A COMPARISON OF THE PERFORMANCE OF HRTF MODELS IN INVERSE FILTER DESIGN FOR CROSSTALK CANCELLATION

J. Hollebon ISVR, University of Southampton, UK
E. C. Hamdan ISVR, University of Southampton, UK
Dr. F. M. Fazi ISVR, University of Southampton, UK

1 INTRODUCTION

Binaural audio signals contain localization cues, such as the Interaural Time Difference (ITD) and Interaural Level Difference (ILD) that allows for the placement of a virtual sound source anywhere around a listener[1]. This results in specific signals for the listener's left and right ears that must be reproduced exactly to maintain the binaural effect. Often, this is achieved by reproducing binaural audio over headphones. However, in many scenarios it is not desirable to wear headphones. Furthermore, with headphones come negative effects such as in-head localisation[2].

Alternatively, binaural audio may be reproduced over loudspeakers through the use of a Crosstalk Cancellation (CTC) system. CTC systems employ soundfield control by the means of inverse filtering, such that at the listener's left and right ears the left and right binaural signals are reproduced correctly. The geometry of a CTC system has important implications on its performance. Hence, whilst research initially focused on two-channel stereo CTC systems[3], this has developed to specific two-channel loudspeaker geometries such as the Stereo Dipole[4] and the Optimal Source Distribution[5]. However, the CTC formulation allows for any number of loudspeakers or listeners[6]. Hence more recent research has focused on using loudspeaker arrays for CTC[7], including for multiple listener reproduction[8,9]. Finally, CTC creates a specific sweet-spot in which the listener must be positioned. However, depending on the method chosen for the filter creation the system may be made adaptive to the listener position, if listener tracking is employed[10,11].

In designing the CTC filters, a plant matrix of acoustic transfer functions between each of the loudspeakers and the listener's ears is required. This may be an accurate, measured plant matrix of the real system including the loudspeaker's response and the listener's Head-Related Transfer Function (HRTF), however this may be impractical to measure. Instead, it is often the case that a model of the system is used, where assumptions are made such as a specific listener HRTF, as well as about how the loudspeakers radiate. However, this introduces a mismatch in the final CTC system, where the plant matrix used to create the CTC filters may not accurately represent the real system. In this case, errors will be introduced in the final reproduced signals at the listener's ears resulting in a degradation of the binaural effect[12]. Often, the loudspeakers are modelled as monopole sources, however there is a range of options for the choice of the HRTF model. Such models have been well compared in the literature[13], and even morphed HRTF models for use in CTC have been suggested[14].

This paper presents a study of the performance of common plant models used in CTC for the design of the inverse filters. These include the freefield shadowless head model for monopole sources and the rigid sphere for monopole sources. The models' performances are assessed in the context of a real two-channel stereo CTC system, where a plant measurement of the real system is used as a reference solution. Next, the models are compared under the presence of perturbations to the system due to small translations and rotations of the listener's head, corresponding to inexact positioning of the listener in the CTC system's sweet-spot. Subsequently, regularisation is included in the CTC filter design and the models' performances are compared with both regularisation and perturbations present. Finally, a suggestion on the most appropriate choice of plant model for creating CTC filters is presented.

2 THEORY

2.1 Crosstalk Cancellation

A two-channel, single listener CTC system is shown in Figure 1, such that there are $L = 2$ loudspeakers controlling the pressure at $M = 2$ ears, or control points. The system may be described in the frequency domain by

$$\mathbf{p} = \mathbf{Cq} = \mathbf{CHd} \tag{1}$$

where $\mathbf{p}$ is the length M vector of reproduced pressures at the ear control points, $\mathbf{C}$ is the $M \times L$ plant matrix of acoustic transfer functions, $\mathbf{q}$ is the length L vector of loudspeaker signals, $\mathbf{H}$ is the $L \times M$ matrix of CTC filters and $\mathbf{d}$ is the length M vector of binaural signals to be reproduced. The goal of the CTC system is through the use of a set of CTC filters, $\mathbf{H}$, the reproduced pressures should equal the binaural signals i.e. $\mathbf{p} = \mathbf{d}$. For this to occur, the CTC filters must be the inverse of the plant matrix such that

$$\mathbf{H} = \begin{cases} \mathbf{C}^{-1} & when\ M = L \\ \mathbf{C}^{+} & when\ M \neq L \end{cases} \tag{2}$$

where the operator $(\cdot)^{+}$ denotes the Moore-Penrose pseudoinverse. A modelling delay is also required to ensure causality[15] but it is omitted here for conciseness. Often, it is the situation that the plant matrix undergoing inversion is ill-conditioned which may lead to very large loudspeaker gains, as well as solutions sensitive to errors in the plant matrix[5]. Hence, Tikhonov regularisation is often employed to improve the conditioning of the matrix at frequencies where the system is ill-conditioned. In this case, and when $M > L$, the CTC filters are given by[15]

$$\mathbf{H} = \mathbf{C}^{\mathrm{H}}[C\,C^{H} + \beta\,I_m]^{-1} \tag{3}$$

where β is the regularisation parameter, I_m is the $M \times M$ identity matrix and the operator $(\cdot)^{\mathrm{H}}$ indicates the Hermitian transpose. Whilst regularisation may be used to limit the loudspeaker gains and improve the robustness of the solution, it comes at the cost of introducing errors in to the reproduced pressures at the listener's ears[16].

