



HAL
open science

Topological Sound Propagation with Reverberation Graphs

Efstathios Stavrakis, Nicolas Tsingos, Paul Calamia

► **To cite this version:**

Efstathios Stavrakis, Nicolas Tsingos, Paul Calamia. Topological Sound Propagation with Reverberation Graphs. *Acta Acustica united with Acustica*, 2008, 94 (6), pp.921-932. 10.3813/AAA.918109 . inria-00606808

HAL Id: inria-00606808

<https://inria.hal.science/inria-00606808>

Submitted on 18 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Topological Sound Propagation with Reverberation Graphs

Efstathios Stavrakis*
REVES-INRIA

Nicolas Tsingos†
REVES-INRIA

Paul Calamia‡
Rensselaer Polytechnic Institute

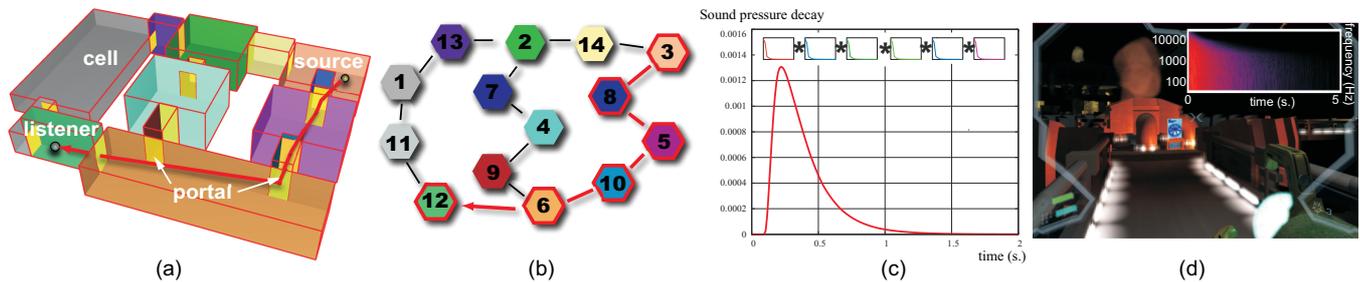


Figure 1: (a) Our approach uses a 3D scene, decomposed into cells and portals, to compute reverberation effects interactively. (b) We estimate global pressure decays along possible sound propagation routes in the cell adjacency graph describing the 3D geometry. (c) The pressure decay for a route (highlighted in red in (b)) is obtained by convolving room-to-room transport operators computed off-line. (d) Our graph-based reverberation engine has been integrated into a 3D game with scalable reverberation processing and supports up to 30 sources interactively. Inset shows the spectrogram of a reconstructed reverberation filter.

Abstract

Reverberation graphs is a novel approach to estimate global sound-pressure decay and auralize corresponding reverberation effects in interactive virtual environments. We use a 3D model to represent the geometry of the environment explicitly, and we subdivide it into a series of coupled spaces connected by portals. Off-line geometrical-acoustics techniques are used to precompute *transport operators*, which encode pressure decay characteristics within each space and between coupling interfaces.

At run-time, during an interactive simulation, we traverse the adjacency graph corresponding to the spatial subdivision of the environment. We combine transport operators along different sound propagation routes to estimate the pressure decay envelopes from sources to the listener. Our approach compares well with off-line geometrical techniques, but computes reverberation decay envelopes at interactive rates, ranging from 12 to 100 Hz.

We propose a scalable artificial reverberator that uses these decay envelopes to auralize reverberation effects, including room coupling. Our complete system can render as many as 30 simultaneous sources in large dynamic virtual environments.

1 Introduction

Modeling of sound propagation is very important for virtual acoustics and virtual reality applications. In particular, indoor reverberation effects due to sound scattering off wall surfaces carry major cues related to the size of the environment and distance to sound sources [Begault 1994]. Therefore, reverberation helps users to establish a better sense of presence in virtual environments and thus is arguably one of the most important audio effects to simulate.

A number of geometrical approaches have been presented to model early specular reflections (and possibly edge diffraction) interactively [Lauterbach et al. 2007; Lentz et al. 2007; Funkhouser et al. 2004]. Such approaches can handle complex geometry but are too computationally intensive to be applicable to interactive reverberation effects. Traditionally, interactive reverberation effects in *single rooms* are implemented by estimating energy decay rates and deriving *artificial reverberation filters* [Jot 1999]. Parameters of the decay envelope can be directly manipulated by the sound designer to achieve the desired effect without requiring any geometrical modeling. Alternatively, the energy decay rates can also be determined from few geometrical parameters using well-known statistical models, such as the Sabine and Eyring formulas [Kuttruff 1973]. Such approaches are widely used in current audio production environments and consumer audio hardware [EAX 2004], but they cannot model the global propagation of sound throughout a complex multi-room environment.

Consider the typical example of a game-like environment containing several connected rooms such as in Figure 1 (a). The late energy decay in the listener’s room will be governed by the trans-

Keywords: Spatial audio rendering, reverberation, geometrical acoustics, statistical acoustics

*Efstathios.Stavrakis@sophia.inria.fr

†Nicolas.Tsingos@sophia.inria.fr

‡Also with the Department of Computer Science, Princeton University.
pcalamia@cs.princeton.edu

mission of energy from the neighboring rooms through the open doors. The corresponding indirect energy contributions will also have strong directional components that need to be rendered correctly from the location of the coupling apertures. In the context of architectural acoustics, coupled-room configurations have been increasingly studied for the design of modern concert halls. While a preference for so-called “double-sloped” decays has not been established [Bradley and Wang 2005; Ermann 2007], they have been shown to be clearly audible under a variety of experimental conditions [Picard 2003; Knight 2003; Ermann 2007] and thus are necessary to replicate for perceptually accurate reverberation effects.

In this paper, we concentrate on interactive, yet physically-based, modeling of reverberation effects in indoor architectural environments containing arbitrarily coupled rooms. Our approach combines characteristics of geometrical and statistical acoustics to achieve realistic results at interactive rates.

Starting with a 3D input model, an adjacency graph that describes the spatial coupling of acoustic spaces is constructed. Using a divide and conquer strategy, energy decay envelopes are then determined *off-line* for each space using path tracing simulations. This decouples complex geometrical calculations from global propagation, which is performed *on-line* by exploring all energy propagation routes along the spatial adjacency graph.

Similar to statistical acoustics, we assume that energy is scattered from space to space through the coupling apertures, which act as diffuse emitters [Summers et al. 2004]. We also assume that the fine-grain temporal structure of reverberation filters can be modeled as statistical noise. From the global propagation routes physically-based artificial reverberation filters can be efficiently constructed and used for real-time spatial audio rendering.

Our contributions are:

- The definition of transport operators for general acoustically coupled rooms.
- An efficient global energy propagation algorithm based on the topology of the room adjacency graph.
- An efficient filter construction algorithm leveraging the coherence of sound propagation routes in the graph.
- A scalable interactive system for auralizing physically-based late reverberation effects for environments with many sources.
- An initial validation against alternative geometrical acoustics solutions.

Our complete pipeline allows for rendering up to 30 sources in complex dynamic environments.

