A NUMERICAL INVESTIGATION OF THE INVERTIBILITY OF ROOM TRANSFER FUNCTIONS

William Putnam, Davide Rocchesso**, Julius Smith**

*CCRMA and Department of Electrical Engineering
Stanford University
Stanford, CA 94309

**Centro di Sonologia Computazionale
Universita′ degli Studi di Padova
35123 Padova -Italy

ABSTRACT

It has been stated that the inverse filtering of room transfer functions is possible using multiple inputs under certain conditions [1]. In this paper we address these conditions from a numerical perspective. A method is developed to determine the numerical precision of multiple input inverse filtering techniques. This method is applied to two simulated rooms, and the numerical precision is estimated as a function of the number of inputs and outputs into the system.

1. INTRODUCTION

Many authors have investigated the application of signal processing techniques to the problem of inverse filtering a room transfer function from a signal. This is done to decrease possible unwanted factors such as reverberation and spectral coloration, which are causes of degraded fidelity and intelligibility.

The inversion of a room transfer function is known to be a difficult task, because of the presence of non-minimum phase zeros, which can not be compensated by causal, stable filters [2]. If the input signal is delayed by a suitable amount, an approximate inversion can be obtained by least squares or homomorphic techniques [3].

The availability of multiple sound sources and/or multiple receivers can help in finding an exact solution to the inversion problem. Miyoshi and Kaneda proposed a method called MINT, which makes use of some fundamental results of multivariable system theory[1]. In practice, the finite precision of measured impulse responses along with the inversion of poorly conditioned matrices, may pose numerical limitations to this method. In this paper, we investigate the numerical behavior of some simulated room transfer functions, in the context of multiple input inversion methods.

2. BACKGROUND

Consider N loudspeakers (inputs) and $M<N$ microphones (outputs) in a room. We want to insert an FIR filter before each of the sources in such a way that a given source signal is transmitted to each of the outputs unaffected by the room response. The scheme is depicted in figure 1.

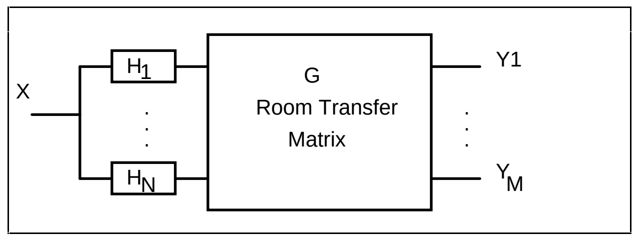

Figure 1: Diagram of a multiple input / output formulation of the inverse filtering problem. G is room transfer matrix whose elements, $G_{i,j}$ are the transfer polynomials from the jth input to the ith output.

$G_{i,j}(z^{-1})$ is the polynomial transfer function of the room from the input j to the output i. $H_j(z^{-1})$ is the FIR compensator to be applied to the j-th input. The problem thus becomes one of finding the solution to a system of polynomial equations:

$$D = \begin{bmatrix} G_{1,1} & \cdots & G_{1,N} \\ \vdots & & \vdots \\ G_{M,1} & \cdots & G_{M,N} \end{bmatrix} \begin{bmatrix} H_1 \\ \vdots \\ H_N \end{bmatrix} = GH \qquad (1)$$

where $D = [1 \quad 1 \quad \cdots \quad 1]^T$, for the problem of inverse filtering. By changing the vector D we might solve other problems such as transmitting a signal only to one output.

