DETECTION AND ESTIMATION OF ARRIVALS IN ROOM IMPULSE RESPONSES BY GREEDY SPARSE APPROXIMATION

*Bob L. Sturm and Guillaume Defrance**

Department of Architecture, Design and Media Technology
Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, Denmark
phone: (+45) 99 40 76 33
email: bst@imi.aau.dk

*Department of Architecture
Crookesmoor Building, Conduit Road
University of Sheffield, Sheffield S10 1FL, U.K.
phone: +44 (0) 114 22 2 03 68
email: defrance.all@gmail.com

ABSTRACT

We investigate the use of greedy sparse approximation for facilitating the time-domain analysis of room impulse responses (RIRs), specifically locating the times and amplitudes of arrivals to not long after the upper bound of the “mixing time,” i.e., the time after which there exists in theory the same number of sound rays per unit volume throughout the room. We compare the performance of two methods of greedy sparse approximation — matching pursuit (MP) and orthogonal MP (OMP) — for estimating arrival times and amplitudes. By using RIRs generated from a stochastic model, we quantify the performance of each estimator using dynamic time warping to optimally pair estimated and true arrivals. We find OMP significantly outperforms MP in estimating both the arrival times and amplitudes, and having fewer erroneous and duplicated arrivals.

1. INTRODUCTION

One measures a room impulse response (RIR) in order to quantitatively study the acoustical characteristics of the room [1]. To measure an RIR in practice, one emits a wideband *source*, e.g., spark gun, into a room, and measure local changes in pressure at specific locations with *receivers*, e.g., microphones. We can model a RIR in the high frequency domain by representing sound as rays that leave a source at the speed of sound, and undergo reflections at boundaries before arriving at each receiver. An *arrival* in this model is a sound ray emitted by the source that has undergone at least one reflection during its journey to the position of the receiver. A relationship between the expected number arrivals received t seconds after excitation is embodied in the relationship [1]

$$\mu_A(t) = \frac{4\pi c^3}{3V} t^3 \tag{1}$$

where V is the volume of the room in cubic meters, and c is the speed of sound in meters per second. We wish to accurately detect arrivals (both their times and amplitudes) in RIRs to study the acoustical characteristics of the room, for instances its volume, its coloration, etc.

Given a measured RIR $r(t)$ assumed to be a superposition of arrivals, we wish to find the time and amplitude of each arrival. This problem has been addressed by a few methods within the field of room acoustics. One approach uses an adaptive thresholding technique [2], which requires empirically testing a range of variables to make the detection algorithm give results assumed to be reasonable. Furthermore, this approach [2] essentially detects local time-domain peaks in the RIR, and then equates those to arrivals. A different approach [3, 4] uses a method of greedy sparse approximation to first decompose the RIR as a linear combination of the estimated direct sound, and then to detect arrivals in a domain more sparse than the original RIR. However, this approach is sensitive to the parameters of the sparse approximation algorithm [3].

In this paper, we quantitatively look at the performance of two estimators of arrival times and amplitudes in RIRs created using methods of greedy sparse approximation [5]: matching pursuit (MP), and orthogonal MP (OMP). For our tests, we generate data using a stochastic model of RIRs [6] — that has previously been validated using various acoustical criteria [3, 7] — with a variety of source-receiver distances, and noise conditions. We calculate the errors in arrival time estimation using a dynamic time warping approach to optimally match the detected arrivals to the true arrivals with respect to three path policies having the same cost. In this way, since we have a ground truth, and because we know the direct sound, we find the best possible performance to be expected for detecting arrivals from real RIRs using these two greedy methods for sparse approximation. Our results conclusively show that OMP provides a much better estimator of arrival times and amplitudes than does MP.

We emphasize that we are interested in detecting arrivals to not long after the *mixing time* of a room, which is the time after which a dynamical system, e.g., a concert hall, is *mixed*. In other words, the mixing time is the time it takes for all sound rays in a room to have the same probability to reach any phase point of the phase space at any time. In room acoustics, the mixing time defines the time region when the early arrivals transition to the late reverberation [4]. We use a heuristic formulation of the mixing time as the time when we expect at any location 10 arrivals within 24 ms, which is described by the relation [8]:

$$T_{\text{mix}} \approx \sqrt{V} \text{ ms}. \tag{2}$$

In this work, we limit ourselves to estimating arrival times not long after the mixing time because that is when the number of arrivals is becoming so large that estimates will likely be poor using greedy sparse approximation [3, 4].