2 Related Work

A large body of work has been devoted to simulating sound propagation effects in virtual environments. In the following section we review the techniques most directly related to our work.

2.1 Statistical acoustics

Statistical acoustics (SA) are widely used to determine energy decay rates in rooms and lead to efficient approaches to model reverberation effects. For instance, the Eyring model describes the energy decay in an enclosure as a decaying exponential:

$$E(t) = E_0 e^{\frac{cS}{4V} t \log(1-\bar{\alpha})}, \quad (1)$$

where c is the speed of sound ($\approx 340m.s^{-1}$), V is the volume of the space, S the total surface area of the walls and $\bar{\alpha}$ the surface-averaged absorption coefficient.

While such approaches can give good results in the case of single rooms, they fail to capture the multi-slope decays created when rooms are connected. In Figure 2, decays of two independent rooms are compared with a decay obtained by coupling the two spaces. Decays were computed using a ray-tracing approach to construct image sources [Vorländer 1989]. One million rays were used to compute a “reference” solution up to 200 reflection orders. Note that the slope of the decay is not linear in dB scale and is a combination of the decay slopes of the individual rooms.

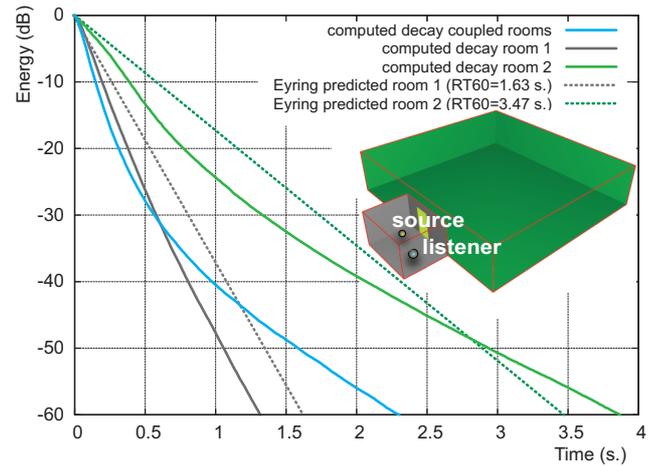


Figure 2: Integrated energy decays for two uncoupled and coupled rooms. The volume of room #1 is $29 m^3$ while room #2 is much larger at $578 m^3$. Energy absorption is uniform on all wall surfaces and was set to 5%. The double decay slope due to coupling is clearly visible on the blue curve.

SA methods specifically designed to predict reverberation time in coupled volumes also exist. The earliest results for *two* coupled rooms are due to Davis [Davis 1925] and Eyring [Eyring 1931]. Further details for the two-room case are given by Lyle [Lyle 1981a; Lyle 1981b] and Cremer and Müller [Cremer and Müller 1982], and Kuttruff extended the technique to an arbitrary number of coupled volumes [Kuttruff 1973]. All of these assume that the reverberant decays for each room in the coupled system are exponentially decaying and can be predicted independently. They also assume that modifications to the single-room decays occur due to diffuse energy exchange through portals between spaces. Summers et al. [Summers et al. 2004] provide an improved SA model for coupled volumes which uses corrections based on geometrical acoustics to increase the accuracy of single-room decay predictions as well as room-to-room energy exchanges.

Our approach can be seen as a generalization of previous SA approaches but also aims at providing a finer-grain estimation of the short-time integrated impulse response so that it can be directly auralized. A solution to more accurate modeling is to use methods based on geometrical acoustics.

2.2 Geometrical acoustics

Geometrical acoustics (GA) is probably the most popular approach for physically-based acoustic modeling. GA is a high-frequency approximation that models sound propagation along ray-paths. Specular reflection paths can be constructed using techniques such as

ray or beam tracing [Lauterbach et al. 2007; Lentz et al. 2007; Funkhouser et al. 2004]. Monte-Carlo path tracing or radiosity-like approaches can be used for diffuse exchanges [Heinz 1993; Dalenbäck 1996; Siltanen et al. 2007].

GA-based techniques have been previously used for the off-line prediction of reverberation in coupled volumes. Using ODEON [Christensen], a commercially available GA-modeling program, Bradley and Wang [Bradley and Wang 2007] evaluated the effects of various architectural parameters on the energy decay in a coupled-volume system. They also have shown good agreement at high frequencies between ODEON simulations and measurements from an existing coupled-volume concert hall [Bradley and Wang 2007]. Summers et al. [Summers et al. 2005] have shown that reverberant tails in coupled volumes can be predicted accurately with ray tracing in the software CATT [Dalenbäck].

In a recent paper [Lentz et al. 2007], Lentz et al. report update times of up to 3 sec. to update full reverberation filters in a Monte-Carlo path tracing context. Early reflection visibility checking takes about 0.7 sec. for up to order 3 in their 105 polygon test model, leading to 300000 potentially visible image sources to test. While acceptable in the context of acoustics design, such update rates are limiting for typical consumer entertainment applications, such as games, where the simulation should typically run between 10 and 60Hz to accommodate fast listener or source motions and geometrical changes in the environments. Another paper [Laine et al. 2008], using a beam-tracing approach, reports constructing all third order reflections of a sound source in 1 msec. for a 1190 polygon model. However, even the fastest interactive GA approaches remain limited to low-orders of reflection (typically 8) and few sources. Hence, they cannot be applied to compute full reverberation effects interactively since reverberation filters typically contain audible reflections of order 100 or more.

To avoid the cost of an on-line GA simulation, Pope et al. [Pope et al. 1999] have proposed to precompute and store the reverberation filters for a number of locations in a virtual environment. However, this approach does not scale well, since it would require computing a number of filters quadratic with the number of source/listener locations or rooms. It also does not support any interactive modification of the environment. In this paper however, we will retain the approach of pre-computing costly GA-based decay envelopes in an off-line process.

2.3 Artificial Reverberators

Statistical acoustics formulations lead to the development of efficient *artificial reverberators*, which are currently the most widespread techniques used to auralize late reverberation effects in interactive virtual environments [Gardner 1998; Rocchesso 2002]. Artificial reverberators do not model the fine-grain temporal structure of a reverberation filter but assume that reverberated components can be modeled as a temporal noise process modulated by slowly-varying pressure-amplitude envelopes in different frequency subbands. Following Eq 1, these envelopes are often considered as exponentially decaying, which leads to the design of efficient recursive *Feedback Delay Network* (FDN) filters [Schroeder 1962; Jot 1999; Gardner 1998; Rocchesso 2002]. This approach of modeling the fine-grain structure of reverberation filters as a noise process has also been used to auralize Monte-Carlo path tracing results [Kuttruff 1993; Schröder et al. 2007]. In this case, a short-time integrated impulse response is constructed via path-tracing and is used to modulate the noise process. In the general case, this approach cannot be directly implemented using the previous FDN formulation due to non-exponential decays, and results in a high computation cost. In this work, we will also be using this latter solution

but show that it can be efficiently implemented within a scalable Fourier-domain framework [Tsingos 2005].

