

The expanded calculation cluster is using the computing hardware listed in Table 3.1.

<table>
<thead>
<tr>
<th>Number of nodes</th>
<th>EASE AURA Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cores per node</td>
<td>32</td>
</tr>
<tr>
<td>Type of processor</td>
<td>6282SE 2.6 GHz (2x 16C)</td>
</tr>
<tr>
<td>RAM per node</td>
<td>32 GB</td>
</tr>
<tr>
<td>Backbone data connection</td>
<td>10 Gbit</td>
</tr>
<tr>
<td>HD primary</td>
<td>800 GB SAS RAID</td>
</tr>
<tr>
<td>HD secondary</td>
<td>3000 GB SATA</td>
</tr>
<tr>
<td>Avg. power consumption (standby / full load)</td>
<td>1800 W / 4500 W (ca.)</td>
</tr>
</tbody>
</table>

TABLE 3.1: Technical specification of expanded calculation cluster.

Software Implementation

It is important that the simulation software is using the most modern compiler technology as e.g. available through MS Visual Studio 10 and Intel Compiler Suite 11. Only this allows fully exploiting modern processor and caching architecture.

In order to support computations using substantially more memory than 2 GB the simulation software is based on 64-bit. As outlined earlier, large projects with many receivers may require 10 GB of RAM and more, even for a single sound source.

In large venues, digitally steered columns and other types of line arrays are often used because of their focused radiation of sound. These advanced loudspeaker systems are modeled in EASE by using DLL and GLL files which are often 32-bit based and partially not thread-safe. These limitations can be overcome by means of a socket-based communication interface between the main program and the loudspeaker plug-in modules.

If the project includes large amounts of receivers or extended audience areas the calculation time for testing if a particle hits a receiver can become dominant. This bottleneck can be removed by means of logarithmic intersection tests, such as binary space partitioning (BSP) as implemented here.

If the project includes many receivers as well as many loudspeakers intermediate results can be very large, on the order of several TB. It is therefore imperative that all data are transferred and stored in a packed format and that the necessary number of data transfer steps is minimized.

Results

Using the described calculation cluster, computation times for typical venues are reduced by a factor of up to 50 relative to contemporary PCs. Table 3.2 and Figure 3.2 display the computation times for a selection of sacral buildings and a shoebox room in comparison. It is particularly remarkable that a full simulation run for the Room D takes about 23 hours whereas the same calculation would take more than a month on a conventional desktop computer.

<table>
<thead>
<tr>
<th>Room</th>
<th>Volume / m³</th>
<th>Time / min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoebox</td>
<td>12 000</td>
<td>10</td>
</tr>
<tr>
<td>Room A</td>
<td>29 000</td>
<td>35</td>
</tr>
<tr>
<td>Room B</td>
<td>63 000</td>
<td>120</td>
</tr>
<tr>
<td>Room C</td>
<td>171 000</td>
<td>426</td>
</tr>
<tr>
<td>Room D</td>
<td>1 690 000</td>
<td>1380</td>
</tr>
</tbody>
</table>

TABLE 3.2: Room volume of selected sacral buildings and computation times at comparable settings.
FIGURE 3.2: Typical calculation times as a function of the room volume according Fig. 2.1. Comparable resolutions and calculation settings have been used.

As a result of this cloud-computation procedure impulse responses of up to 6 s length may be calculated. Schroeder showed that these full impulse responses may be used to derive the corresponding STI values describing the achievable speech intelligibility [6].

For analysis and presentation purposes, auralization of results is a powerful tool in addition to quantitative evaluation. The software can provide impulse response and transfer function data in monaural and binaural output formats as well as newly in B-format.

4. Absorption of carpets in combination WITH different worshiper postures

4.1 General

In order to have a room acoustically correct model for simulation, the absorption behavior of typical floor materials in such venues has to be known like worshipers in church pews or sitting or kneeling on carpets in mosques. This absorption provided by the floor area is often the only one in sacral venues to reduce the reverberation time. For mosque projects absorption coefficients of persons in typical postures are unknown, so measurements have been done according to the reverberation room method.

In a first step the sound absorption coefficients of a designated carpet had been measured according the reverberation room method as described in ISO 354: 2003 “Acoustics - Measurement of sound absorption in a reverberation room” [7]. The size of the carpet sample available for the measurement as test specimen was around 10 m² and thus according to ISO 354. The thickness of the carpet was approx. 10 mm.