2. ARRIVAL ESTIMATION BY GREEDY SPARSE APPROXIMATION

We can model ideally the arrivals at one location a distance $d \geq 0$ meters from the source by

$$h(t) = A_0\delta(t - d/c) + \sum_{k=1}^{\infty} A_k \delta(t - T_k) \tag{3}$$

where $\mathcal{A} = \{A_k : \mathbb{R} \to \mathbb{R}\}_{k\in\mathbb{N}}$ and $\mathcal{T} = \{T_k \geq T_{k-1} > d/c : \mathbb{R} \to \mathbb{R}\}_{k\in\mathbb{N}}$ are non-stationary and dependent random processes describing the behaviors of arrival amplitudes and times, respectively. The value $A_0^2 \propto d^{-2}$ scales the energy of the *direct sound*, or the unreflected source signal. Since real sources are band-limited signals, we can synthesize an RIR by convolving $h(t)$ with a band-limited excitation $d(t)$:

$$r(t) = (h \star d)(t) + n(t). \tag{4}$$

where $n(t)$ is stationary and independently distributed zero mean white noise with power $\sigma_n^2 \geq 0$. (Note that in this synthesis we do not consider any filtering of sound rays by the boundaries of the room.) If $d(t) = \delta(t - \tau)$ and $\sigma_n^2 = 0$ in (4), we can trivially determine the arrival amplitudes $\mathcal{A}$ and times $\mathcal{T}$ from $r(t)$ since all

© EURASIP, 2010 ISSN 2076-1465

i	t_i	T_i		t_i	T_i		t_i	T_i
1	141	141		141	141		141	141
2	2015	2016		2015	2016		2015	2016
3	2967	2155		2015	2155		–	2155
4	3004	2966	DTW $\Longrightarrow$	2967	2966	PAIR $\Longrightarrow$	2967	2966
5	3128	3128		3004	2966		3004	–
6	4049	3895		3128	3128		3128	3128
7		4049		4049	3895		–	3895
8				4049	4049		4049	4049

Table 1: Example of how we generate unique pairings of true $\{T_i\}$ and estimated $\{t_i\}$ arrival times. The times at left are paired optimally (center) by dynamic time warping with respect to the absolute time difference (16). We remove repetitions to enforce pairings to be unique or absent (right).

we need to do is find the times and values where $r(t) > 0$. Otherwise we must use a different method, such as localized thresholding [2], or greedy sparse approximation [3,4].

Let us define a *dictionary* $\mathcal{D}$ that contains all translations of the estimated direct sound $\tilde{d}(t)$; and define the *amplitude compensated RIR* as

$$\hat{r}(t) \overset{\Delta}{=} e^{\beta t} r(t) \qquad (5)$$

for some specified $\beta \geq 0$, assuming the excitation occurs at $t = 0$. With this scaling we attempt to compensate the natural exponential decay of power in the signal so that we can focus on finding arrivals throughout the RIR independent of their powers. We wish to express $\hat{r}(t)$, which we write as the vector $\mathbf{r}$, as a linear combination of n elements from $\mathcal{D}$:

$$\mathbf{r} = \sum_{i=1}^{n} a_i \mathbf{b}_i + \mathbf{e}(n) = \mathbf{B}(n)\mathbf{a}(n) + \mathbf{e}(n) \qquad (6)$$

where each $\mathbf{b}_i \in \mathcal{D}$ is a translation of $\tilde{d}(t)$, $\mathbf{B}(n) = [\mathbf{b}_1|\mathbf{b}_2|\cdots|\mathbf{b}_n]$, a_i is a weight and the ith element of $\mathbf{a}(n)$, and $\mathbf{e}(n)$ is the nth-order error signal (residual). We define the nth-order representation of $\mathbf{r}$ as the set of elements $\mathcal{R}_n = \{\mathbf{B}(n), \mathbf{a}(n), \mathbf{e}(n)\}$.

