forum acusticum 2026
8-12 September · Graz, Austria

ANALYSIS OF PLANT MATRIX PERTURBATIONS ON CROSSTALK CANCELLATION PERFORMANCE BASED ON SINGULAR VALUE DECOMPOSITION

Francesco Veronesi[1], Filippo Maria Fazi[1,2], Jacob Hollebon[2]
[1] *Institute of Sound and Vibration Research, University of Southampton, United Kingdom*
[2] *Audioscenic Ltd., Southampton, United Kingdom*

Abstract

This work presents a closed-form formula for predicting Crosstalk Cancellation (CTC) performance under random, uncorrelated additive perturbations of the plant matrix. Binaural reproduction over loudspeakers via CTC is highly sensitive to plant modelling errors, commonly introduced by the simplified plant models used in practice, which degrade the reproduced spatial audio. Tikhonov regularisation stabilises the inverse solution but reduces CTC performance, and predicting this trade-off under perturbations has not previously been addressed. This work derives, via Singular Value Decomposition (SVD) of the inverse problem in the frequency domain, a closed-form expression for the expected CTC performance loss caused by random plant perturbations under regularisation, validated against numerical simulations. The perturbation model, representing generic uncorrelated uncertainty such as diffuse reflections, provides an average-case bound rather than a predictive model for structured, deterministic error sources.

Keywords: *Crosstalk Cancellation, Binaural Audio Reproduction, Singular Value Decomposition, Perturbation Analysis*

1. Introduction

Crosstalk Cancellation (CTC) enables binaural audio reproduction over loudspeakers by controlling the sound field through inverse filtering [1], [2]. Its performance depends on system geometry, loudspeaker characteristics, and the accuracy of the plant model used to design the CTC filters, which ideally requires a precise matrix of acoustic transfer functions between each loudspeaker and the listener's ears, including individual Head-Related Transfer Functions (HRTFs) [3].

Direct measurement of the plant matrix is, however, often impractical, requiring complex, time-consuming calibration, and remains tied to the specific scenario in which it was acquired. It is therefore not robust to natural variations of the plant, such as environmental changes (e.g. temperature fluctuations [4]), HRTF individualisation [3], spatial errors [5], [6], the mismatch between anechoic measurement conditions and reflective real-world environments [7], or transducer variability across loudspeaker arrays [8], [9]. Simplified plant models, typically assuming idealised HRTFs and monopole loudspeakers, are therefore commonly used instead [3], [10]. Understanding the impact on CTC performance of the mismatch between such simplified models and the true physical plant is at the heart of this investigation.

Since the plant matrix may become ill-conditioned at certain frequencies, Tikhonov regularisation is applied to obtain robust CTC filter estimates [3], governing the trade-off between inversion accuracy and loudspeaker effort [11]: as regularisation increases, the solution shifts from minimising performance error to reducing loudspeaker effort [1].

Robustness to perturbations in the plant matrix is commonly investigated via the Singular Value Decomposition (SVD), which expresses the plant matrix through pairs of orthonormal (unitary) bases, with a set of singular values quantifying the efficiency of energy transfer between each corresponding pair of basis vectors, and collectively characterises the system's conditioning [2], [12]. In CTC, the SVD has been used to compute the condition number and analyse robustness [2], [7], [10], or relate results to the structure of the perturbation [11], [13], but not to predict performance under combined regularisation and perturbation constraints. This work presents preliminary progress in this direction, introducing a statistical

Corresponding author: *F. Veronesi@soton.ac.uk.*

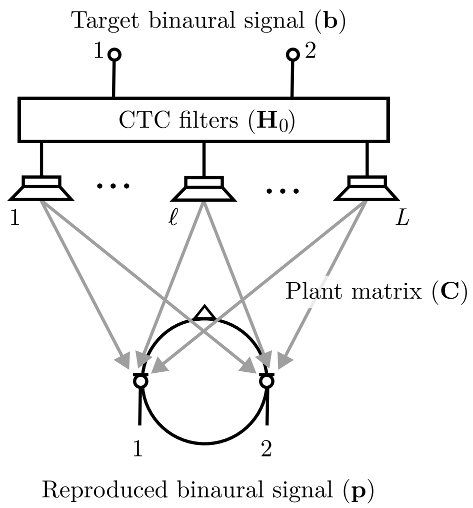

Figure 1: Diagram of a crosstalk cancellation system.