In order to make use of standard results for polynomial matrices, equation 1 is rewritten as:

$$\begin{aligned} D = &\begin{bmatrix} G_{1,1} & \cdots & G_{1,M} \\ \vdots & & \vdots \\ G_{M,1} & \cdots & G_{M,M} \end{bmatrix} \begin{bmatrix} H_1 \\ \vdots \\ H_M \end{bmatrix} + \\ &\begin{bmatrix} G_{1,M+1} & \cdots & G_{1,N} \\ \vdots & & \vdots \\ G_{M,M+1} & \cdots & G_{M,N} \end{bmatrix} \begin{bmatrix} H_{M+1} \\ \vdots \\ H_N \end{bmatrix} = G_1 H_1 + G_2 H_2 \end{aligned} \qquad (2)$$

A necessary and sufficient condition for the existence of a solution to equation 2 is that every left greatest common divisor of G_1 and G_2 is a left divisor of D[4]. Now, consider the Bezout identity for the polynomial matrices G_1 and G_2:

$$I = G_1 W_1 + G_2 W_2 \tag{3}$$

where I is the M x M identity matrix, W_1 is a M x M matrix, and W_2 is a $(N\text{-}M)$ x M matrix. If a solution to the Bezout identity exists, then we can form a solution to equation 2 by adding the columns of W_1 and W_2 to form the vectors H_1 and H_2 respectively. More generally, if a solution to equation 3 exists, we can find the solution to equation 2 for any D, by right-multiplication by W_1 and W_2.

A polynomial solution to the Bezout identity exists if and only if the two polynomial matrices G_1 and G_2 are mutually coprime, i.e., the matrix G never drops rank[4]. Namely, the Bezout identity admits a solution if and only if there exists two unimodular matrices U and V such that

$$G = USV \tag{4}$$

with $S = \begin{bmatrix} I_M & 0 \end{bmatrix}$. Equation 4 is called the Smith Canonical form of G.

When this condition holds, a block Toeplitz time domain matrix, formed of the impulse response measurements of each of the transfer functions of the system, can be inverted in order to solve for the appropriate filter coefficients, i.e. $\mathcal{H} = \mathcal{G}^{-1}\mathcal{D}$ [1]. In this equation, script variables are used to denote the time domain counterparts of the matrices in equation 1.

It has been stated that coprimeness is usually achieved "except for some symmetrical positions."[1] As far as we know, nobody has investigated the validity of this statement and its implications in the context of practical realizations. Strictly speaking, we need to determine if the matrix $\mathcal{G}$ is full rank. A more appropriate question regards the condition number of the matrix. Due to noise inherent in measurements, the matrix will likely be full rank, hence we are interested in cases where the matrix exhibits near singular behavior. As will be shown, this will impact the numerical precision of the filters designed with these methods.

This type of analysis involves the singular value decomposition of a matrix. Due to the high dimensionality of the time domain matrix, it is impractical to do this directly. In what follows, we develop the relation between the condition number of the time domain matrix and the transfer matrices evaluated over frequency. In this way, the analysis is reduced in dimension to a set of N x M matrix decompositions. Using these techniques, the singular behavior is studied as the number of inputs and outputs are varied.

3. NUMERICAL ANALYSIS

The condition number of a matrix is defined as the ratio of the maximum and minimum singular values of a matrix. In order to determine the condition number of the time domain matrix, we must bound the singular values. The maximum and minimum singular values are defined as:

$$\sigma_{\max}(\mathcal{G}) = \max_x \frac{|\mathcal{G}x|}{|x|} \qquad \sigma_{\min}(\mathcal{G}) = \min_x \frac{|\mathcal{G}x|}{|x|}, \tag{5}$$

where the norms used in the above equations are the Euclidean or L_2 norms. Another way to state this is that given any vector, we can form a bound on the maximum and minimum singular values in the following manner:

$$\sigma_{\max}(\mathcal{G}) \geq \frac{|\mathcal{G}x|}{|x|} \qquad \sigma_{\min}(\mathcal{G}) \leq \frac{|\mathcal{G}x|}{|x|} \tag{6}$$

Consider a finite length time domain vector v, comprised of a complex exponential with unitary Euclidean norm and frequency f_0. This can be used to construct a vector x_{f_0} in the following manner:

$$x_{f_0} = \begin{bmatrix} \alpha_1 v \,|\, \alpha_2 v \,|\cdots|\, \alpha_n v \end{bmatrix}^T \tag{7}$$