Matching pursuit (MP) [5] is a greedy iterative descent approach to sparse approximation that updates $\mathcal{R}_n$ by

$$\mathcal{R}_{n+1} = \left\{ \begin{array}{l} \mathbf{B}(n+1) = [\mathbf{B}(n)|\mathbf{b}_{n+1}], \\ \mathbf{a}(n+1) = [\mathbf{a}^T(n), \mathbf{b}_{n+1}^T \mathbf{e}(n)/||\mathbf{b}_{n+1}||_2^2]^T, \\ \mathbf{e}(n+1) = \mathbf{r} - \mathbf{B}(n+1)\mathbf{a}(n+1) \end{array} \right\} \qquad (7)$$

using the atom selection criterion

$$\mathbf{b}_{n+1} \overset{\Delta}{=} \arg\max_{\mathbf{d}\in\mathcal{D}} |\mathbf{d}^T \mathbf{e}(n)|/||\mathbf{d}||_2. \qquad (8)$$

An alternative to MP is Orthogonal MP (OMP) [5], which retains the atom selection criterion of MP, but updates the representation by

$$\mathcal{R}_{n+1} = \left\{ \begin{array}{l} \mathbf{B}(n+1) = [\mathbf{B}(n)|\mathbf{b}_n], \\ \mathbf{a}(n+1) = \mathbf{B}^{\dagger}(n+1)\mathbf{r}, \\ \mathbf{e}(n+1) = \mathbf{r} - \mathbf{B}(n+1)\mathbf{a}(n+1) \end{array} \right\} \qquad (9)$$

where $\mathbf{B}^{\dagger}(n) \overset{\Delta}{=} [\mathbf{B}^T(n)\mathbf{B}(n)]^{-1}\mathbf{B}^T(n)$ is the psuedoinverse of $\mathbf{B}(n)$. This optimization guarantees that the nth-order residual will be orthogonal to the nth-order representation basis $\mathbf{B}(n)$.

Finally, given an nth-order representation $\mathcal{R}_n$ of an amplitude compensated RIR $\hat{r}(t)$, we can estimate n arrival times and amplitudes with the sets [3,4]

$$\widetilde{\mathcal{T}_n} \overset{\Delta}{=} \{\tau(\mathbf{b}_i) : \mathbf{b}_i \in \mathbf{B}(n), i = 1, \dots, n\} \qquad (10)$$

$$\widetilde{\mathcal{A}_n} \overset{\Delta}{=} \{[\mathbf{a}(n)]_i : i = 1, \dots, n\} \qquad (11)$$

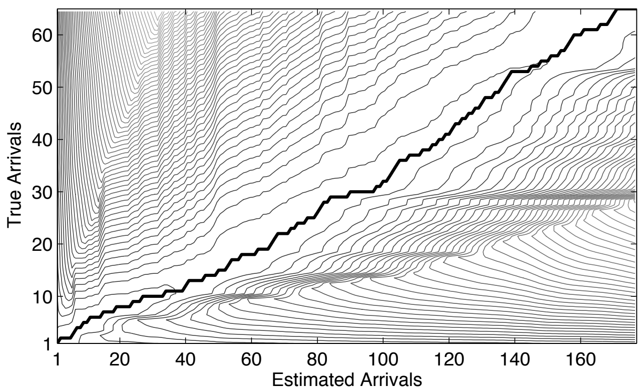


Figure 1: Best path (thick black) matching a set of 65 true and 178 estimated arrival times. Contours denote distance surface.

where $\tau(\mathbf{b}_i)$ gives the time translation of the estimated direct sound in $\mathbf{b}_i$ from the dictionary $\mathcal{D}$, and $[\mathbf{a}(n)]_i$ is the ith element of $\mathbf{a}(n)$. Note that with both MP (7) and OMP in (9) the amplitudes can be negative and positive, and thus allow for modeling arrivals as inversions of the direct sound.

Given the sets of sorted estimated and true arrival times and amplitudes produced by either method of sparse approximation

$$\widetilde{\mathcal{T}_n} = \{t_1, t_2, \dots, t_n : t_{m+1} \geq t_m\} \qquad (12)$$

$$\widetilde{\mathcal{A}_n} = \{a_1 e^{-\beta t_1}, a_2 e^{-\beta t_2}, \dots, a_n e^{-\beta t_n}\} \qquad (13)$$

$$\mathcal{T} = \{T_1, T_2, \dots, T_N : T_{m+1} \geq T_m\} \qquad (14)$$

$$\mathcal{A} = \{A_1, A_2, \dots, A_N\} \qquad (15)$$

we now define a measure with which we gauge the quality of the estimations. Note that we rescale the estimated amplitudes by applying the inverse of (5). We first find the optimal pairing of members from each set of arrival times by performing dynamic time warping with three simple path policies, each carrying the same cost: 1) $(1,0)$, i.e., skip true arrival; 2) $(0,1)$, i.e., skip estimated arrival; and 3) $(1,1)$, i.e., match estimated with true arrival. We use the distance measure defined as the absolute time difference:

$$D(t_i, T_j) \overset{\Delta}{=} |t_i - T_j|. \qquad (16)$$

This process returns the best path through the two sets of arrival times, an example of which is shown by Fig. 1. From the pairings of these times, we then enforce unique pairs such that each estimated arrival time is uniquely paired to one true arrival time, or none at all; and similarly, that each true arrival time is paired to one estimated

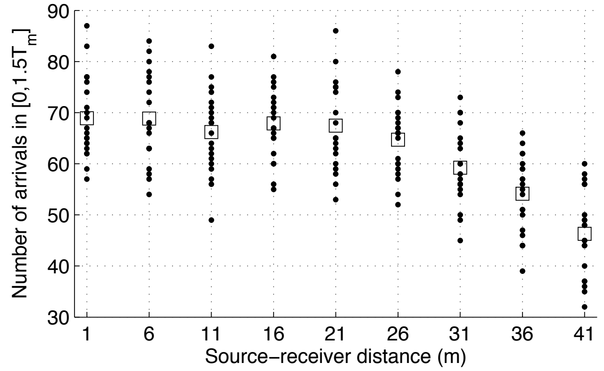
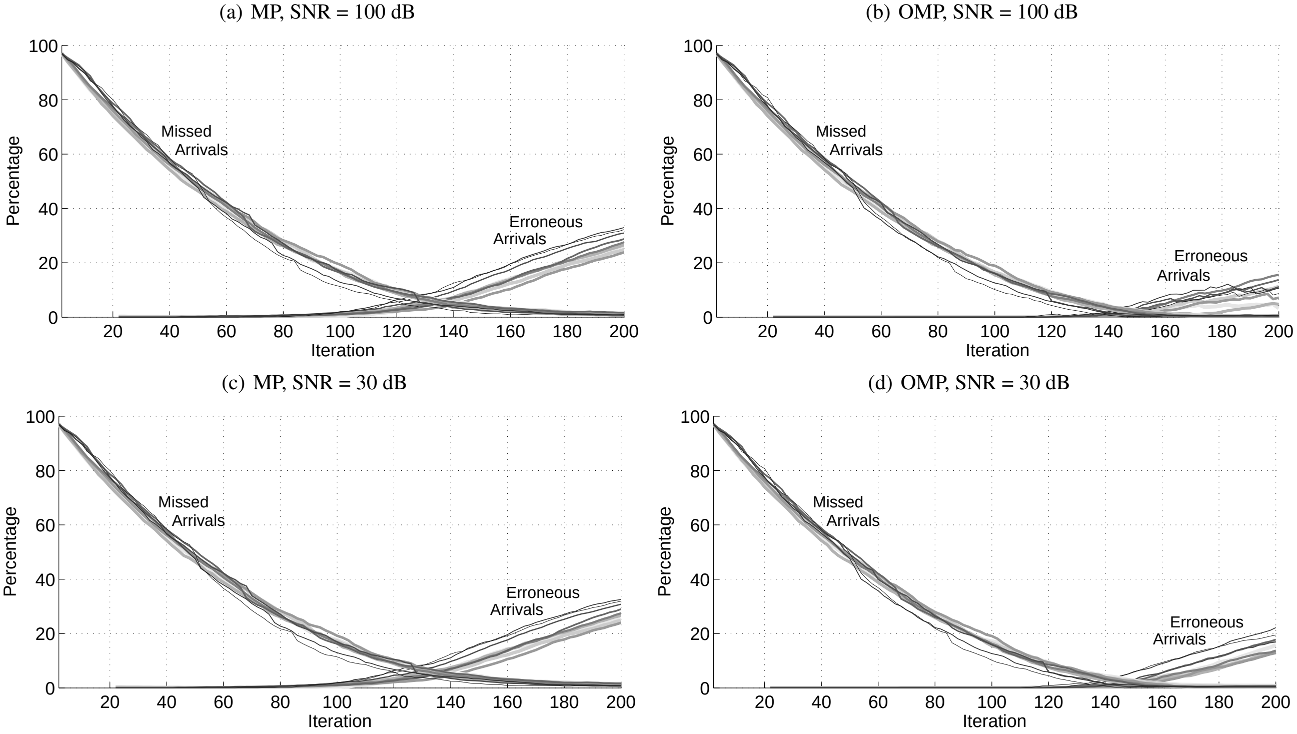


Figure 2: Number of arrivals in $[0, 1.5T_{\text{mix}}]$ for each realization of each source-receiver distance. Squares denote means.