model in which each entry of the perturbation matrix is treated as a zero-mean, circularly symmetric complex Gaussian random variable, uncorrelated across matrix entries and with a common variance. This enables a closed-form prediction of the reproduced pressure and CTC performance as a function of the regularisation parameter and the perturbation variance, providing an average-case, statistically grounded bound on CTC degradation under generic, uncorrelated sources of uncertainty, rather than a predictive model for structured, deterministic error sources.

2. Theoretical Framework

2.1. Multichannel Crosstalk Cancellation

In a typical L-channel CTC system designed for a single listener, the sound pressure at two points corresponding to the listener's ears is controlled using L loudspeakers. The system is commonly described in the frequency domain as [1]

$$\mathbf{p} = \mathbf{C}\mathbf{H}_0\mathbf{b}, \tag{1}$$

where $\mathbf{p} \in \mathbb{C}^2$ denotes the complex pressure at the ears, $\mathbf{C} \in \mathbb{C}^{2\times L}$ is the plant matrix representing the transfer functions between each loudspeaker and ear, $\mathbf{H}_0 \in \mathbb{C}^{L\times 2}$ is the matrix of CTC filters, and $\mathbf{b} \in \mathbb{C}^2$ is the binaural signal to be reproduced. A schematic of the system is shown in Figure 1. Unless otherwise noted, all terms are considered frequency-dependent. The filters $\mathbf{H}_0$ are typically derived by minimising the squared error $\|\mathbf{p} - \mathbf{b}\|^2$. A modelling delay is usually introduced to maintain causality during filter design [1], though this is omitted here for brevity. As $\mathbf{C}$ is usually non-square and knowledge of the exact transfer functions between all loudspeakers and ears is not known, the inverse problem is solved by computing the Moore-Penrose pseudoinverse of a simplified model of the physical plant $\mathbf{C}$, defined as $\mathbf{C}_0$ [3]. To improve robustness against plant modelling inaccuracies and perturbations, and control the ill-conditioning of the system, Tikhonov regularisation is applied to obtain a robust estimate of the CTC filters [3], resulting in [1]:

$$\mathbf{H}_0 = \mathbf{C}_0^H[\mathbf{C}_0\mathbf{C}_0^H + \beta\mathbf{I}]^{-1}, \tag{2}$$

where β is the regularisation parameter, $\mathbf{I}$ is the 2×2 identity matrix, and $[\cdot]^H$ denotes the Hermitian transpose.

The objective of this study is to investigate how CTC performance is affected when the system is subject to modelling errors, i.e., when there exists a mismatch between $\mathbf{C}_0$ and $\mathbf{C}$. To this end, plant uncertainty is modelled by assuming that the response to the physical plant, $\mathbf{C}$, is equal to that of the modelled plant, $\mathbf{C}_0$, plus a matrix of additive plant uncertainties, $\Delta\mathbf{C}$, so that

$$\mathbf{C} = \mathbf{C}_0 + \Delta\mathbf{C}. \tag{3}$$

With this assumption, the pressure reproduced by the CTC system at the listener's ears is given by:

$$\mathbf{p} = \mathbf{C}\mathbf{H}_0\mathbf{b} \equiv \mathbf{C}\mathbf{C}_0^H\left(\mathbf{C}_0\mathbf{C}_0^H + \beta\mathbf{I}\right)^{-1}\mathbf{b}. \tag{4}$$

2.2. Application of the SVD

The Singular Value Decomposition (SVD) is adopted as the primary analytical tool to analyse the perturbation, or uncertainty, problem described here. The plant matrix $\mathbf{C}_0$ can be expressed in the form [12]

$$\mathbf{C}_0 = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{V}_0^H. \tag{5}$$

where the matrix $\mathbf{\Sigma}_0$ is of size $2\times L$, and is given by:

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \sigma_1 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 \end{bmatrix}, \tag{6}$$

and where σ_1 and σ_2 denote the singular values of the matrix $\mathbf{C}_0$, which are real-valued, non-negative and ordered in decreasing magnitude, i.e., $\sigma_1 \geq \sigma_2$. The matrix $\mathbf{U}_0$ is of size 2×2, and its columns comprise the left singular vectors of $\mathbf{C}_0$, while the matrix $\mathbf{V}_0$ is of size $L\times L$, and its columns comprise the right singular vectors of $\mathbf{C}_0$. The matrices $\mathbf{U}_0$ and $\mathbf{V}_0$ are unitary and satisfy the properties $\mathbf{U}_0^H\mathbf{U}_0 = \mathbf{U}_0\mathbf{U}_0^H = \mathbf{I}$ and $\mathbf{V}_0^H\mathbf{V}_0 = \mathbf{V}_0\mathbf{V}_0^H = \mathbf{I}$ [12]. From a physical perspective, the basis vectors forming the matrices $\mathbf{U}_0$ and $\mathbf{V}_0$ correspond to the *field pressure modes* and the *source strength modes* of the CTC system, respectively, while the squared singular values, σ_m^2, which are the eigenvalues of the Gram matrix $\mathbf{C}_0\mathbf{C}_0^H$, represent the *modal radiation efficiencies* [2].

The key idea of this analysis is to apply the SVD introduced in Eq. (5) to the matrices $\mathbf{C}$ and $\Delta\mathbf{C}$, and consequently to Equations (3) and (4). Accordingly, let the matrices $\mathbf{\Sigma}$ and $\Delta\mathbf{\Sigma}$ be introduced such that

$$\mathbf{\Sigma} = \mathbf{U}_0^H\mathbf{C}\mathbf{V}_0 = \underbrace{\mathbf{U}_0^H\mathbf{C}_0\mathbf{V}_0}_{\mathbf{\Sigma}_0\in\mathbb{R}^{2\times L}} + \underbrace{\mathbf{U}_0^H\mathbf{\Delta C}\ \mathbf{V}_0}_{\mathbf{\Delta\Sigma}\in\mathbb{C}^{2\times L}}. \tag{7}$$

Note that, in general, the matrix of singular value perturbations, $\mathbf{\Delta\Sigma}$, will not be diagonal [13].

Defining the transformed variables $\widetilde{\mathbf{p}} = \mathbf{U}_0^H \mathbf{p}$ and $\widetilde{\mathbf{b}} = \mathbf{U}_0^H \mathbf{b}$, the pressure reproduced at the listener's ears is given by

$$\begin{aligned} \widetilde{\mathbf{p}} &= (\mathbf{\Sigma}_0 + \mathbf{\Delta\Sigma})\mathbf{\Sigma}_0^T(\mathbf{\Sigma}_0\mathbf{\Sigma}_0^T + \beta\mathbf{I})^{-1}\widetilde{\mathbf{b}} = \\ &= \mathbf{\Sigma}_0^2(\mathbf{\Sigma}_0^2 + \beta\mathbf{I})^{-1}\widetilde{\mathbf{b}} + \mathbf{\Delta\Sigma}\cdot\mathbf{\Sigma}_0^T(\mathbf{\Sigma}_0^2 + \beta\mathbf{I})^{-1}\widetilde{\mathbf{b}} \end{aligned} \tag{8}$$

where $\mathbf{\Sigma}_0^2 = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2\right)$.
To keep the analysis simple, it is assumed that the CTC system is symmetric (i.e. the listener is centred), so that [2]

$$\mathbf{U}_0 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{9}$$

and that the binaural target signal, $\mathbf{b}$, is given by $\mathbf{b} = [1, 0]^T$, which also implies that $\widetilde{\mathbf{b}} = [1, 1]^T/\sqrt{2}$. These assumptions allow to rewrite Eq. (8) as

$$\widetilde{\mathbf{p}} = \frac{1}{\sqrt{2}}\begin{bmatrix} \sigma_1\zeta_1 \\ \sigma_2\zeta_2 \end{bmatrix} + \frac{1}{\sqrt{2}}\begin{bmatrix} \zeta_1\Delta\sigma_{11} + \zeta_2\Delta\sigma_{12} \\ \zeta_1\Delta\sigma_{21} - \zeta_2\Delta\sigma_{22} \end{bmatrix}, \tag{10}$$