In the above expression, the α_i*'s* represent a complex weight applied to the vector v at each input of the system. We need to consider the Euclidean norm of the vector y_{f_0} given by.

$$y_{f_0} = \mathcal{G}x_{f_0} \tag{8}$$

We can transform this matrix equation to an equivalent equation involving the transfer matrix, through multiplication by the DFT matrix. This operation preserves the Euclidean norm since the DFT matrix is unitary. After doing this, we have:

$$\tilde{Y}_{f_0} = \tilde{G}\tilde{X}_{f_0} \tag{9}$$

In the above expression $\tilde{X}_{f_0}$ is of the form:

$$\tilde{X}_{f_0} = \begin{bmatrix} 0\cdots\alpha_1\cdots 0 | 0\cdots\alpha_2\cdots 0 | \cdots | 0\cdots\alpha_n\cdots 0 \end{bmatrix} \tag{10}$$

Since each section of the vector x_{f_0} consists only of a single frequency, all components of $\tilde{X}_{f_0}$ are zero, except for those elements corresponding to the frequency f_0. Due to the sparse nature of this vector, we simplify it to the equivalent, lower dimensional problem:

$$Y_{f_0} = G(f_0) X_{f_0} \tag{11}$$

where we consider only the nonzero elements of $\tilde{X}_{f_0}$ and $\tilde{Y}_{f_0}$, and the corresponding columns in the matrix $\tilde{G}$. This results in the following:

$$X_{f_0} = [\alpha_1 \alpha_2 \cdots \alpha_n] \quad Y_{f_0} = [\beta_1 \beta_2 \cdots \beta_n]$$

$$G(f_0) = \begin{bmatrix} G_{1,1}(f_0) & \cdots & G_{1,N}(f_0) \\ \vdots & & \vdots \\ G_{M,1}(f_0) & \cdots & G_{M,N}(f_0) \end{bmatrix} \tag{12}$$

Hence, G is simply the transfer matrix evaluated at the frequency f_0. Note that the vectors X_{f_0} and x_{f_0} have the same norm since they are related by the DFT which is unitary.

As stated previously, the condition number of the time domain matrix is given by:

$$\kappa(\mathcal{G}) = \frac{\sigma_{\max}(\mathcal{G})}{\sigma_{\min}(\mathcal{G})} = \frac{\max\limits_{x} \frac{|\mathcal{G}x|}{|x|}}{\min\limits_{x} \frac{|\mathcal{G}x|}{|x|}} \tag{13}$$

where $\sigma_{\max}(\mathcal{G})$, $\sigma_{\min}(\mathcal{G})$ are the maximum and minimum singular values. We can bound the numerator term in equation 13 as the maximum over both frequency and the α values in equation 7.

$$\max_{x} \frac{|\mathcal{G}x|}{|x|} \geq \sup_{f} \max_{\alpha} \frac{|\mathcal{G}x_f|}{|x_f|} = \sup_{f} \max_{X_f} \frac{|GX_f|}{|X_f|} \tag{14}$$

The equality in equation 14 makes use of the fact that X_{f_0} and x_{f_0} have the same lengths. A similar relation can be derived for the minimum singular value, which results in the following bound for the condition number:

$$\kappa(\mathcal{G}) \geq \frac{\sup\limits_{f} \max\limits_{X_f} \frac{|GX_f|}{|X_f|}}{\inf\limits_{f} \min\limits_{X_f} \frac{|GX_f|}{|X_f|}} = \frac{\sup\limits_{f} \sigma_{\max}(G(f))}{\inf\limits_{f} \sigma_{\min}(G(f))} \tag{15}$$

κ can now be used to make a statement concerning the numerical precision of the inverted matrix. In practice, the time domain matrix is constructed from impulse response measurements determined through experimental means, and hence exhibit finite numerical precision. We can express the inherent numerical error in the following manner.