(a) MP, SNR = 100 dB

(b) OMP, SNR = 100 dB

(c) MP, SNR = 30 dB

(d) OMP, SNR = 30 dB

Figure 3: Mean percentages of true arrivals missing from those estimated, and of estimated arrivals that are erroneous or duplicated, as a function of decomposition iteration, or model order, for both MP and OMP. Time region analyzed is $[0, 1.5T_{\text{mix}}]$.

arrival time, or none at all. We demonstrate this process in Table 1 for an RIR with 7 true arrivals in which we have estimated 6 arrivals. The right portion of this table shows the best pairing with respect to the distance measure (16). Here we see that two of the seven true arrival times are missing from the six estimated arrival times, and one of the six arrival times is either a repetition or an erroneous arrival. From these pairings, we can calculate, for instance: the percentage of true arrivals missing from those estimated (which for the example in Table 1 is $2/7$); and the percentage of erroneous and replicated arrivals in those estimated ($1/6$ in the same example).

3. COMPUTER SIMULATIONS

We now describe how we synthesize RIRs to generate testing data with known arrival times and amplitudes, and then present the results of several simulations using MP and OMP for estimating the arrivals in this synthetic data.

3.1 RIR Test Data Synthesis

We synthesize a set of 180 RIRs from realizations of a stochastic process that models the arrivals and their amplitudes at one receiver location for a room of a given volume, mean absorption coefficient, and reverberation time [6]. It has been shown that the number of arrivals in a time of duration $\Delta \geq 0$ at a time $t > 0$ after the excitation can be modeled as a non-stationary Poisson process where the probability of $k \geq 0$ arrivals is given by [6]

$$P[N(t,\Delta) - N(t,0) = k] = \frac{[\lambda(t,\Delta)\Delta]^k e^{-\lambda(t,\Delta)\Delta}}{k!} \qquad (17)$$