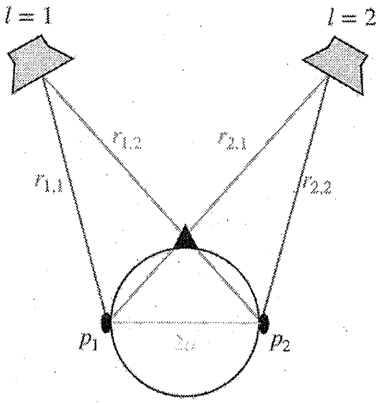


Figure 1: Reference geometry of a two-channel, single listener, CTC system.

2.2 Head-Related Transfer Function Models For The Plant Matrix

To ensure the CTC filters accurately reproduce the correct pressures at the control points, the plant matrix used in creating the CTC filters must be as similar to that of the real CTC system as possible. It is possible to measure the transfer functions between each loudspeaker of the system and the listener's head using either in-ear binaural microphones to use the listener's own Head-Related Transfer Function (HRTF), or a binaural mannequin microphone. However, these measurements are often impractical. Furthermore, using measured plant matrices means the CTC system may not easily be made adaptive to listener movements.

The most commonly used plant model of a CTC system is the freefield shadowless head model for monopole sources, herein referred to as the freefield monopole model. In this case, the loudspeakers are modelled as acoustic monopoles. The HRTF used is two diametrically opposed points in free-space, separated by a head diameter of $2a$, i.e. there is no model of the diffraction effects the listener's head contributes to the problem. Hence, this model provides an estimate of the ITD of the HRTF, but not the ILD due to shadowing of the head – however there is a small ILD due to distance attenuation of the monopole source. Let $r_{m,l}$ be the distance from loudspeaker l to control point m. Then the pressure due to the loudspeaker at the control point may be written as a Greens function, normalized by the pressure at the centre of the head, such that the entries for the plant matrix are given by[7]

$$c_{m,l} = \frac{e^{-jkr_{m,l}}}{4\,\pi\, r_{m,l}} \cdot \frac{4\,\pi\, r_{c,l}}{e^{-jkr_{c,l}}} \qquad (4)$$

where $c_{m,l}$ is entry of index m, l of the plant matrix, k is the wavenumber and $r_{c,l}$ is the distance from loudspeaker l to the centre of the head. This model is extremely simple and quick to calculate - an advantage for real-time adaptive CTC systems. Furthermore, the CTC filters may be calculated adaptively by measuring $r_{m,l}$ at any given time and updating the filters accordingly.

A second model, classically proposed by Lord Rayleigh but less used in CTC, is the rigid sphere model due to monopole sources. Here the loudspeakers are again modelled as acoustic monopoles, but now the head is modelled as a rigid sphere of radius a, where the ears are diametrically opposed. Hence the pressure field now includes the acoustic scattering due to the rigid sphere. This model therefore includes an estimate of the ITD due to diffraction of the incident wave around the sphere, as well as an ILD estimate due to the shadowing of the rigid sphere. Therefore, the pressure at the left ($m = 1$) and right ($m = 2$) ear control points due to loudspeaker l is given by[17]

$$\begin{aligned} c_{1,l} &= \frac{r_{c,l}}{ka^2 e^{-jkr_{c,l}}} \sum_{n=0}^{\infty} \frac{(2n+1)\, h_n\left(kr_{c,l}\right) (-1)^n\, P_n(\sin\theta_c)}{h_n'(ka)} \\ c_{2,l} &= \frac{r_{c,l}}{ka^2 e^{j-kr_{c,l}}} \sum_{n=0}^{\infty} \frac{(2n+1)\, h_n\left(kr_{c,l}\right) P_n(\sin\theta_c)}{h_n'(ka)} \end{aligned} \qquad (5)$$

where h_n is the n^{th} order spherical Hankel function of the second kind, P_n is the n^{th} order Legendre polynomial, θ_c is the angle the loudspeaker is positioned at measured from the median plane and the operator $(\cdot)'$ denotes differentiation. Eqn. (5) is a series that for any calculation must be truncated to a given N summations. The rigid sphere model is more computationally expensive than the freefield monopole model, however it is a more accurate representation of the human head than the freefield monopole model and may still be made adaptive to listener position. Notably, the rigid sphere does not model the effect of the pinna.

Both models have dependence on the size of the head, given by its radius a. For a fair comparison to a real HRTF, the head radius must be tuned such that the models give as similar results as possible to the real HRTF, noting that a real head is not symmetric. Furthermore, it has been shown that at low frequencies the rigid sphere HRTF is equivalent to using the freefield shadowless head model except with an enlarged head radius[13], such that $a' = 3a/2$ is used in this region. Hence, using an enlarged head radius at low frequencies will give better results for the shadowless head model, however will skew its results at high frequencies. Despite this, here for consistency one constant value for the head radius was used for all the models at all frequencies.