3 Overview

Our approach requires a 3D model of the environment to simulate. In a first pre-processing step, we decompose the 3D environment into a cell-and-portal structure [Teller 1992; Lefebvre and Hornus 2003; Haumont et al. 2003]. We build an adjacency graph that encapsulates the physical connectivity of the individual building blocks of the geometry, i.e. the cells and portals. The nodes of the graph correspond to the rooms and the edges to the different coupling apertures (see Figure 1 (b)).

In contrast to previous GA work [Lentz et al. 2007; Funkhouser et al. 2004], we do not use the cell-portal decomposition to speed up visibility checks. Rather, we use the topology of the corresponding adjacency graph to decouple the acoustic spaces and treat them independently. As such, our work shares some similarities with the approach concurrently suggested in [Foale and Vamplew 2007].

By following this divide-and-conquer strategy, we seek to model reverberation as an energy transfer between rooms, using operators similar in spirit to form-factors [Siltanen et al. 2007]. These operators encode not only direct propagation from room to room but also higher order reflections within rooms that encounter the portals and thus lead to additional, delayed energy exchange. For a more accurate modeling of the acoustical behaviour of each subspace, these operators are pre-computed for each individual room using a GA approach. Section 4 describes the operators and their off-line evaluation.

During an on-line simulation, global room-to-room propagation is solved by traversing the adjacency graph and constructing all possible propagation routes from all rooms to the current listening room (Figure 1 (c)). These routes will be re-used to auralize all audible sound sources in the listening room. Once sound-propagation routes have been determined, global short-time integrated pressure decays are computed for each source room. This involves convolving adjacent transport operators as energy radiates from subspace to subspace along each route. We propose to compute the required convolutions numerically in the Fourier domain. In addition, we propose a convolution caching mechanism that leverages the coherence between all propagation routes in the graph. This process is described in Section 5.

A set of matching artificial reverberation filters are then generated for auralization (Figure 1(d)). In Section 6, we show that efficient auralization of the reverberation effects can be achieved using a scalable rendering approach, similar to [Tsingos 2005], which we extend to include spatial rendering of the reverberated components.

Finally in Section 7, we show that our approach compares well with full GA solutions and that our rendering engine can deal with complex environments at interactive rates.

4 Transport operators for coupled spaces

In this section, we introduce *transport operators*, or TOPs, describing global energy exchanges inside and in-between spaces. Transport operators model how energy decays locally within a single building block of the virtual environment due to reverberation. We also introduce operators modeling the response of incoming sound at the interface between each building block, i.e. portals. A specific pressure decay function is computed and associated with each transport operator.

4.1 Definitions

We define several transport operators for any given room in the environment. To handle general environments, we pre-compute transport operators using Monte-Carlo path tracing, see Figure 3 for an example.

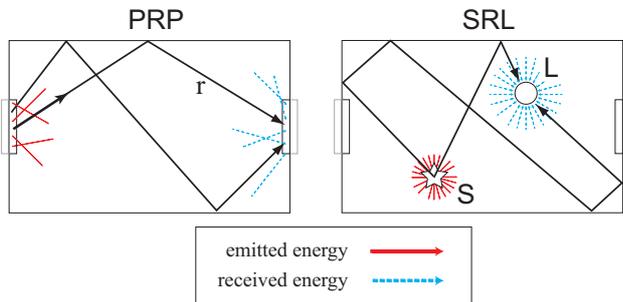


Figure 3: Illustration of the path tracing process used to compute a PRP and SRL transport operator.

The Source-Room-Listener (SRL) operator is the point-to-point response of a source in a room for a listener in the same room (Figure 3 right). It is used for sources that directly share the same space as the listener. This operator typically can be obtained by summing up the contributions of image-sources identified by the path tracer [Vorländer 1989]. In our current implementation, SRL operators do not include the direct contributions to the listener, which we maintain separately for faster update rates (see Section 6).

The Source-Room-Portal (SRP) operator is the response of a point source in a room, integrated on a given portal in the same room. In our current implementation, we use a uniform solid angle sampling from the point source and compute the energy response of the SRP operator as:

$$SRP(t) = \frac{4\pi}{nbrays} \sum_{i=1}^{nbrays} test_i \prod_k (1 - \alpha_k) \delta(t - r/c),$$

where $nbrays$ is the total number of rays fired and $test_i = 1$ if the ray intersects portal P and $test_i = 0$ otherwise. c is the speed of sound, r the length of the path, δ the Dirac function. The α_k 's are the surfaces' energy absorption coefficients along the path, with k being the reflection order. For better integration, russian roulette could alternatively be used to handle reflection coefficients.

The Portal-Room-Portal (PRP) operator is the response of a portal acting as a diffuse emitter integrated for all points on a portal of the same room (Figure 3 left). Note that the same portal can act as both a source and a receiver in a PRP operator, in which case only the indirect paths are sampled. To compute a PRP operator, we sample simultaneously the surface of the source portal P_i and the hemisphere of shooting directions and obtain the energy at the destination portal P_j as:

$$PRP_j(t) = \frac{2\pi}{nbrays} \frac{A_i}{A_j} \sum_{i=1}^{nbrays} test_i \prod_k (1 - \alpha_k) \delta(t - r/c),$$

where A_i and A_j are the respective areas of the source and destination portals.

The Portal-Room-Listener (PRL) operator is the response of a portal acting as a diffuse emitter for a point receiver in the same room. In our current implementation PRL operators are computed in the same way as SRP operators.

Portals can be fully absorbent (e.g., open door), fully reflective (e.g., closed thick door) or a mixture of both (e.g., a thin wall between rooms). If the portals are fully absorbent, a ray hitting the portal is propagated no further.

We also assume that operators do not depend on the position of the source and listener within a given space and estimate them for a representative source and listener located at the centroid of each room. However, this is not a hard limitation since multiple operators could be precomputed for several source or receiver locations within each space at the expense of more processing and a larger memory footprint.

As a result, for a room with N_p portals, $N_p^2 + 2N_p + 1$ operators must be determined, corresponding to N_p^2 PRPs, N_p PRLs, N_p SRPs and one SRL. Note that this number is a function of the local number of portals in each room and is independent of the total number of rooms, portals or the complexity of the underlying geometry present in each room. Example operators are shown in Figure 4.

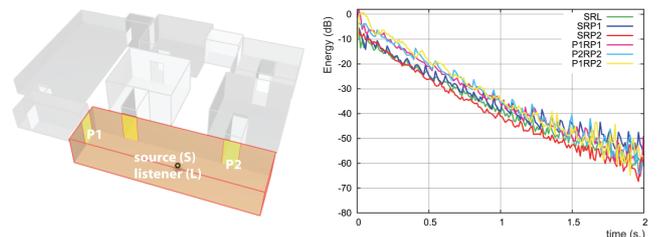


Figure 4: Example transport operators pre-computed for the high-lighted room.

4.2 Pre-computing and storing transport operators

Our current implementation supports specular reflections, which we compute for up to 200 orders, but could be extended to include diffuse scattering as well for improved modeling [Vorländer 1995]. A regular grid was used to optimize ray-triangle intersections. Any alternative GA prediction approach supporting point and area receivers could be used. Each operator is computed for a number of frequencies to account for absorption and propagation effects. In our current implementation we used four chosen as 200, 1000, 8000 and 16000 Hz.