(and zero for $k < 0$), and with Poisson parameter given by

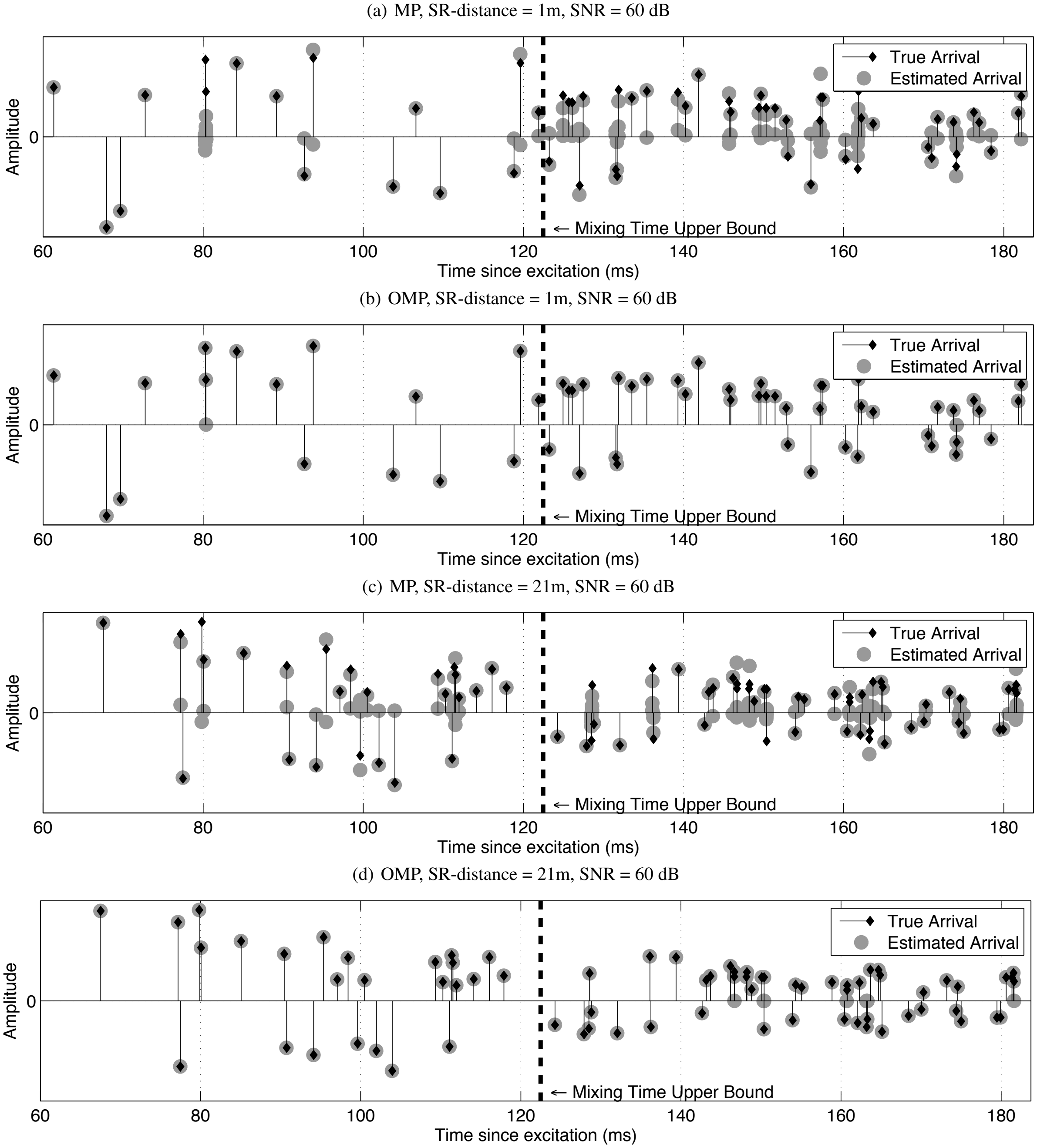
$$\lambda(t,\Delta) = \mu_A(t+\Delta) - \mu_A(t) = \frac{4\pi c^3}{3V}(\Delta^3 + 3t^2\Delta + 3t\Delta^2) \qquad (18)$$

using the mean number of arrivals in (1). We model the magnitude of an arrival as a random process, with parameters that are related to the room volume, mean absorption coefficient, and reverberation time. The sign of the arrival amplitude is a discrete binary random variable distributed on ± 1 with equal probability mass [9]. This stochastic model of RIRs has been validated on a set of 8 concert halls by comparing sets of room acoustical indices [3, 7]. With this stochastic model, we generate 20 realizations of $h(t)$ in (3) for each of nine source-receiver distances ($\{1, 6, 11, \ldots, 41\}$ m) for a room with $V = 15000$ m^3, reverberation time $R_T = 1.5$ s, and mean absorption coefficient $\alpha = 0.3$, at a sampling rate of 48 kHz. We generate each RIR $r(t)$ in (4) by setting $d(t)$ (and $\tilde{d}(t)$) to be the direct sound recorded from a pistol shot, and then adding white Gaussian noise at a specific power. The distribution of the numbers of arrivals within $[0, 1.5T_{\text{mix}}]$ for each of these realizations are shown in Fig. 2. We generate each amplitude compensated RIR by using $\beta = 4.8$, set such that (5) appears to roughly compensate the natural power decay of each RIR.

3.2 Estimation Results

In Fig. 3 we show the results of using MP and OMP for estimating arrival times with respect to the mean percentage of true arrivals missing from, and the number of erroneous and replicated arrivals, in those estimated. For each source-receiver distance we average the results from all 20 realizations. We see that for both MP and OMP, our estimations miss fewer arrivals as we push the model order in (6) higher, but our detected arrivals begin to include more erroneous and repeated arrivals. It is clear that both of these figures of merit are jointly minimized at higher model orders for OMP than for MP at a high SNR: between orders 135 and 155 for OMP, and between orders 120 and 137 for MP. These orders do not change significantly even with low SNR (tested to 15 dB — which is low considering that the standardized SNR for measuring room acoustics characteristics is at least 30 dB), demonstrating the noise robustness of this approach to greedy sparse approximation. However, it is clear that of the two estimations OMP provides the best results in terms of capturing true arrivals, and suppressing erroneous and repeated arrivals.