$$\hat{\mathcal{G}} = \mathcal{G} + \Delta\mathcal{G} \tag{16}$$

where $\hat{\mathcal{G}}$ is the matrix of measured impulse responses, $\mathcal{G}$ is the actual value, and $\Delta\mathcal{G}$ is the error. As stated in an earlier section, the filters which are applied to the inputs of the multiple input system are given by:

$$\mathcal{H} = \hat{\mathcal{G}}^{-1}\mathcal{D} \tag{17}$$

We can express the relative error in $\mathcal{H}$ as[5]:

$$\frac{\|\Delta\mathcal{H}\|}{\|\mathcal{H}\|} \leq \kappa(\mathcal{G}) \frac{\|\Delta\mathcal{G}\|}{\|\mathcal{G}\|}, \tag{18}$$

where $\|\mathcal{G}\|$ denotes the spectral norm of the matrix, which is defined as its maximum singular value. This can be interpreted directly in terms of bits of precision of the resultant filters,

$$\#bits(\mathcal{H}) = \#bits(\mathcal{G}) - \log_2(\kappa(\mathcal{G})) \tag{19}$$

ROOM:	1			2		
	x	**y**	**z**	**x**	**y**	**z**
dimension	6	7.1	8.6	8	10	6
Source1	7.0	.50	1.4	.93	.50	.97
Source2	7.0	.50	7.9	.93	.50	2.7
Source3	5.2	.50	1.4	6.9	.50	.97
Source4	5.2	.50	7.9	6.9	.50	2.7
Source5	.35	.50	4.3	.46	.50	3.0
Source6	3.0	.50	8.3	4.0	.50	5.7
Source7	3.0	.50	7.0	4.0	.50	.48
Source8	5.6	.50	4.3	7.5	.50	.19
Rcvr1	1.9	5.9	1.0	2.6	8.2	1.0
Rcvr2	4.0	5.9	1.0	5.4	8.2	1.0
Rcvr3	1.9	6.5	2.0	2.6	9.1	2.0
Rcvr4	4.0	6.5	2.0	5.4	9.1	2.0

Table 1: Room geometry and source / receiver locations for the simulations (in meters).

4. SIMULATION

A. Description

In order to study the singular behavior of the transfer matrix, simulations were performed using two rooms with rectangular geometry. Impulse responses were generated using the image method [6]. For the purpose of this experiment, a sampling rate of 1KHz was used, hence only low frequency behavior was studied. Each impulse response was calculated for a duration of 1 second.

Eight Sources and four receivers were used, and impulse responses were generated for each of the 32 input output combinations. The room dimensions, and source / receiver locations are shown in table 1. The first room used the ratios 1:1.186:1.439 to determine the length width and height. These have been shown to be optimal in the sense that they minimize the deviation of the modal distribution from the asymptotic hyperbolic curve[7]. The second room used the ratios of 3:4:5. Both rooms used reflection coefficients of .7 on all boundary surfaces. Sources and receivers were assumed to be omnidirectional in nature.

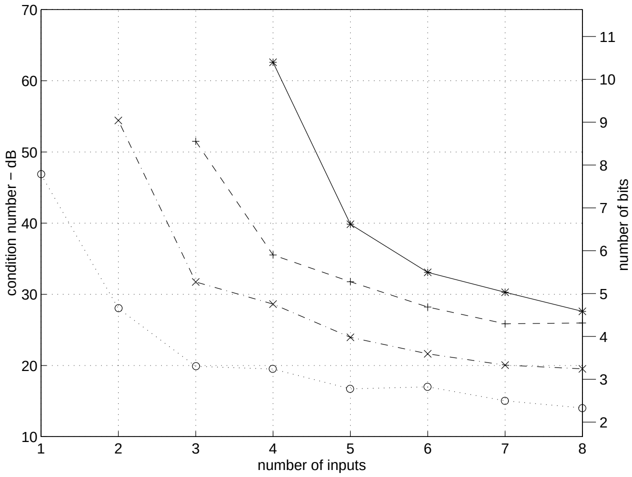

Figure 2: Plot of condition number versus number of inputs for room 1. The curves are for the cases of inverse filtering at 1 ('o') through 4 (' * ') outputs.