where $\zeta_m = \sigma_m/(\sigma_m^2 + \beta)$, while $\Delta\sigma_{m,\ell}$ is the m,ℓ-th complex element of the matrix $\mathbf{\Delta\Sigma}$. Note, the dependency of ζ_m from the regularisation parameter, β, has been dropped to simplify the notation. Given that $\mathbf{p} = \mathbf{U}_0\widetilde{\mathbf{p}}$, it follows that

$$\mathbf{p} = \frac{1}{2}\begin{bmatrix} \sigma_1\zeta_1 + \sigma_2\zeta_2 \\ \sigma_1\zeta_1 - \sigma_2\zeta_2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \zeta_1\Delta_1^+ + \zeta_2\Delta_2^+ \\ \zeta_1\Delta_1^- + \zeta_2\Delta_2^- \end{bmatrix}, \tag{11}$$

where $\Delta_m^\pm = \Delta\sigma_{1m} \pm \Delta\sigma_{2m}$. In Eq. (11), the first term includes the regularisation error, while the second term represents the perturbation error. It is furthermore important to note that the equalisation term $(\sigma_1\zeta_1 + \sigma_2\zeta_2)/2$ ensures that the desired signal is reproduced at the listener's left ear, while the cancellation term $(\sigma_1\zeta_1 - \sigma_2\zeta_2)/2$ guarantees that the undesired crosstalk is cancelled, becoming exactly zero when $\beta = 0$ and the system is perfectly conditioned, i.e., $\sigma_1 = \sigma_2$ [2]. Within the perturbation matrix, the diagonal terms $(\Delta\sigma_{11}, \Delta\sigma_{22})$ represent how much the perturbation modifies the radiation efficiency of each mode, without energy leaking between them, whereas the off-diagonal terms $(\Delta\sigma_{12}, \Delta\sigma_{21})$ represent the cross-coupling between modes, i.e., the energy that the perturbation causes to leak from one mode into the other. Physically, this cross-coupling is what breaks the clean separation between the in-phase and out-of-phase modes that the nominal (unperturbed) system relies on for cancellation.

2.3. Modelling Perturbations as Random Variables

The random perturbation model adopted in this work assumes that each element of the plant perturbation matrix, $\mathbf{\Delta C}$, is a circularly symmetric complex Gaussian random variable, uncorrelated with the other entries of the matrix, with zero mean and variance s^2 (i.e. $\Delta C_{m\ell} \sim \mathcal{CN}(0, s^2)$). This implies a Rayleigh-distributed magnitude, with second moment $\mathrm{Var}[\Delta C_{m\ell}] \equiv \mathbb{E}[|\Delta C_{m\ell}|^2] = s^2$, and a phase uniformly distributed over $[0, 2\pi)$ [14]. Moreover, since the distinct $\Delta C_{m\ell}$ are assumed uncorrelated, the cross-covariance between any two entries reduces to $\mathbb{E}[\Delta C_{m\ell} \cdot \Delta C^\star_{m'\ell'}] = s^2 \cdot \delta_{mm'}\delta_{\ell\ell'}$, equal to s^2 for the same element and zero for any distinct pair of elements. This choice corresponds to the statistical model of a diffuse sound field, composed of plane waves coming from all directions, with statistically independent and uniformly distributed phases and zero-mean, uncorrelated amplitudes [15].
Since the singular matrices $\mathbf{U}_0$ and $\mathbf{V}_0$ are unitary, then the elements of the transformed perturbation matrix, $\mathbf{\Delta\Sigma} = \mathbf{U}_0^H\mathbf{\Delta C}\ \mathbf{V}_0$, are linear combinations of the elements of $\mathbf{\Delta C}$ with unitary weights, and thus they remain complex random variables characterised by identical properties [14], namely:

(i) $\mathbb{E}[\Delta\sigma_{m\ell}] = 0$

(ii) $\mathrm{Var}[\Delta\sigma_{m\ell}] \equiv \mathbb{E}[|\Delta\sigma_{m\ell}|^2] = s^2$