In the case of PRP and SRP operators, energy of all paths hitting the portal is integrated over short time-frames consistent with human hearing resolution (typically $\Delta T \approx 10$ ms).

For PRL and SRL operators, we use the rays to construct image-sources that we render at the listener's location [Vorländer 1989]. We also integrate the energy of the corresponding image-sources using the same ΔT . To limit aliasing we use a 25cm-radius detection sphere which is roughly consistent with the size of the listener's head. Aliasing-free approaches, such as pyramidal beam tracing, could alternatively be used.

From the obtained energy responses, we take the square-root to obtain a short-time averaged pressure response. Table 1 shows pre-computation time for all operators in environments of varying complexity. As can be seen, the timings remain reasonable even for our highly conservative settings and pre-computation time largely depends on the topological complexity of the scene.

Operators are stored as short-time integrated pressure histograms (fixed size and ΔT), sampled at less than 200 Hz. We use a size of 1024 samples and $\Delta T \approx 10$ ms., which can encode decays up

to 10 sec., and requires a storage of 4Kb per histogram assuming 32-bit floating point values. To account for frequency-dependent absorption effects, we store 4 echograms corresponding to different frequency subbands. Table 1 shows memory requirements to store the operators for environments of various complexities. As it can be seen, memory requirements remain reasonably low even for scenes with high polygon counts since transport operators only depend on the topological complexity of the scene.

In comparison to the brute force solution of storing all frequency responses for possible couples of source and listening rooms, this approach is more efficient, since we do not have to propagate through portals and thus can trace fewer rays. In addition, storing short time-integrated pressure decay envelopes, instead of full audio sampling-rate data (e.g., 44.1 KHz) reduces storage and memory footprint.

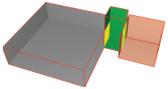
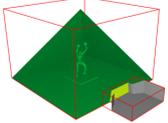
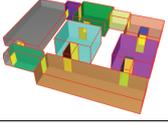
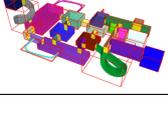
Environment				Time	Memory
3D Model	#cells(tris)	#portals	#TOPs	(mins)	(Kb)
	3(39)	2	17	6	272
	2(12729)	1	8	22	128
	14(195)	15	140	42	2240
	29(1614)	33	339	151	5424

Table 1: Precomputing transport operators. Statistics collected on an Intel Xeon 2.33GHz by tracing 300000 rays/operator, with 200 the maximum number of reflections per ray. Memory statistics are provided for operators with 4 frequency bands with a length of 1024 samples. We show color-coded cells, and portals are highlighted in yellow.

4.3 Combining operators

When sound propagates from room to room, we successively apply the transfer function of each operator along a desired sound propagation route. For example, Figure 1 (c) shows a decay envelope obtained by convolving all operators along a possible route. Combining operators is achieved by convolving their decay envelopes. An analytical solution for convolution of two exponential decay envelopes has been proposed in [Svensson 1998]. To handle more than one convolution and support non-exponential decays, we compute the convolution numerically. Individual transport operators are pre-stored in the frequency domain so that fast convolutions can be achieved using only complex multiplies.

We further assume that all decay envelopes start at $t = 0$ and account for possible initial time-delay gaps separately. The time-delays of all operators convolved along a route can be factored out and added together leading to the total time-delay for the route. This delay can be subsequently applied to the result of the convolution. This limits

aliasing problems, especially for small time-delay gaps.

We also need an estimation of energy contributed by a given propagation route. This energy is used as an oracle to direct graph traversal during the simulation of the global propagation and can also be used to prioritize rendering of propagation routes.

There is no general formula to estimate the energy of a convolution. However, an upper bound is given by:

$$\|f * g\|_2 \leq \min(\|f\|_2 \|g\|_1, \|f\|_1 \|g\|_2), \quad (2)$$

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1, \quad (3)$$

where $\|f\|_2 = \sqrt{\sum_i f^2(i)}$ and $\|f\|_1 = \sum_i |f(i)|$.

Given the starting energy of an *SRP* operator, we can successively apply Eq. 2 for successive *PRP* operators to estimate the energy carried along the route. Note that Eq. 3 is required to successively estimate the $\|f\|_1$ term in Eq. 2.

5 On-line topological sound propagation

Considering the initial cell-portal connectivity graph is already established and the transport operators for each room are computed, we describe in this section how global coupling is solved on-line by propagating energy through the graph and combining successive operators along all possible propagation routes. The pseudo-code of the complete algorithm can be seen in Algorithm 1.

5.1 Graph traversal

To model sound reverberation in the environment, we emit energy from the listener's room and, according to the connectivity graph, we select and apply the energy decays of respective transport operators to the propagated energy. In this way we approximate the exchanges that take place between the rooms of the environment.

The tree formed by selecting the listening room as the root node enables a fast, depth-limited, traversal algorithm to be applied. The criterion driving our traversal algorithm is that the propagated energy remains above a conservative energy threshold (see Eq. 2). In our simulations, we used a 10^{-10} threshold (i.e., a 100 dB dynamic range).

The traversal algorithm takes advantage of the connectivity graph between the rooms and portals of the environment. Each time we enter a new room, we identify and recursively follow all possible adjacent nodes. At each step we update the propagating energy by applying the energy decay of the respective transport operator (Eq. 2). Followed routes are typically coded as transport operator sequences, e.g. $P_n R_v P_{n-1} * \dots * [P_{i-1} R_m P_i] * P_i R_m L$. Note that since operators already include all reflections within rooms, edges in the reverberation graph model coupling between adjacent spaces and always connect different nodes. As a result, two successive operators in a sequence always correspond to different rooms.

This traversal only extracts the longest possible routes that energy can follow from room to room until it decays below our minimal threshold, and is independent of the actual location of the actual sound sources that need to be auralized. To account for the potential presence of sound sources in any room along these routes, we simply consider that a unique virtual source is present in each room. The initial extracted routes are further subdivided to construct all possible routes pairing these virtual sources with the listener, e.g. routes of the form $SR_v P_{n-1} * \dots * P_i R_m L$.

Since the transport operators of a route will later be convolved to construct decay filters, the computational requirement of creating

a filter along a route, made up of N_{TOPS} number of TOPs, will be equal to performing $N_{TOPS} - 1$ convolutions. One way to further control the performance in an interactive system is to set a global budget and prune all routes that have a higher computational payload based on the estimated energy they carry.

5.2 Convolution caching

Once the significant propagation routes have been determined, we must calculate how energy is spread through time along each route. This involves combining the transport operators along a route, which corresponds to a set of convolutions as described in Sec. 4.3.

To estimate the global decay shape for all source rooms to the listening room, we cluster these routes first by source room, then by incoming portal in the listening room. The sum of all decay envelopes ($SR_V P * \dots * P_i RL$) must be computed for each portal P_i in the listener's room and a source room R_V .

Convolutions are computed in the Fourier domain. Operators' FFT data is pre-cached so that only the complex multiplies required for convolution are performed on-line. Since the number of samples used to encode the operators is small, this is efficient. Typically, 850 convolutions can be performed at 60 Hz on a Intel Xeon 2.33GHz processor using an unoptimized implementation. However, as the environments become more complex, convolving the operators can become the bottleneck of the approach.