2.3 Performance Metrics

Two key performance metrics for the analysis of CTC systems are the array effort, AE, and the crosstalk cancellation spectrum, CTC. The array effort is the norm of the loudspeaker signals divided by the norm of the input signal, q_s, where q_s is the loudspeaker signal required by a single loudspeaker to reproduce the same pressure as the CTC system in a given ear. Hence,

$$AE = \frac{q^H q}{|q_s|^2} \tag{6}$$

where the array effort is a dimensionless quantity that is typically presented in decibels. The array effort gives a measure of the gains of the loudspeaker filters, where large loudspeaker gains are related to ill-conditioning in the plant matrix[5]. Hence, the array effort is a useful metric for deciding the value of the regularisation parameter, β.

The CTC spectrum for a single listener is defined as the ratio of the squared reproduced pressures at the listener's ears for a target binaural signal $\boldsymbol{d} = [1,0]^T$, which equates to reproducing an impulse in the listener's left ear and a null in the listener's right ear. The CTC spectrum is therefore a measure of the amount of CTC achieved by the system between the two ears of the listener. Therefore, the CTC spectrum is given by.

$$CTC = 10\log_{10}\left(\frac{|p_1|^2}{|p_2|^2}\right). \tag{7}$$

In general, a minimum of approximately 20 dB of CTC is required between the two ears to maintain the binaural effect[18].

3 HRTF COMPARISON

To perform a comparison between the different models for the CTC filter creation, the transfer functions of a two-channel stereo CTC system were measured, as shown in Figure 2. The system consisted of two Genelec 8020D loudspeakers, at a distance of 1.25 m from the center of the head and with a span of 60 degrees. To measure the transfer functions, sine sweeps were played by a control computer through an RME Fireface audio interface, which drove each of the loudspeakers. A Neumann KU100 binaural microphone was used to record the sweep, which was deconvolved to give the transfer function of each loudspeaker to each control point[20]. To ensure freefield conditions, the measurements took place in the anechoic chamber at the Institute of Sound and Vibration Research (ISVR), University of Southampton.

The measured transfer functions were used to populate a reference plant matrix, C_{ref}. Next, three sets of CTC filters were made using the freefield shadowless head model, the rigid sphere model, and the measured 'in situ' transfer functions, respectively. No regularisation was used in creating the CTC filters. A head radius of $a = 0.0875$ m was used for the freefield shadowless head model and the rigid sphere, as this gave an optimal fit to the KU100 transfer functions.

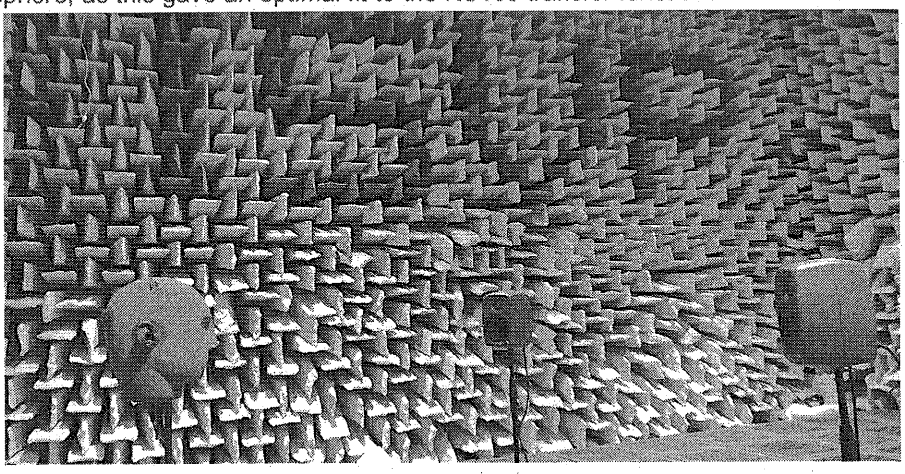
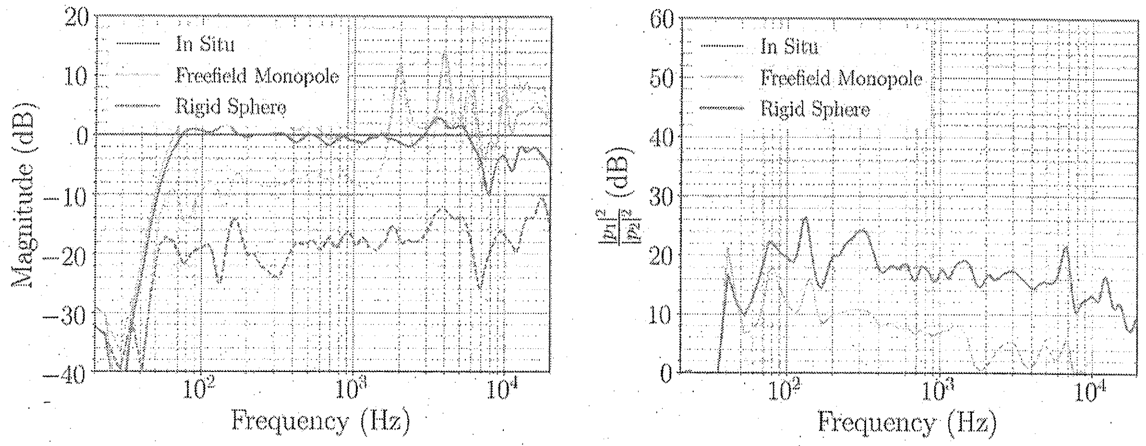