A special rubber foam underlay of 8 mm thickness for enhanced worshipers’ physical comfort was used for additional measurement runs. Here a certain low frequency absorption boost can be expected due to the physical principle behind supplementing the porous absorption principle of a carpet by the resonance (plate oscillation) absorption principle. All the measurements with and without persons on the carpet have been done with and without the 8 mm thick rubber foam underlay for the carpet. In addition to the measurement of only the carpet (with and without the special rubber foam underlay) different person arrangements on the carpet have been tested as well. For the measurements of different person arrangements on the carpet a number of 14 persons as
test persons have been recruited, for the measurement arranged on the carpet in 2 rows (each 7 persons) as follows:

- standing
- sitting on the floor
- praying posture (kneeling on the floor, crouched down, bent forward)

The 14 persons (6 men, 8 women) were dressed with long trousers, no shoes and the women wore head scarfs.

In order to obtain practice-oriented data for the use of simulation, a surrounding reflective barrier was necessary preventing lateral sound incidence into the person arrangements. This reflective barrier has been constructed with wooden (MDF) plates of 16 mm thickness (12 kg/m²) and 1 m height, standing on the floor surrounding directly the carpet as test specimen.

Equivalent absorption areas \( A \) of the test specimen derived from the measured reverberation times (based on impulse responses) have been calculated according the reverberation room method as described in ISO 354 “Acoustics - Measurement of sound absorption in a reverberation room”. In order to get absorption data for simulation purposes, sound absorption coefficients \( \alpha \) have been derived by dividing \( A \) by the floor area within the reflective barrier.

### 4.2. Measurement results

The first two figures show the results (absorption coefficient vs. frequency) for the different person arrangements as curve-parameter (standing, sitting, praying posture) and for the carpet with underlay (Fig 4.1), the other one for the carpet without underlay (Fig. 4.2). As well the measurement results for the carpet without any persons for both runs, with and without the rubber foam underlay, are displayed in Fig. 4.3.

**FIG. 4.1:** Absorption coefficients of the different person arrangements placed on the carpet with underlay, as well as the data for the unoccupied carpet (measured without reflecting barrier).
FIG. 4.2: Absorption coefficients of the different person arrangements placed on the carpet without underlay, as well as the data for the unoccupied carpet (measured without reflecting barrier).

FIG. 4.3: Absorption coefficients of the carpet with and without foam rubber underlay (no surrounding reflective barrier was used here)
4.3. Summary and conclusion of the absorption measurements

The results show that the different person arrangements lead to different values. This can be explained by the different body surface areas of standing, sitting and praying people. In the case of standing worshipers, the surface area (able to absorb) is maximum size, decreasing in the case of sitting and even more in praying posture. On the same hand the hidden area of the carpet by the worshipers is minimum in case of standing and maximum in case of praying posture.

The results lead to the conclusion that data for absorption of persons as given in the acoustical standard literature (normally given only for standing persons or persons sitting on chairs [8], [9]) is inappropriate for practice orientated simulations in mosques: The absorption ability of persons sitting on floor or even more being in Muslim specific praying posture is significantly smaller than the data given in acoustical standard literature for standing and chair seated persons.

The rubber foam underlay for the carpet made for enhanced physical comfort leads to a significant boost of absorption ability in the low-mid frequency range (approx. 200-800 Hz).

4. Conclusions and Outlook

The prediction of acoustic properties in rooms and open spaces is today a normal procedure, but the results are quite often not precise enough as measurements later show. By measuring the impulse response on a listener place you obtain a kind of an acoustic fingerprint. By using this response you may derive objective acoustic criteria characterizing subjective properties at that place. But, because of limitations in time and memory often only approximations are done in the computer model, which leads to wrong conclusions about the acoustic quality of a room or a sound system.

In complicated room structures with a huge amount of different sound sources the new cloud-computing algorithm will allow calculating full-size impulse responses without any assumptions or approximations. Of course the frequency range around 100 Hz and lower is not yet included. This is one of next fields of enhancement of our tools.

These impulse responses have to be calculated in reasonable short time, otherwise the procedure loses the function as a design tool. Measurements are done in minutes and simulations need still too much time or they are only approximations. Until now for simulations in complicated room structures with a lot of sources we need still hours or a day but by enhancing this approach we want to reduce the calculation time to one or two hours or even less.

5. Bibliography