A key aspect of formulating the energy propagation as a sequence of operators that can be convolved together is that we can take advantage of the associative and commutative properties of convolution to significantly reduce the number of convolutions actually performed. Recall that energy will sequentially propagate between the rooms of the environment in an iterative manner, thus in many routes the same operators may occur multiple times. The commutativity of convolution enables us to arbitrarily re-order transport operators, by using a pre-assigned unique numeric value that acts as an identification number of each transport operator. Furthermore, the associative property of convolution allows us to apply a simple dictionary-based compression scheme in which transport operators are convolved pair-wise and new transport operators, denoting this convolution, are stored in a cache. Therefore, when filters are created for each route, this *convolution cache* is queried for pre-computed convolutions of operators. The high correlation of transport operators in the different sound propagation routes, emanating from the listener's room, provides a significant decrease in the number of required convolutions, as can be seen in Table 2. We obtained compression ratios of more than 10 to 1 in some of our test cases.

5.3 Handling changes in portal state

Our approach allows for modifying the topology of the graph interactively, for instance to model opening or closing doors in the environment. However, closing a door in a room will also affect transport operators which must then be recomputed. Our current solution to smoothly blend between portal states is to pre-compute each operator for all possible portal state (i.e., open/closed) in the room. We smoothly interpolate between states by linear combination of the transport operators. If a portal becomes fully opaque, we can remove the edge from the graph to speed up propagation.

6 Real-time auralization

Once our pressure-response envelopes have been computed, they can be used to auralize sound sources with consistent reverberation

Algorithm 1 Graph-based topological sound propagation

```

// * denotes a convolution.
routes ← traverse(graph)
routes → split()
routes → unique()
for Route ∈ routes do
  Route → reorder(sort)
  TOPRoute ← new Route
  for pair < TOPi, TOPi+1 > ∈ Route do
    TOPimp → setID(generateID(TOPi*TOPi+1))
    if TOPimp ∉ ConvolutionCache then
      TOPimp ← TOPi * TOPi+1
      ConvolutionCache → add(TOPimp)
    end if
  TOPRoute ← TOPR * ConvolutionCache → get(TOPimp)
end for
end for

```

effects. For auralization, the decay envelopes are transformed back to time-domain and used as the envelope of a noise process, similar to artificial reverberators. To optimize the reverberation process, we use a progressive Fourier-domain approach, similar to [Tsingos 2005], to perform block-based convolution of the input audio signals with the reverberation filters.

The next sections detail our spatial audio rendering model and our implementation of a global pipeline, shown in Figure 5, where energy propagation and spatial audio processing are performed asynchronously. A key feature of this pipeline is to decouple the propagation performed from room to room from the actual sources that need to be processed, which can be much larger.

6.1 Spatial audio rendering model

Consider a sound source in room R_S that is audible in the listener's room R_L . We auralize three components: 1) the direct sound to the listener, 2) the directional incoming contributions through each portal in the listener's room, and 3) the reverberated contributions including the reverberation effect of the listening room. Our current implementation uses binaural rendering.

The *direct sound* to the listener is rendered using positional audio processing. Our current framework uses Head Related Transfer Functions (HRTF) from the *LISTEN* database (<http://www.ircam.fr/listen>). If the sound source is not in the listener's room, we use the shortest route from the graph propagation to derive an approximate diffracted component, which we render from the centroid of the portal, using a simplified Geometrical Theory of Diffraction formalism [Tsingos et al. 2001].

For each portal P_i in the listening room R_L , we render an additional *directional source* located at the centroid of P_i . The source signal, in this case, is first convolved by a *monophonic* reverberation filter and then positioned using the corresponding HRTF. The decay envelope of the reverberation filter corresponds to the *sum* of all $SR_S P * PRP * \dots * PRP_i$ routes ending at this portal. The obtained decay envelope is then converted back to time-domain to be used by our scalable artificial reverberation engine. This involves N_P inverse Fourier transforms for the N_P portals in the current listener's room.

Finally, we render *spatially diffuse* reverberated components for both the direct sound and the portal contributions. We convolve the signal of the source with a 2-channel random noise process, whose decay envelope includes the reverberation effect of the listener's room R_L . For sources located in R_L , this is directly a pre-

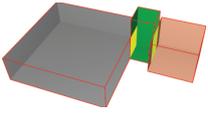
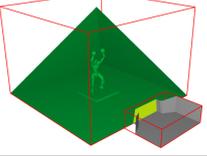
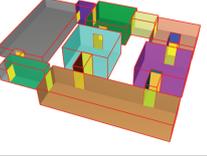
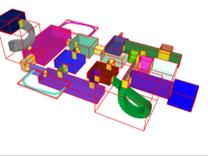
								
Traversal Time (msecs)	0.2		0.07		7.25		1.71	
Route Construction Time (msecs)	0.36		0.05		12.93		9.32	
#Routes	23		9		256		59	
Convolution Cache	off	on	off	on	off	on	off	on
#Convolutions	123	22	43	10	1819	420	852	62
Convolutions Time (msecs)	2.82	1.79	1.19	0.83	40.23	37.65	15.69	6.73
Compression ratio	–	5.6:1	–	4.3:1	–	4.3:1	–	13.7:1
Cache size (Kb)	–	432	–	208	–	7200	–	1152
Total Time (msecs)	3.38	2.35	1.31	0.95	60.41	57.83	26.72	17.76

Table 2: Topological sound propagation statistics, averaged over 600 runs on an Intel Xeon 2.33Ghz.

computed SRL envelope. For portal contributions, we generate the corresponding decay envelope as the sum of all $SR_S P * PRP * \dots * PRP_i * P_i R_L L$ routes. This only involves a single additional convolution of the previously computed directional source contribution with the final $P_i R_L$ operator. The obtained envelope is also converted back to time-domain before being transferred to the artificial reverberation engine.

For multiple sources in the same room the same decay envelopes are used, thus making our approach scalable. Another advantage of separating portal and diffuse contributions is that the relative level of the various components (portal, diffuse, room) can be individually adjusted by a sound designer which gives more control on the desired effect in the context of entertainment applications.

6.2 Implementation of the rendering pipeline

We implemented our real-time audio processing pipeline using two separate processes running asynchronously (see Figure 5).

The first process handles the energy propagation and construction of reverberation decay envelopes. If the listener changes room or the portal state is modified, we re-propagate energy in the environment from the room containing the listener (see Section 5).

In the second process, the artificial reverberation engine retrieves time-domain decay envelopes for each source and applies block-based FIR convolution to the input sound signals. FIR convolution is done directly in the Fourier domain following the approach introduced in [Tsingos 2005]. This approach performs lossy block-based convolution, accounting for masking between reverberation blocks of all sources. Additionally, the engine dynamically allocates processing to blocks contributing more energy to the final mix according to a user-specified computational budget. Figure 7 illustrates how masking and budget allocation allows for controlling the computational load.

We process 1024 sample blocks at 44.1 KHz sampling-rate and use 50% overlap-add blending to avoid artefacts in the reconstruction. Hence, our audio processing thread runs at 86Hz outputting frames of 512 reconstructed samples.