Figure 2: Measurement of the two-channel CTC system in the anechoic chamber at the ISVR, University of Southampton



(a) Reproduced control point pressures. *(b) CTC spectrum.*

Figure 3: Reproduced pressures and CTC spectrum for the three filter sets. For the reproduced pressures, the solid and dashed lines indicate the left and right ears respectively.

Next, using offline numerical simulations, Eqn. (1) was evaluated using the reference measured plant matrix however each time changing the CTC filters to one of the three filter sets in turn. Hence the reproduced pressure due to filter set i is given by

$$\boldsymbol{p}^i = \boldsymbol{C}_{ref}\boldsymbol{H}^i\boldsymbol{d} \tag{8}$$

where a target binaural signal of $\boldsymbol{d} = [1,0]^T$ was used. This equates to using the real CTC system, except with different design processes for each of the filter sets. Furthermore, this is reliant on the assumption that the system is both linear and time-invariant. The results of the reproduced pressures and CTC spectrums are shown in Figure 3. As there is no regularisation, the in situ filter set corresponds to a perfect inversion and hence the left and right ear achieve the desired target signals of $\boldsymbol{d} = [1,0]^T$, or 0 dB and $-\infty$ dB, respectively. The low frequency roll-off seen in all other filter sets below ~80 Hz corresponds to the roll-off of the loudspeakers below their resonance frequency.

The freefield monopole model achieves a small amount of low frequency CTC, averaging 10 dB between ~100 Hz and ~1000 Hz. However, at high frequencies the CTC completely breaks down, with large peaks in the pressure spectrum equating to strong colouration in the reproduced signals as well as little to no CTC in this region. This is a phenomenon well studied in the literature, and equates to frequencies at which the plant matrix is ill-conditioned[5]. This occurs due to difficulty in recreating the in-phase or out-of-phase mode of the binaural signals at the listener's ears at these frequencies, dictated by the geometry of the CTC system[5,19]. The rigid sphere model performs considerably better, and across the whole spectrum achieves close to the required 20 dB of CTC. In the spectrum of the magnitude of the reproduced pressures, it is clear that the HRTF of the KU100 in the reference plant matrix causes some colouration, particularly at high frequencies with a pinna notch at 7000 Hz.

Whilst the in situ filters clearly perform the best, providing perfect inversion, the rigid sphere outperforms the freefield monopole model. This is likely because the rigid sphere is a more accurate HRTF in the low and mid frequencies compared to the freefield monopole model. This is partly due to the inclusion of an ILD estimate in the HRTF of the rigid sphere, where in the mid-high frequencies the ILD cue is strong and featured in an accurate HRTF. However, at high frequencies the rigid sphere is less accurate, particularly due to its lack of modelling the listener's ears, and hence performs worse in these regions. Furthermore, whilst the in situ filters perform the strongest, it is important to remember that often measured CTC filters use the actual loudspeakers of the system but the HRTF of a binaural mannequin microphone, not that of the listener. Hence for a real listener there would be a mismatch between the HRTF used for the filter creation and that in plant matrix of the actual CTC system, which would result in a drop in the amount of CTC achieved.

4 PERTURBATION STUDY

4.1 Mismatches In Listener Position

So far, the CTC system has been tested assuming perfect alignment of the loudspeakers and listener. However, with a real system it is likely that there will be misalignments due to setting up the system in a different space, or movement of the listener, such that the modelled plant matrix assumes the wrong listener and loudspeaker positioning. To account for this, perturbations where applied to the listener's position in order to investigate the achievable CTC when the listener is not in the exact sweet spot of the CTC system. Hence, as before the transfer functions of the CTC system were measured except now the KU100 microphone was displaced. The perturbations were random and included both translations and rotations of the microphone. The perturbations were limited to a maximum radial displacement of 5 cm and a maximum rotation of ± 15 degrees. These values were chosen to represent realistic small perturbations of the listener about the listening sweet spot. Transfer functions were then measured for a total of 30 perturbations.