Figure 4 shows the estimations of the arrivals (both in time and


(a) MP, SR-distance = 1m, SNR = 60 dB

(b) OMP, SR-distance = 1m, SNR = 60 dB

(c) MP, SR-distance = 21m, SNR = 60 dB

(d) OMP, SR-distance = 21m, SNR = 60 dB

Figure 4: Comparisons of true and estimated arrival times (abscissa) and relative amplitudes (ordinate) using MP and OMP for two source-receiver distances. Time and amplitude of direct sound is not shown.

amplitude) in two RIRs at different source-receiver distances (SR-distance) using MP and OMP at 60 dB SNR. Note that we show the estimated relative amplitudes of each arrival as well (below the center line is negative amplitude). In Fig. 4(a) we see that MP detects numerous arrivals around 80 ms where there exists two true arrivals that are nearly overlapping. We see this same behavior at several times after the mixing time. Figure 4(b) shows for the same RIR that this behavior has nearly disappeared in the estimates found by OMP. Additionally, the amplitudes of all arrivals are estimated much better. We see the same results in Fig. 4(c-d) for an RIR synthesized for a source-receiver distance of 21m.

3.3 Discussion

Essentially, the greedy sparse approximation methods MP and OMP attempt to recover $h(t)$ from $r(t)$ in (4) by approximating its deconvolution given an estimated direct sound $\tilde{d}(t)$, and by making the assumption that $h(t)$ can be sparsely represented as a linear combination of a finite number of discrete translations of $\tilde{d}(t)$. One might attempt to perform a deconvolution by least squares minimization of the norm residual, i.e.,

$$\mathbf{r} = \begin{bmatrix} \tilde{\mathbf{d}} & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{d}} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \tilde{\mathbf{d}} \end{bmatrix} \mathbf{a} + \mathbf{n} = \mathbf{D}\mathbf{a} + \mathbf{n} \qquad (19)$$

$$\Longrightarrow \mathbf{a}_{\text{LS}} = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\mathbf{r} = \mathbf{D}^{\dagger}\mathbf{r} \qquad (20)$$

where $\tilde{\mathbf{d}}$ is a column vector of the estimated direct sound, $\mathbf{n}$ is unknown white Gaussian noise, and $\mathbf{D}$ is a full-column rank matrix

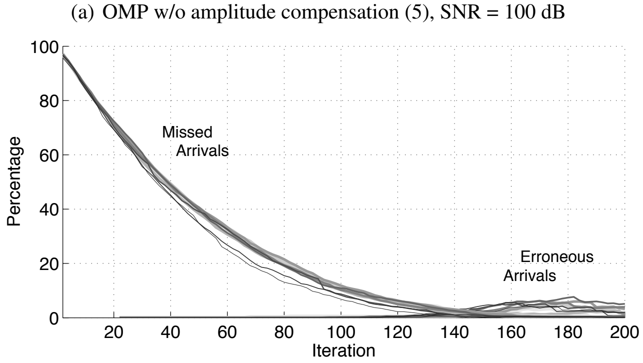


Figure 5: Without amplitude compensation (5), the mean percentages of true arrivals missing from those estimated, and of estimated arrivals that are erroneous or duplicated, as a function of decomposition iteration, or model order, for OMP. Time region analyzed is $[0, 1.5T_{\text{mix}}]$.

composed of all possible discrete translations of $\tilde{\mathbf{d}}$. However, not only will this solution be affected by noise (the degree to which depends on the amount $\mathbf{n}$ points into the column space of $\mathbf{D}$), the solution will likely not be sparse since it only guarantees by construction to minimize the error $||\mathbf{r} - \mathbf{D}\mathbf{a}||_2$, e.g., the variance of the estimated noise signal, and not the number of non-zero components of $\mathbf{a}$, i.e., its sparsity. In such a scenario, the resulting "activation vector" $\mathbf{a}_{\text{LS}}$ will not be nearly as sparse as 4. As a consequence then, we would have to perform additional processing, such as thresholding [2], to find and estimate the true arrivals using all the elements of $\mathbf{a}_{\text{LS}}$ that are not zero.