(iii) $\mathbb{E}[\Delta\sigma_{m\ell} \cdot \Delta\sigma^\star_{m'\ell'}] = s^2 \cdot \delta_{mm'}\delta_{\ell\ell'}$

2.4. Predicting CTC Performance

The properties introduced above are used to analyse the statistics of the sound pressure reproduced by the CTC system, and predict the CTC performance achieved by the system when it is subject to the modelled random perturbations. To this end, given the chosen target binaural signal, CTC performance is investigated by analysing the ratio of the expected acoustic energies at the bright and dark ears,

$$\mathrm{CTCL}(\beta, s^2) = 10\log_{10}\left(\frac{\mathbb{E}[|p_1|^2]}{\mathbb{E}[|p_2|^2]}\right), \tag{12}$$

where the expectation operator, $\mathbb{E}[\cdot]$, is taken over the ensemble of plant perturbations. The metric $\mathrm{CTCL}(\beta, s^2)$, defined in Eq. (12) as the ratio between the expected acoustic energies received at the two ears, differs from the expected ratio of the reproduced pressures obtained for each individual perturbation, $\mathbb{E}[|p_1|^2/|p_2|^2]$. The latter definition would be more consistent with the commonly adopted definition of acoustic contrast, or channel separation, used to quantify the degree of crosstalk attenuation achieved between the listener's ears in the absence of perturbations [3]. The former, however, resembles the definition of acoustic contrast adopted in the research field of personal sound zones [11]. These two definitions correspond to different estimators of CTC performance: $\mathrm{CTCL}(\beta, s^2)$, as defined in Eq. (12), is less sensitive to perturbations that produce exceptionally large or small contrast values, since the expectation is taken over the acoustic energies themselves rather than over their ratio. This choice is further justified by the mathematical tractability it affords, following directly from the properties of the modelled perturbations introduced above. Future work will investigate whether,