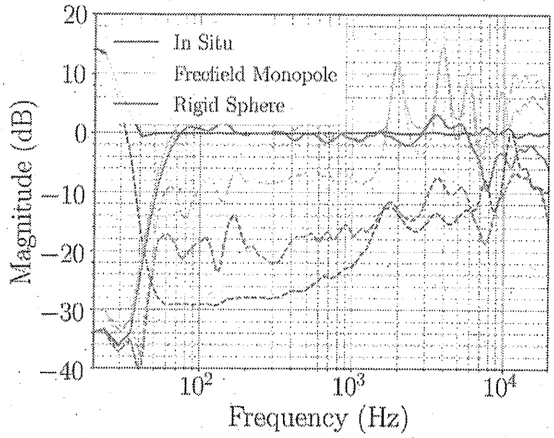
Next, the forward problem was simulated as before for the three different CTC filter sets that assumed the listener was positioned in the sweet spot. However, the problem was run 30 times for each filter set where the reference plant matrix used was changed to each of the measured perturbed plant matrices in turn. The reproduced pressures were then averaged over the 30 perturbed forward problems to allow for comparison between the filter sets in a similar manner as before.

The reproduced control point pressures and CTC spectrum for the perturbed problem are shown in Figure 4 (a) and (b). Compared to the scenario where there is no perturbation, there is a reduction in the amount of CTC achieved using the in situ filter set, as this no longer corresponds to a perfect inversion due to the shift in the position of the listener's head. However, in the low to mid frequencies the in situ filter set still performs the strongest. There is a boost in the reproduced pressure at very low frequencies for the in situ filters – as the measured transfer functions are mostly noise in this region the CTC filters and resulting pressures are henceforth meaningless here. In practise, a high-pass filter would often be employed to account for this. At high frequencies there is a drastic loss of CTC for all the filter sets. This is expected as perturbations will have a greater effect when the wavelength is comparable to the size of the perturbation. Here, the rigid sphere and in situ filters perform similarly.

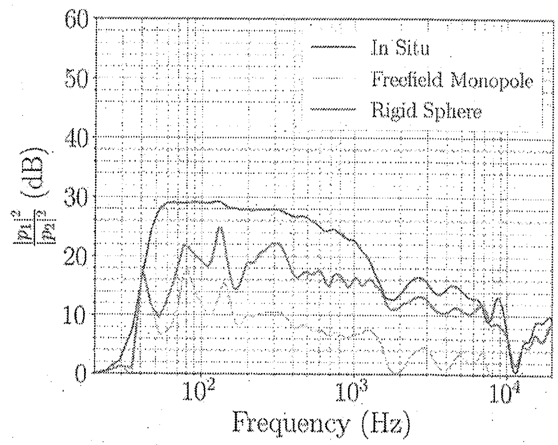
Interestingly, the freefield monopole and rigid sphere filters achieve a very similar amount of CTC to the unperturbed problem in Figure 3. This is likely because the effect of mismatches in the model of the CTC system, including the loudspeaker's response and the listener's HRTF, are of a much larger magnitude than the physical perturbations enforced in the measurements. Therefore, these dominate the errors and the perturbed results appear similar to those that are unperturbed. Henceforth, as before the rigid sphere achieves a larger channel separation across the whole spectrum than the freefield monopole model. The freefield monopole model also exhibits the same characteristic peaks in the high frequencies which result in a lack of CTC and colouration in the reproduced binaural signals at these frequencies.

4.2 Regularisation

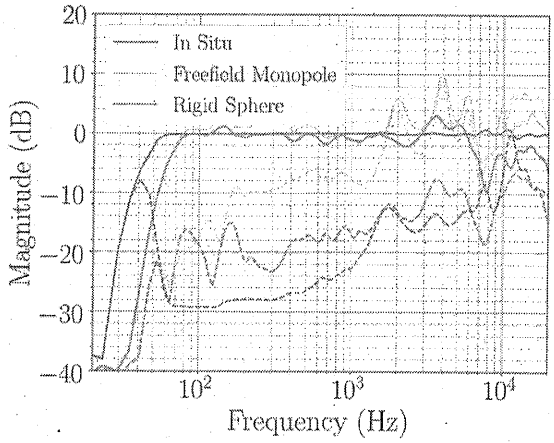
To improve the conditioning of the plant matrix under inversion in the creation of the CTC filters, as well as increase robustness to perturbations, Tikhonov regularisation was used in the filter creation as per Eqn. (3). Hence, the perturbed problem was repeated for all 30 of the perturbed plant matrices, except the CTC filters now included a set amount of regularisation. The AE was used to decide on the amount of regularisation to ensure a consistent application across all the different filter sets. Hence, the regularisation parameter, β, was chosen for three different scenarios such that the AE did not exceed $+\infty$ (no regularisation), 10 dB and 5 dB respectively. The values of β for all the filter sets and AE limits are shown in Table 1, whilst the AE of the filter sets having applied regularisation is shown in Figure 5.



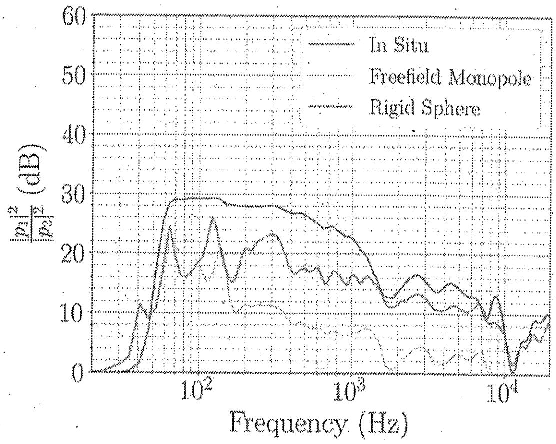
(a) Reproduced pressures with no regularisation.