To optimize processing, all our input signals are pre-stored in the Fourier frequency domain so that only a single inverse FFT is required to reconstruct an audio frame. To synthesize the reverberation filters, we use a circular array of successive white-noise blocks, which we also pre-transform into the Fourier domain. To render spatially diffuse reverberation components, we generate uncorre-

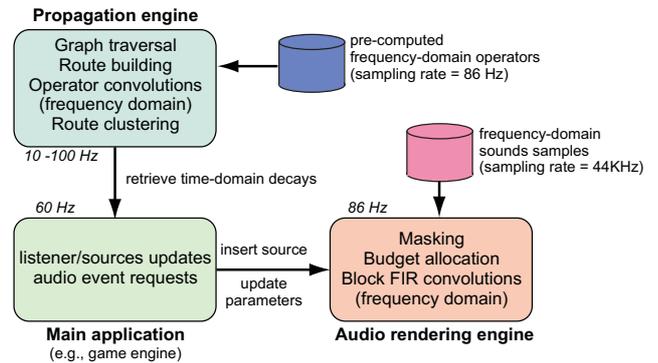


Figure 5: Overview of the rendering pipeline. Each box runs as an asynchronous process.

lated noise sequences by picking a random starting block in the array for each output channel. Each block of noise is equalized in the Fourier domain using the values stored in the time-domain decay envelopes obtained after propagation. Hence, the synthesized impulse response corresponds to a noise sequence multiplied by a frequency-dependent staircase envelope with a staircase-length equal to our frame size (11 msec. in this case). To avoid audible artefacts, we use a 15-band equalization by interpolating the 4 frequency-dependent values stored in the envelopes at each time-frame (see Section 4.2). The 15 obtained gains are directly used to multiply the complex Fourier coefficients of the noise data within each band prior to the complex multiplication with the source signal. Figure 6 illustrates our reverberation process. Although we do not use zero-padding on the source signal (thus performing circular convolution), we found that the use of a Hann window limits audible aliasing. Overlap-add reconstruction also introduces time-domain smoothing and eliminates possible artefacts due to our discontinuous staircase envelope.

7 Results

We integrated our approach in the *Penta G* game engine (<http://www.gebauz.com/index.php?page=penta-g>) developed in the context of the EU IST Gametools project (<http://www.gametools.org/>). A snapshot of the game is shown in Figure 1 (d). Please visit <http://www-sop.inria.fr/reves/projects/revGraphs/> for additional material

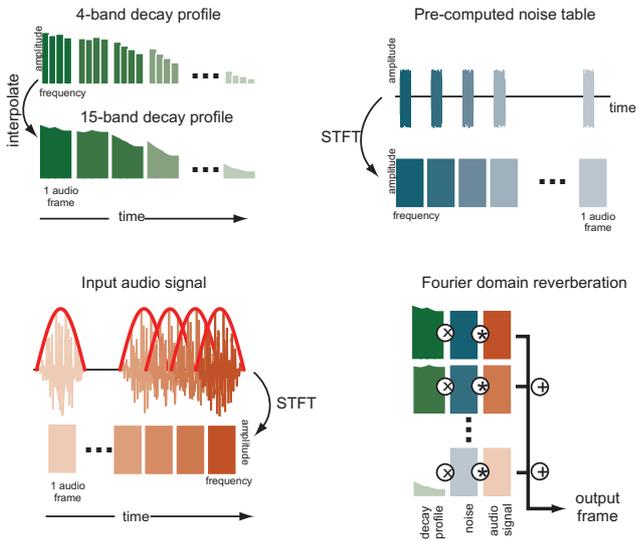


Figure 6: Illustration of our block-based, Fourier-domain, artificial reverberator. A noise sequence is first weighted by the real-valued reverberation decay profile before its complex multiplication with the input source signal. Several past blocks of the input audio signal undergo this process and add-up to construct a frame of reverberated audio signal.

including demo sounds and videos.

7.1 Performance

As can be seen in Table 2, our topological sound propagation approach can compute global reverberation filters at interactive rates for environments exhibiting complex coupling structure. Using our approach, we were able to compute the contributions of 256 propagation routes, resulting in 1800 convolutions at 17 Hz. Although the compression ratio of our current caching algorithm is very good, performance does not increase sub-linearly due to the overhead of our current algorithm. However, we believe this could be significantly improved. Clustering of the obtained decay filters according to source room and destination portal and final time-domain reconstruction were achieved at 30 Hz in this environment. Inverse Fast Fourier Transforms (IFFT) are only performed on-demand for rooms that contain the sources to auralize. One inverse FFT is required per active room and per portal in the current listener’s room.

Our scalable reverberation engine runs at 86 Hz and is able to render about a 100 individual reverberation filters at each time-frame. Figure 8 shows timings for each component of our scalable reverberation approach: masking, budget allocation and processing.

On a dual core Intel Xeon 2.33GHz workstation, our complete pipeline allows for rendering in real-time all reverberated components obtained from our propagation algorithm for 30 individual sources.

7.2 Comparison to path-tracing solutions

We compared our interactive propagation algorithm to an off-line image-source simulation implemented using a ray-casting process in different test environments. To generate these solutions 3 million rays were fired from the source and traced up to 200 orders of reflection. Valid image-sources were collected at the listening point using a 25 cm-radius detection sphere. Figures 9 and 10 show comparisons of pressure decay envelopes obtained using the image-source

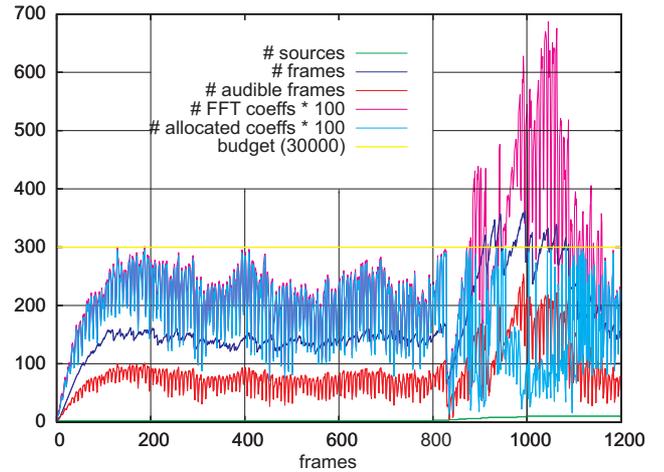


Figure 7: Statistics for our scalable reverberation engine on a sequence of 1200 audio processing frames where up to 20 sources are active. The plot displays the number of sources, number of frames of input audio data mixed for block-based convolution, and corresponding number of FFT coefficients before and after budget allocation (we set a global budget of 30000 coefficients in this example).