By using greedy sparse approximation for *n*-iterations, however, we obtain an *n*-sparse solution $\mathbf{a}(n)$ that, while not guaranteed to minimize the error $||\mathbf{r} - \mathbf{D}\mathbf{a}||_2$, is sparse by construction, and thus removes the need to apply thresholding to find arrivals because we can consider everything that is not zero to be an arrival. The assumption here is that after *n*-iterations the greedy sparse approximation method has found *n* arrivals. Thus the thorny issue remains of deciding when to stop the decomposition process, or in other words, how to choose the best model order in (6). Previous work [4] has looked at stopping the decomposition process using room acoustical indices of the reconstructed RIR, and perceptual listening tests; but these essentially consider the precision of the time-domain approximation, and not the model of the underlying structure of the RIR. Figure 3 tells us to stop the decomposition when the two curves are at their joint minimum — but this requires knowledge of the true arrivals. Regardless, our work here provides insight into only how well greedy sparse approximation methods can detect arrivals in a measured RIR — even when there is a significant amount of noise — in the best case scenario, i.e., we know the direct sound, and that there is no filtering at reflection boundaries.

4. CONCLUSION

We have investigated the differences between two methods of greedy sparse approximation, specifically MP and OMP, in detecting arrivals in RIRs and estimating their arrival times and amplitudes, up to a time not long after the mixing time. We clearly see that both methods are highly robust to SNR, but that OMP performs significantly better than MP with respect to estimating both the time and amplitude of each arrival in our set of simulated RIRs. This performance difference is due to the ability of OMP to modify the amplitudes in the model each time it adds a new element, i.e., the use of orthogonal projection in (9). In further experiments, we found that preprocessing the RIRs with the amplitude compensation in (5) appears to reduce the accuracy of arrival detection using OMP. We see a much smaller percentages of replicated and missing arrivals in Fig. 5 for two noise levels that those seen in Fig. 3, which uses amplitude compensation before decomposition. This suggests that, at least in our relatively synthetic experiements, we need not worry about the effects of the natural decay of an RIR in the atom selection criterion of greedy approximation methods, e.g., (8) for MP and OMP.

Our current work concerns the problem of order selection in the model (6); comparing these greedy approaches to estimation with those using other techniques, such as, localized time-domain thresholding [2]; extracting the direct sound from the RIR itself; and how to incorporate the filtering occurring at room boundaries into the dictionary. We are also considering the impact on these results of the "corrections" that are inherent to greedy sparse approximation methods [10].

5. ACKNOWLEDGMENTS

The work by G. Defrance was supported in part by Marie Curie TOK No. 042656.

REFERENCES

[1] L. Cremer, H.A. Muller, and T.J. Schultz, *Principles and Applications of Room Acoustics*, vol. 1, Applied Science Publishers Ltd, London and New York, 1982.

[2] M. Kuster, "Reliability of estimating the room volume from a single room impulse response," *J. Acoust. Soc. of Am.*, vol. 124, no. 2, pp. 982–993, Aug. 2008.

[3] G. Defrance, L. Daudet, and J-D. Polack, "Using Matching Pursuit for estimating mixing time within Room Impulse Responses," *Acta Acustica united with Acustica*, vol. 95, no. 6, pp. 1082–1092, Dec. 2009.

[4] G. Defrance, *Characterization of mixing within room impulse responses. Application to the experimental estimation of the mixing time.*, Ph.D. thesis, Université Pierre et Marie Curie, Paris 6, Paris, France, Nov. 2009.

[5] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, Academic Press, Elsevier, Amsterdam, The Netherlands, 3rd edition, 2009.

[6] J-D. Polack, "Playing Billiards in the Concert Hall: The Mathematical Foundations of Geometrical Room Acoustics," *Applied Acoustics*, vol. 38, pp. 235–244, Feb. 1993.

[7] J-D. Polack, "Reverberation time and mean absorption in concert halls," in *Proc. Institute of Acoustics*, Copenhagen, Denmark, May 2006, vol. 28.

[8] J-D. Polack, *La transmission de l'énergie sonore dans les salles*, Ph.D. thesis, Thèse de doctorat d'Etat, Université du Maine, Le Mans, France, Dec. 1988.

[9] H. Kuttruff, "Auralization of Impulse Responses Modeled on the Basis of Ray-Tracing Results," *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 876–880, Nov. 1993.

[10] B. L. Sturm and J. J. Shynk, "Sparse approximation and the pursuit of meaningful signal models with interference adaptation," *IEEE Trans. Acoustics, Speech, Lang. Process.*, vol. 18, no. 3, pp. 461–472, Mar. 2010.