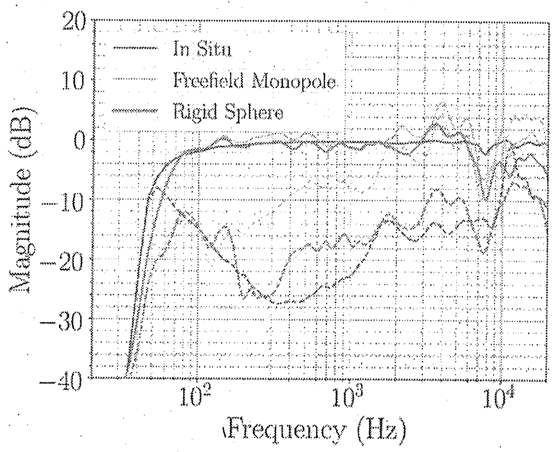
(b) CTC spectrum with no regularisation.



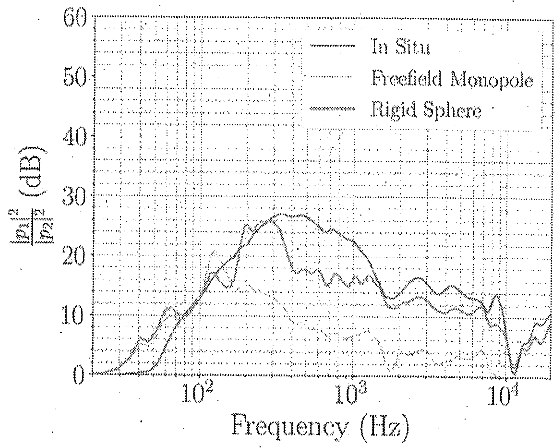
(c) Reproduced pressures with regularisation such that the AE does not exceed 10 dB.



(d) CTC spectrum with regularisation such that the AE does not exceed 10 dB.

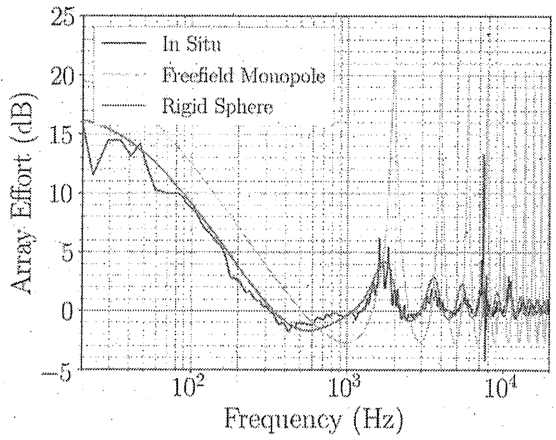


(e) Reproduced pressures with regularisation such that the AE does not exceed 5 dB.

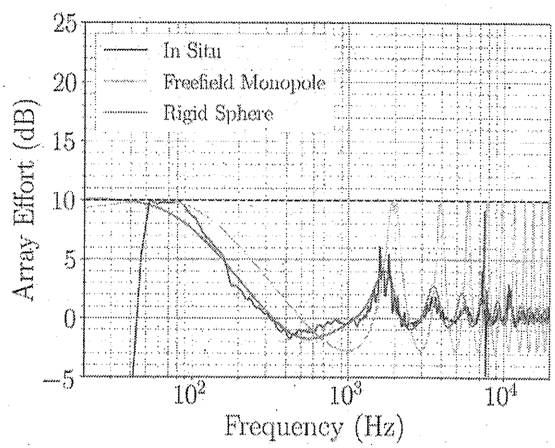


(f) CTC spectrum pressures with regularisation such that the AE does not exceed 5 dB.

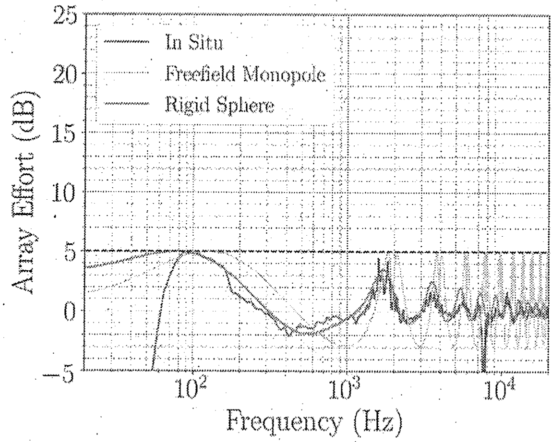
Figure 4: Average control point responses for the perturbed forward problem using three CTC filter sets, with varying levels of regularisation. For the reproduced pressures the solid and dashed lines indicate the left and right ears respectively.



(a) AE with no regularisation.



(b) AE with regularisation for a 10 dB limit.



(c) AE with regularisation for a 5 dB limit.

Figure 5: AE for the three filter sets with varying amounts of regularisation to enforce different limits in the AE.