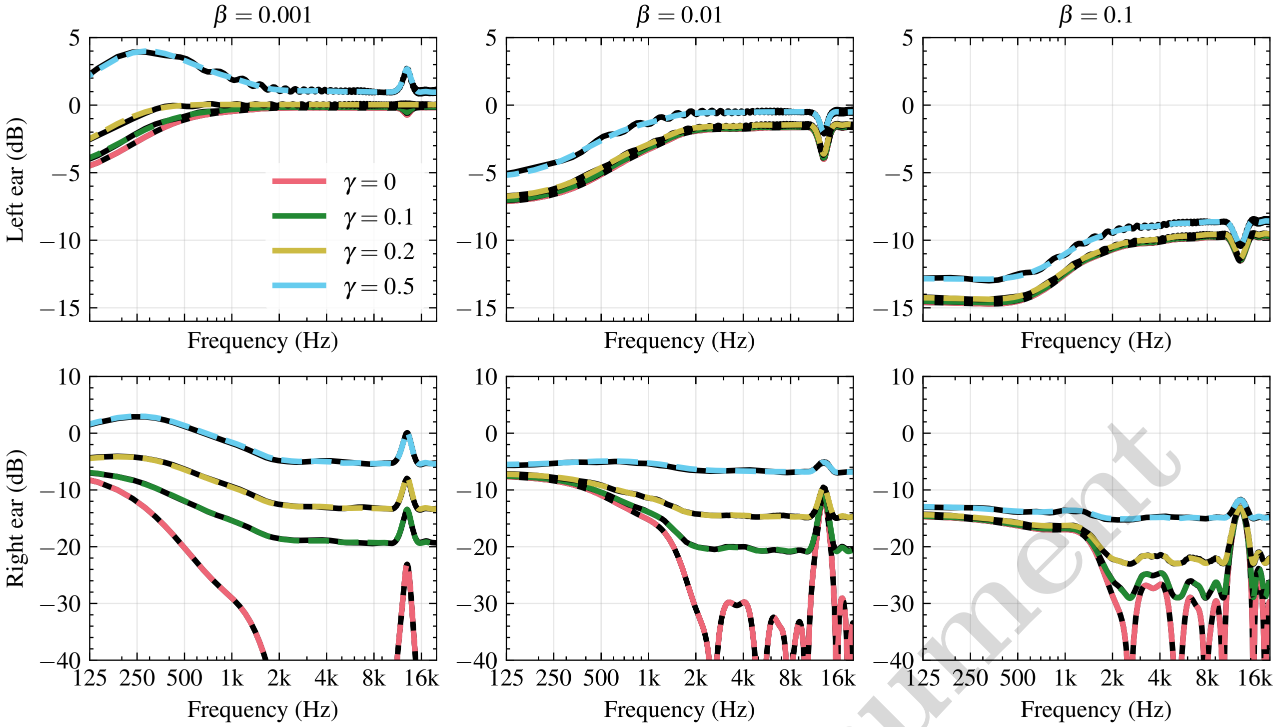

Figure 2: Comparison between predicted and simulated reproduced pressure for different regularisation values in the presence of random and uncorrelated complex additive perturbations, whose variance s^2 is controlled by the factor γ. The black solid lines consistently denotes the reference averaged reproduced pressure.

and how, this analysis can be extended to the expected acoustic contrast.

A closed-form solution to Eq. (12) can be obtained by combining the result presented in Eq. (11) with the properties *(i)*, *(ii)*, and *(iii)* previously defined. In particular, for the pressure at the left and right ears, it can be shown that:

$$\begin{cases} \mathbb{E}\left[|p_1|^2\right] = \mathbb{E}\left[\left|\frac{1}{2}\left(\sigma_1\zeta_1 + \sigma_2\zeta_2 + \zeta_1\Delta_1^+ + \zeta_2\Delta_2^+\right)\right|^2\right] \\ \qquad\qquad = \frac{1}{4}\left(\sigma_1\zeta_1 + \sigma_2\zeta_2\right)^2 + \frac{s^2}{2}\left(\zeta_1^2 + \zeta_2^2\right), \\ \mathbb{E}\left[|p_2|^2\right] = \frac{1}{4}\left(\sigma_1\zeta_1 - \sigma_2\zeta_2\right)^2 + \frac{s^2}{2}\left(\zeta_1^2 + \zeta_2^2\right). \end{cases} \tag{13}$$

It follows that

$$\frac{\mathbb{E}[|p_1|^2]}{\mathbb{E}[|p_2|^2]} = \frac{\frac{1}{4}\left(\sigma_1\zeta_1 + \sigma_2\zeta_2\right)^2 + \frac{s^2}{2}\left(\zeta_1^2 + \zeta_2^2\right)}{\frac{1}{4}\left(\sigma_1\zeta_1 - \sigma_2\zeta_2\right)^2 + \frac{s^2}{2}\left(\zeta_1^2 + \zeta_2^2\right)}, \tag{14}$$

which can be used to compute $\mathrm{CTCL}(\beta, s^2)$ as defined in Eq. (12). Eq. (14) shows that the performance of a regularised CTC system subject to random and uncorrelated complex additive perturbations to the plant matrix can be predicted from the singular values of the unperturbed plant matrix, $\mathbf{C}_0$, provided that the variance of the random perturbations, s^2, is known.

3. Numerical Validation

In the following, the validity of Equation (14) is tested for a linear array of five loudspeakers with an inter-channel distance of 15 cm. For simplicity, the plant matrix is based on monopole-like loudspeakers and a rigid-sphere head model of radius 87.5 mm, positioned 1 m away from the array in the central symmetric location, with head scattering accounted for through the analytical solution to the scattering problem of spheres [3]. It is further assumed that the CTC system is subject to perturbations that satisfy the properties defined above, with s defined as the root-mean-square value of the matrix 2-norm of the unperturbed $\mathbf{C}_0$ computed over all frequencies, and scaled by a positive factor γ. In the numerical simulations the values $\gamma = 0$, 0.1, 0.2, and 0.5 are considered, with the expectation approximated by an average over 10^3 repetitions. Given the known dependency of CTC performance on regularisation, the simulations are further conducted for three regularisation values: $\beta = 10^{-3}$, 10^{-2}, and 10^{-1}.