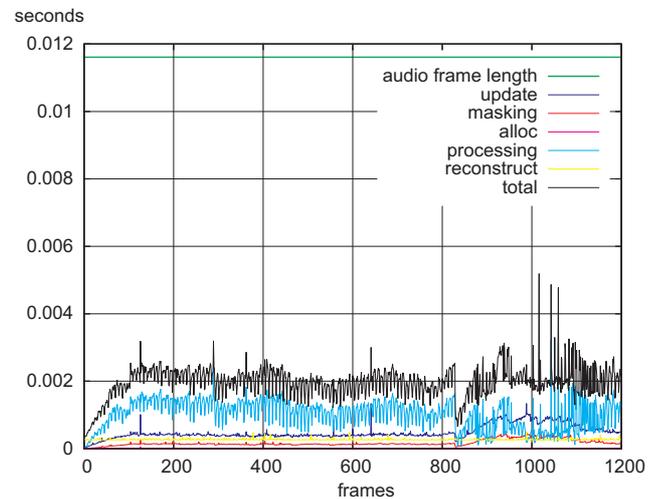


Figure 8: Breakout times for the various components of our scalable reverberation engine for the same sequence as Figure 7. Note that the masking and budget allocation introduce very little overhead.

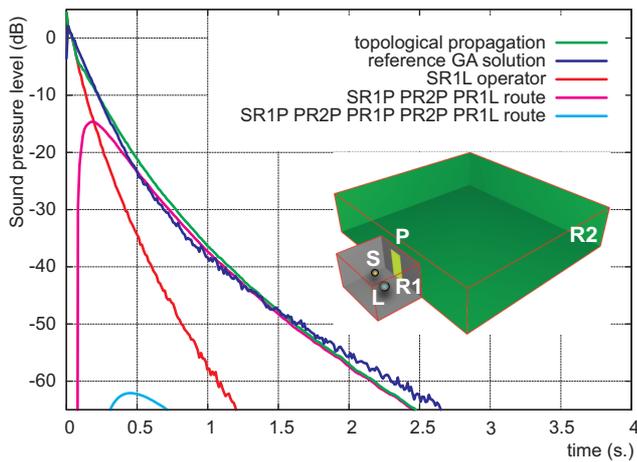


Figure 9: Comparison of our approach to an image-source simulation in a simple two-room environment. We show the result of our topological sound propagation and most representative indirect routes.

simulation and our approach. As can be seen, both approaches lead to similar decay envelopes, although our results slightly overestimates the slope of the late decay.

8 Discussion and limitations

The fundamental limitation of our approach is that portals act as diffuse emitters. As a result, our approach does not accurately model early reflections. However, it would be possible to pre-compute a room-to-room matrix of possible early reflection paths or image-sources [Lentz et al. 2007], and utilize our approach to render higher order reflections. Our current implementation only includes specular reflection and does not model diffuse reflection on surfaces, nor scattering/diffraction at portals which have been shown to significantly improve acoustical modeling [Vorländer 1995; Funkhouser et al. 2004]. Our pre-computing step could be extended to include these effects. Another expected benefit of introducing diffuse reflection would be to reduce the error introduced by our “diffuse-portal” assumption by generating more diffuse exchanges between spaces. We also currently pre-computed operators for a single representative source and listener position in each space. If the geometry of the space is complex, for instance if a room and a corridor have been merged into the same space, this might lead to unrealistic results. A solution is to define cell and portals so that the geometry of the space boundaries is convex which will limit this problem.

The performance of our approach and quality of the results is also directly influenced by the temporal sampling resolution of the decay envelopes. Sampling resolution is related to the modulation spectrum of the modeled transfer functions [Schroeder 1981]. In particular, properly taking into account early reflections would require a finer resolution or preferably a separate modeling. Our choice of a 10 msec. integration window is also directly related to the processing frame size for our auralization engine (currently running at 83 Hz). For improved efficiency, we currently use the same DT (11.6 msec.) for both the length of an output audio frame and the sampling rate of the decay envelopes. Hence, we can directly obtain the decay value at each audio frame without re-sampling the envelopes.

Currently we render our reverberation effects as decorrelated noise signals that do not capture the fine-grain spatial structure of the reverberation within each room. We believe an approach simi-

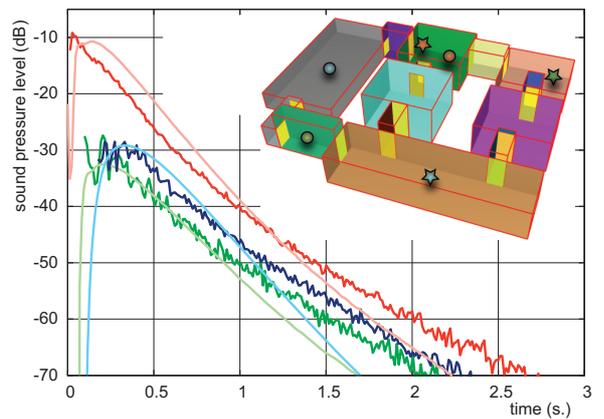


Figure 10: Comparison of our approach to an image-source simulation for different configurations of sources and listeners. Stars and spheres indicate source and listener positions, respectively. Graph colors correspond to $\langle \text{source}, \text{listener} \rangle$ pairs for which the simulations were computed. Darker colors correspond to the reference path tracing simulation while light colors are the result of our topological graph-based approach.

lar to the *Spatial Impulse Response Rendering* [Merimaa and Pulkki 2004] could be used to introduce finer grain modeling of the spatial aspects of the reverberation effect by pre-computing time-dependent average incidence directions with SRL and PRL operators.

Our current work focuses on indoor architectural environments but we believe possible application to more general situations, such as outdoor scenes (cities, etc.). These raises several issues, for instance how to define cells/portals and their connectivity graph in this context. Our assumption of diffuse energy exchanges might also not be valid in this particular case.

9 Conclusion

We presented a novel graph-based topological sound propagation algorithm that can compute interactive reverberation effects in complex coupled environments. In our auralization system we use auditory masking and scalable Fourier domain processing to render a large number of reverberated components.

Our approach supports multiple moving sources and listener. It also allows on-line topology changes and some geometrical modifications and would be suitable for procedurally-generated environment, e.g. using pre-defined building blocks as in [Merrell 2007].

Results obtained using our interactive approach compare well to off-line geometrical acoustics solutions. In the future, the results of our technique could be improved by incorporating diffuse scattering, instead of pure specular reflections. While further numerical and perceptual validation is necessary, we believe the proposed method opens interesting perspectives for real-time auralization that could find applications both for acoustical design and virtual environments.

Acknowledgments

The authors would like to thank Fanouris Moraitis for modeling the test 3D environments, Michael Wimmer and Chris Chiu for making the *Penta-G* game engine available to us and providing support, David Alloza for initial discussions on graph-based reverber-

ation and Xavier Granier. We also would like to thank the anonymous reviewers for their helpful suggestions. This research was funded by the EU FET Open project IST-014891-2 CROSSMOD (<http://www.crossmod.org>).