The AE gives an indication of the conditioning of the plant matrix under inversion, hence large values in the AE indicate large loudspeaker gains and ill-conditioning at these frequencies. From Figure 5 (a) it is clear that at very low frequencies all the filter sets are ill-conditioned. In general, the AE for the rigid sphere is very similar to the in situ filters, except for a noticeable peak near 8000 Hz which is due to the effect of the pinna included only in the in situ filters. However, the freefield monopole model includes well-defined peaks throughout the spectrum where ill-conditioning occurs. These are well documented in CTC and correspond to frequencies dictated by the geometrical set-up of a two-channel CTC system, at which the system is unable to provide CTC[5,19]. It is the presence of these peaks that have given rise to more complicated loudspeaker geometries for CTC, such as the Optimal Source Distribution and loudspeaker arrays.

From the AE plots, it is clear that applying regularisation to enforce a set AE limit has a significant effect only at frequencies where the magnitude of the AE exceeds the limit. All of the three filter sets are affected by regularisation at low frequencies. However, the rigid sphere and in situ filters are generally well conditioned elsewhere. Hence even when applying heavy regularisation, i.e. setting the 5 dB AE limit, the regularisation has negligible affect at any other frequencies, the one exception

Filter Model	AE max +∞	AE max 10 dB	AE max 5 dB
In Situ	$\beta = 0.000000$	$\beta = 0.00095$	$\beta = 0.04750$
Freefield Monopole	$\beta = 0.000000$	$\beta = 0.01350$	$\beta = 0.04410$
Rigid Sphere	$\beta = 0.000000$	$\beta = 0.01350$	$\beta = 0.04420$

Table 1: Regularisation parameters for each of the filter sets for the three AE limits.

being the high frequency peak due to the pinna in the in situ filters. However, with the freefield monopole the regular high frequency peaks are strongly regularised as they are strongly ill-conditioned. Hence, the freefield monopole model is both a less accurate HRTF model and more strongly ill-conditioned than the rigid sphere and in situ measurements. Therefore, with the freefield monopole filters regularisation must be applied to account for an issue that does not occur to the same degree as in the real physical system.

The reproduced control point pressures and CTC spectrum for the perturbed problem, including regularisation, are shown in Figure 4. As seen from examining the AE regularisation only affects frequencies where the AE magnitude exceeds the set limit. When this is this case, peaks in the magnitude of the frequency response at each control point are flattened reducing the colouration in the binaural signals, however little to no CTC is achievable at these frequencies. As the ill-conditioning of all the models is similar at low frequencies, as seen in Figure 5, all the models see a low frequency limit below which CTC is not achieved. Increasing the amount of regularisation results in raising the frequency of this limit. However, even with heavy regularisation the rigid sphere and in situ filters see little or no change in performance above the low frequency limit. It is the freefield monopole filters that benefit the most from regularisation, which results in a trade-off in an even lower achievable amount of CTC for a less coloured reproduced binaural signal. Despite this, it remains the worst performing model.

5 CONCLUSIONS

In conclusion, the performance of three well-established plant models for CTC filter creation have been compared in the context of a real two-channel stereo CTC system. The three plant models investigated were measured transfer functions of the real CTC system ('in situ' filters), the freefield shadowless head model for acoustic monopole sources and the rigid sphere for acoustic monopole sources. The measurements of the CTC system were used as a reference plant matrix. Both theoretical models assume that the loudspeakers radiate as monopoles. The freefield shadowless head model utilises a HRTF of two diametrically opposed points in free space, such that an ITD is estimated and shadowing due to the head is not modelled. The rigid sphere assumes the HRTF is that of a perfectly rigid sphere, hence includes an estimate for both the ITD and ILD of the HRTF.

Offline numerical simulations of the CTC system using each of the filter sets in turn were employed to assess the performance of each filter set. Assuming the system is perfectly aligned and with no regularisation applied, the in situ filters provide perfect CTC. However, it is important to remember here the in situ filters used the HRTF of a binaural mannequin microphone – a HRTF mismatch would be introduced if a real listener replaced the binaural microphone. The freefield monopole filters performed poorly, achieving only 10 dB of CTC in the low to mid frequencies whilst at high frequencies exhibiting strong colouration in the reproduced pressures at the listener's ears as well as little to no CTC in this region. The rigid sphere performs well across the whole spectrum, in general reaching 20 dB of CTC. However, at high frequencies there is a reduction in the CTC achieved due to the absence of any model of the pinna.

Next, the filters were assessed in the presence of perturbations to the listener position including translations and head rotations, as well as the addition of regularisation in the filter creation. Including perturbations, the in situ filters no longer provide perfect CTC, however still outperform the two theoretical models at all frequencies. The freefield monopole and rigid sphere filters are less affected by the presence of perturbations, as likely the differences between the models and the real plant of the CTC system correspond to larger errors than the perturbations enforced in the measurements.