3.1. Reproduced Pressure

Figure 2 presents the predicted and simulated expected pressure reproduced at the listener's ears, for the different choices of regularisation and perturbation gain γ. Coloured lines represent the quantities predicted by Eq. (13), while black lines show the reference values obtained from numerical simulation and averaging. The close agreement between the two con-

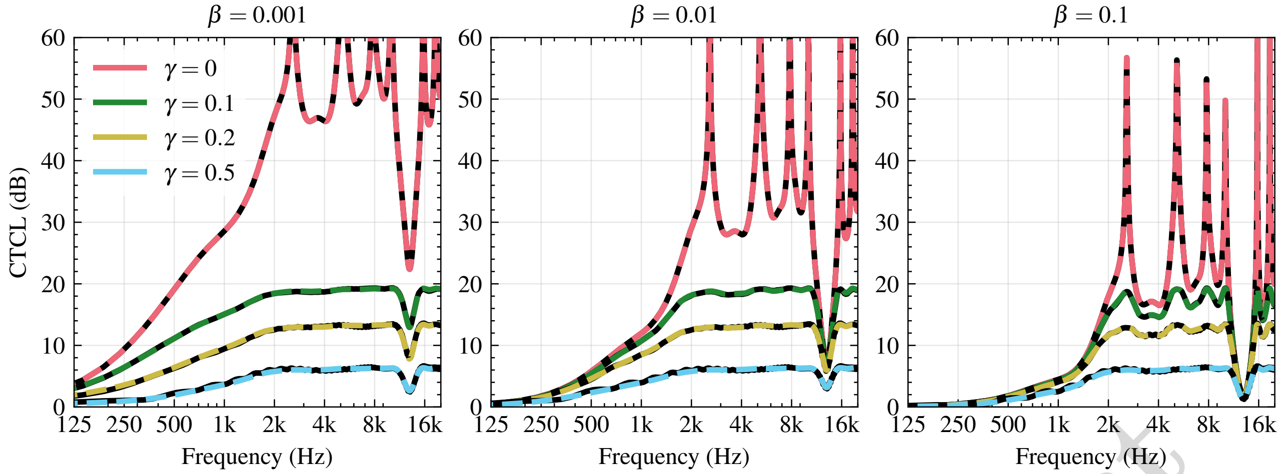

Figure 3: Comparison between predicted and simulated CTC performance for different regularisation values in the presence of random and uncorrelated complex additive perturbations, whose variance s^2 is controlled by the factor γ. The black solid lines consistently denotes the reference averaged CTC level.

firms the validity of the analytical model in Eq. (13). Beyond this validation, the figure also reveals several interesting properties of the regularised system when subject to perturbations, particularly when read alongside Figure 4, which shows the singular values σ_1, σ_2 of the unperturbed plant matrix $\mathbf{C}_0$.

For small perturbations, the reproduced pressure and CTC performance are governed mainly by the smallest singular value, σ_2, associated with the out-of-phase mode, whose efficiency is known to be poor at low frequencies [2], [10]. This can be shown by noting that, for $\sigma_1 \gg \beta$ (cf. Figure 4) which implies $\sigma_1^2 + \beta \approx \sigma_1^2$, and for $s^2 = 0$, Eq. (13) gives

$$\begin{cases} \mathbb{E}\left[|p_1|^2\right] = \left(1 - \dfrac{\beta/2}{\sigma_2^2 + \beta}\right)^2, \\ \mathbb{E}\left[|p_2|^2\right] = \left(\dfrac{\beta/2}{\sigma_2^2 + \beta}\right)^2 \end{cases} \tag{15}$$

For small regularisation values, the reproduced pressure is governed by both the perturbations and the conditioning of the system. In particular, the energy of the reproduced pressure is shaped mainly by the condition number, $\kappa = \sigma_1/\sigma_2$, of the unperturbed plant matrix $\mathbf{C}_0$, and scaled by the perturbation variance, s^2. This can be shown by assuming $\beta = 0$, which gives

$$\begin{cases} \mathbb{E}\left[|p_1|^2\right] = 1 + \dfrac{s^2}{2\sigma_1^2}\left(1 + \kappa^2\right), \\ \mathbb{E}\left[|p_2|^2\right] = \dfrac{s^2}{2\sigma_1^2}\left(1 + \kappa^2\right) \end{cases} \tag{16}$$