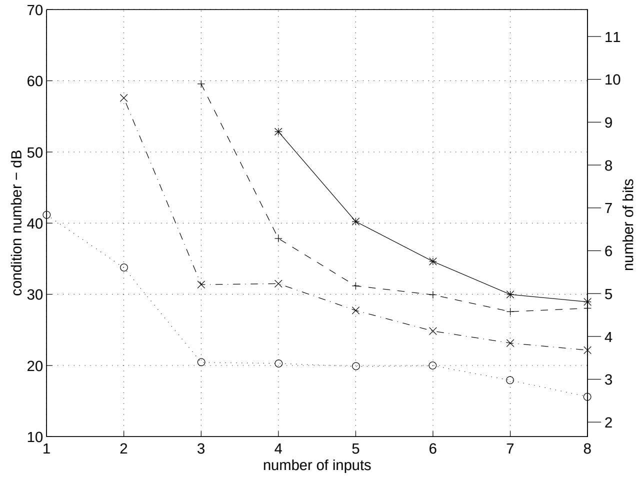

Figure 3: Plot of condition number versus number of inputs for room 2. The curves are for the cases of inverse filtering at 1 ('o') through 4 ('*') outputs.

B. Results

Subsets of the 32 impulse responses were studied for each room in order to examine the behavior of the system as inputs or outputs were added. For each of these subsets of data, a transfer matrix was constructed, and its singular values were calculated over a range of frequencies. Equation 15 could then be used to relate this data to the condition number of the corresponding time domain matrix.

Figures 2 and 3 show the results for each of 2 rooms. These plots should be interpreted as the loss in numerical precision of the filter coefficients, resulting from the inversion of the time domain matrix. This number can be expressed both in dB as well as bits of precision. For example, in room 2, we notice a loss in precision of approximately 18 dB or 3 bits when 4 inputs are used with 1 output. Hence, the resulting precision of the compensation filters would be 3 bits less than that of the measured impulse responses.

Not surprisingly, the condition number decreases, and hence numerical performance is enhanced as the number of inputs is increased. Conversely, the condition number increases at the rate of approximately 1 bit for each output added. Both rooms exhibited similar in this respect.

5. CONCLUSION

Two rooms were studied in the context of inverse filtering using multiple input techniques. The analysis provides an estimate of the loss in numerical precision of the compensation filters, which occurs due to the inversion of a matrix. It was shown that the condition number of the time domain matrix is related to the singular values of the transfer matrix evaluated over frequency. This provides an efficient way to study the singularities of the larger time domain matrix. Furthermore, this study provides insight into the level of benefit the inverse filtering operation receives when the number of inputs is increased.

REFERENCES

1. Miyoshi, M., and Kaneda, Y. "Inverse Filtering of Room Acoustics," *IEEE Trans. ASSP* 36, pp. 145-152, 1988.
2. Neely, S.T., and Allen, J.B. "Invertibility of a Room Impulse Response," *J. Acoust. Soc. Am.*, Vol. 66, July 1979.
3. Mourjoupolous, J., Clarkson, P.M., Hammond, J.K., "A Comparative Study of Least-Squares and Homomorphic Techniques for the Inversion of Mixed Phase Signals," *Proc. IEEE Int. Conf. on ASSP* (1982), pp.1858-1861.
4. Kailath, T., *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ, 1980.
5. Golub, T., Van Loan, C. , *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1989.
6. Allen, J.B., Berkeley, "Image Method for Efficiently Simulating Small-Room Acoustics," *J. Acoust. Soc. Am.*, Vol. 65, pp. 943-950, 1979.
7. Milner, J., Bernhard, R., "An Investigation of the Modal Characteristics of Nonrectangular Rooms," *J. Acoust. Soc. Am.*, Vol. 68, No. 2, February 1989.