Finally, the use of regularisation reduces all the filter set's ability to provide CTC at low frequencies. However, it is shown that the in situ and rigid sphere plant matrices are generally better conditioned elsewhere, and hence require no regularisation at all outside of the low frequency region. The one exception is at high frequencies around notches due to the pinna in the in situ transfer functions. The freefield monopole model requires heavy regularisation at mid to high frequencies, however this exaggerated ill-conditioning is due to the lack of modelling the acoustic effects of the listener's head which is not a physical reality, and hence not seen in the in situ or rigid sphere plants. Therefore, the commonly used freefield monopole model requires the use of regularisation to account for a problem that does not exist to the same degree in a real CTC system.

Henceforth, the rigid sphere is shown to strongly outperform the freefield shadowless head model, which is a widely used plant model for CTC filter creation. However, in situ measurements still provide a substantial improvement over the rigid sphere plant. Despite this, it may still be advantageous to use the rigid sphere model. This is because the rigid sphere model both may be made adaptive to listener movements and requires no experimental measurements of the CTC system. Furthermore, at high frequencies it may be more beneficial to not use a real HRTF with pinna notches, as inverting these notches introduces ill-conditioning as seen with the in situ measurements.

6 REFERENCES

1. Blauert, J. (1997). *Spatial hearing.* Cambridge: MIT Press.
2. Xie, B. (2013). *Head-related transfer function and virtual auditory display.* J. Ross Publishing.
3. Schroeder, M. and Atal, B. (1963). Computer simulation of sound transmission in rooms. *Proceedings of the IEEE*, 51(3), pp.536-537.
4. Kirkeby, O., Nelson, P. and Hamada, H. (1998). Local sound field reproduction using two closely spaced loudspeakers. *The Journal of the Acoustical Society of America*, 104(4), pp.1973-1981.
5. Takeuchi, T. and Nelson, P. (2002). Optimal source distribution for binaural synthesis over loudspeakers. *The Journal of the Acoustical Society of America*, 112(6), pp.2786-2797.
6. B. Bauck, J. and Cooper, D. H. (1996). Generalized Transaural Stereo and Applications. *Journal of the Audio Engineering Society,* 44(9), pp.683–705.
7. Simón Gálvez, M. and Fazi, F. M. (2015). Loudspeaker Arrays For Transaural Reproduction. *The 22nd International Congress of Sound and Vibration, Florence.*
8. Simón Gálvez, M. and Fazi, F. M. (2017). A Loudspeaker Array for 2 People Transaural Reproduction. *24th International Congress On Sound and Vibrations, London.*
9. Hollebon, J,, Simón Gálvez, M. and Fazi, F. M. (2019). Multiple Listener Crosstalk Cancellation Using Linear Loudspeaker Arrays For Binaural Cinematic Audio. *Audio Engineering Society Conference: AES International Conference on Immersive and Interactive Audio.*
10. Simón Gálvez, M., Menzies, D. and Fazi, F. M. (2019). Dynamic Audio Reproduction with Linear Loudspeaker Arrays. *Journal of the Audio Engineering Society*, 67(4), pp.190-200.
11. Simón Gálvez, M., Takeuchi, T. and Fazi, F. M. (2017). Low-Complexity, Listener's Position-Adaptive Binaural Reproduction Over a Loudspeaker Array. *Acta Acustica united with Acustica*, 103(5), pp.847-857.
12. Akeroyd, M., Chambers, J., Bullock, D., Palmer, A., Summerfield, A., Nelson, P. and Gatehouse, S. (2007). The binaural performance of a cross-talk cancellation system with matched or mismatched setup and playback acoustics. *The Journal of the Acoustical Society of America*, 121(2), pp.1056-1069.
13. Kuhn, G. (1977). Model for the interaural time differences in the azimuthal plane. *The Journal of the Acoustical Society of America*, 62(1), pp.157-167.
14. Xu, H., Xia, R., Li, J. and Yan, Y. (2017). An improved free-field cross-talk cancellation method based on the spherical head model. *Applied Acoustics*, 123, pp.47-54.
15. Kirkeby, O., Nelson, P., Hamada, H. and Orduna-Bustamante, F. (1998). Fast deconvolution of multichannel systems using regularization. *IEEE Transactions on Speech and Audio Processing*, 6(2), pp.189-194.
16. Fazi, F. M. and Hamdan, E. C. (2018). Stage Compression in Transaural Audio. *Audio Engineering Society Convention 144.*
17. Duda, R. and Martens, W. (1998). Range dependence of the response of a spherical head model. *The Journal of the Acoustical Society of America*, 104(5), pp.3048-3058.
18. Lacouture Parodi, Y. and Rubak, P. (2011). A Subjective Evaluation of the Minimum Channel Separation for Reproducing Binaural Signals over Loudspeakers. *Journal of the Audio Engineering Society.* 59(7/8), pp.487-497.
19. Hamdan, E. C. and Fazi, F. M. (2019). Three-channel Crosstalk Cancellation Mode Efficiency for Sources in the Far-Field. *Audio Engineering Society Conference: AES International Conference on Immersive and Interactive Audio.*
20. Farina, A. (2007). Advancements in Impulse Response Measurements by Sine Sweeps. *Audio Engineering Society Convention 122.*