References

- BEGAULT, D. R. 1994. *3D Sound for Virtual Reality and Multimedia*. Academic Press Professional.
- BRADLEY, D. T., AND WANG, L. M. 2005. The effects of simple coupled volume geometry on the objective and subjective results from nonexponential decay. *J. Acoustical Soc. Am.* 118, 3, 1480–1490.
- BRADLEY, D., AND WANG, L. 2007. Comparison of measured and computer-modeled objective parameters for an existing coupled volume concert hall. *Building Acoustics* 14, 2, 79–90.
- CHRISTENSEN, C. L. Odeon Room Acoustics Software. <http://www.odeon.dk/>.
- CREMER, L., AND MÜLLER, H. 1982. *Principles and Applications of Room Acoustics*. Applied Science.
- DALENBÄCK, B.-I. CATT Acoustic. <http://www.catt.se/>.
- DALENBÄCK, B.-I. L. 1996. Room acoustic prediction based on a unified treatment of diffuse and specular reflection. *J. Acoustical Soc. Am.* 100, 2 (Aug.), 899–909.
- DAVIS, A. H. 1925. Reverberation equations for two adjacent rooms connected by an incompletely sound-proof partition. *Philos. Mag.* 50, 75–80.
- EAX, 2004. Environmental audio extensions 4.0, Creative©. <http://www.soundblaster.com/eaudio>.
- ERMANN, M. 2007. Double sloped decay: Subjective listening test to determine perceptibility and preference. *Building Acoustics* 14, 2, 91–107.
- EYRING, C. F. 1931. Reverberation time measurements in coupled rooms. *J. Acoustical Soc. Am.* 3, 2, 181–206.
- FOALE, C., AND VAMPLEW, P. 2007. Portal-based sound propagation for first-person computer games. In *The Proceedings of the 4th Australasian Conf. on Interactive Entertainment*, ACM.
- FUNKHOUSER, T., TSINGOS, N., CARLBOM, I., ELKO, G., SONDHI, M., WEST, J. E., PINGALI, G., MIN, P., AND NGAN, A. 2004. A beam tracing method for interactive architectural acoustics. *J. Acoustical Soc. Am.* 115, 2, 739–756.
- GARDNER, W. 1998. *Reverberation Algorithms, Applications of Digital Signal Processing to Audio and Acoustics, Chapter 6*. Mark Kahrs and Karlheinz Brandenburg Ed., Kluwer Academic Publishers.
- HAUMONT, D., DEBEIR, O., AND SILLION, F. 2003. Volumetric cell-and-portal generation. In *Computer Graphics Forum*, Blackwell Publishers, vol. 3-22 of *EUROGRAPHICS Conference Proceedings*.
- HEINZ, R. 1993. Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics* 38, 145–159.
- JOT, J.-M. 1999. Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems* 7, 1, 55–69.
- KNIGHT, D. 2003. *Audibility of non-exponential energy decay in running reverberation*. Master's thesis, Rensselaer Polytechnic Institute.
- KUTTRUFF, H. 1973. *Room Acoustics (1st edition)*. Elsevier Applied Science.
- KUTTRUFF, K. 1993. Auralization of impulse responses modeled on the basis of ray-tracing results. *J. Audio Eng. Soc.* 41, 11 (Nov.), 876–880.
- LAINE, S., SILTANEN, S., LOKKI, T., AND SAVIOJA, L. 2008. Accelerated beam tracing algorithm. *Applied Acoustics (to appear)*. Available online at <http://dx.doi.org/10.1016/j.apacoust.2007.11.011>.
- LAUTERBACH, C., CHANDAK, A., AND MANOCHA, D. 2007. Interactive sound rendering in complex and dynamic scenes using frustum tracing. *Transactions on Visualization and Computer Graphics* 13, 6 (Nov.-Dec.), 1672–1679.
- LEFEBVRE, S., AND HORNUS, S. 2003. Automatic cell-and-portal decomposition. Tech. Rep. 4898, INRIA, July.
- LENTZ, T., SCHRÖDER, D., VORLÄNDER, M., AND ASSENMACHER, I. 2007. Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal on Advances in Signal Processing* 2007, Article ID 70540, 19 pages. doi:10.1155/2007/70540.
- LYLE, C. D. 1981. The prediction of steady state sound levels in naturally coupled enclosures. *Acoust. Lett.* 5, 16–21.
- LYLE, C. D. 1981. Recommendation for estimating reverberation time in coupled spaces. *Acoust. Lett.* 5, 35–38.
- MERIMAA, J., AND PULLKI, V. 2004. Spatial impulse response rendering. *Proc. of the 7th Intl. Conf. on Digital Audio Effects (DAFX'04), Naples, Italy* (Oct.).
- MERRELL, P. 2007. Example-based model synthesis. In *I3D '07: Proceedings of the 2007 symposium on Interactive 3D graphics and games*, ACM, New York, NY, USA, 105–112.
- PICARD, D. 2003. Audibility of non-exponential reverberation decays. Rapport de Stage D'Option Scientifique, École Polytechnique, Paris, France.
- POPE, J., CREASEY, D., AND CHALMERS, A. 1999. Realtime room acoustics using ambisonics. In *The Proceedings of the AES 16th Intl. Conf. on Spatial Sound Reproduction*, Audio Engineering Society, 427–435.
- ROCCHESO, D. 2002. *Spatial Effects, DAFX - Digital Audio Effects, Chapter 6*. Udo Zölzer Ed., Wiley.
- SCHRÖDER, D., DROSS, P., AND VORLÄNDER, M. 2007. A fast reverberation estimator for virtual environments. In *Proceedings of the AES 30th Intl. Conf., Saariselkä, Finland*.
- SCHROEDER, M. 1962. Natural sounding artificial reverberation. *J. Audio Eng. Soc.* 10, 3, 219–223.
- SCHROEDER, M. 1981. Modulation transfer function: Definition and measurement. *Acustica* 49, 179–182.
- SILTANEN, S., LOKKI, T., KIMINKI, S., AND SAVIOJA, L. 2007. The room acoustic rendering equation. *J. Acoustical Soc. Am.* 122, 3 (Sept.), 1624–1635.
- SUMMERS, J., TORRES, R., AND SHIMIZU, Y. 2004. Statistical-acoustics models of energy decay in systems of coupled rooms and their relation to geometrical acoustics. *J. Acoustical Soc. Am.* 116, 2, 958–969.
- SUMMERS, J., TORRES, R., SHIMIZU, Y., AND DALENBÄCK, B.-I. 2005. Adapting a randomized beam-axis-tracing algorithm to modeling of coupled rooms via late-part ray tracing. *J. Acoustical Soc. Am.* 118, 3, 1491–1502.
- SVENSSON, U. P. 1998. Energy-time relations in a room with an electroacoustic system. *J. Acoustical Soc. Am.* 104 (Sept.), 1483–1490.
- TELLER, S. 1992. *Visibility Computations in Densely Occupied Polyhedral Environments*. PhD thesis, Computer Science Div., University of California, Berkeley.

- TSINGOS, N., FUNKHOUSER, T., NGAN, A., AND CARLBOM, I. 2001. Modeling acoustics in virtual environments using the uniform theory of diffraction. *ACM Computer Graphics, SIGGRAPH'01 Proceedings* (Aug.), 545–552.
- TSINGOS, N. 2005. Scalable perceptual mixing and filtering of audio signals using an augmented spectral representation. In *Proceedings of the Intl. Conf. on Digital Audio Effects*. Madrid, Spain.
- VORLÄNDER, M. 1989. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoustical Soc. Am.* 86, 1, 172–178.
- VORLÄNDER, M. 1995. International round robin on room acoustical computer simulations. In *Proc. of the Internat. Congress on Acoustics (ICA'95), Trondheim, Norway*, 689–692.