For regularisation values greater than the smallest singular value, CTC becomes poorly effective and reproduces the same acoustic pressure at both ears, whose energy decreases for larger β and is amplified for larger perturbation variance, s^2. As widely reported in the literature, regularisation improves numerical conditioning and system robustness, and limits loudspeaker effort and dynamic range, at the cost of penalising the effective performance and sensitivity of the CTC system [10], [16]. This can be shown by assuming $\beta \gg \sigma_2$, which gives

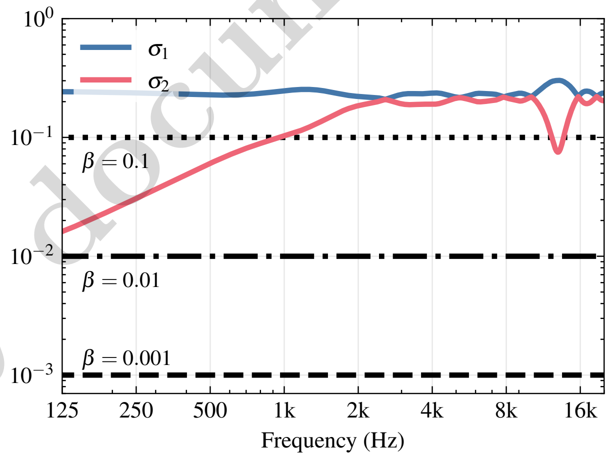

Figure 4: Singular values σ_1, σ_2 of the unperturbed plant matrix $\mathbf{C}_0$ compared with the selected values of the regularisation parameter β.

$$\mathbb{E}\left[|p_1|^2\right] = \mathbb{E}\left[|p_2|^2\right] = \frac{1}{4}\left(\frac{\sigma_1^2}{\sigma_1^2 + \beta}\right)^2 + \frac{s^2}{2}\left(\frac{\sigma_1}{\sigma_1^2 + \beta}\right)^2 \tag{17}$$

Figures 2 and Figure 4 show that the conditions that enable Eq. (17) are met for the plant matrix $\mathbf{C}_0$ under investigation at frequencies below 1 kHz.

3.2. CTC Level

Figure 3 presents the comparison between the predicted and simulated CTC performance obtained for the plant under investigation and for the different selected values of β and s^2. The results demonstrate that Equation (14) accurately predicts the modelled $\mathrm{CTCL}(\beta, s^2)$ across the different regularisation values and perturbation variances considered, confirming the consistency between the analytical formulation and the numerical simulations.

4. Conclusions

This work introduced a statistical model of plant perturbations for CTC systems, treating each entry of the perturbation matrix as a zero-mean, circularly symmetric complex Gaussian random variable, uncorrelated across matrix entries and with a common variance, s^2. This formulation, consistent with the statistical model of a diffuse sound field, enabled a closed-form prediction of the reproduced pressure and CTC performance as a function of the regularisation parameter β and the perturbation variance s^2, validated against numerical simulations. The analysis revealed that regularisation governs a trade-off between CTC performance and robustness to perturbations, with two distinct regimes. For small β, performance is dominated by the conditioning of the unperturbed plant, through its condition number, with perturbations scaling this contribution through s^2. As β increases beyond the smallest singular value, this dependence on conditioning is progressively suppressed, and the perturbation contributes equally to the pressure reproduced at both ears, at the cost of an increasingly poor CTC performance. These results provide a practical criterion for selecting regularisation values that balance CTC performance against robustness to plant perturbations, and motivate future work extending the analysis to structured, correlated perturbation sources, such as listener displacement, loudspeaker misplacement or directivity, inaccurate HRTF modelling, which the present perturbation model does not capture.

5. Acknowledgments

This work was supported by Audioscenic Limited as part of the project "Understanding, Modelling and Improving Various Aspects of Crosstalk Cancellation Systems" (Contract No 9747105).

References

[1] O. Kirkeby, P. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 189–194, 1998.

[2] E. C. Hamdan and F. M. Fazi, "Ideal focusing and optimally-conditioned systems in sound field control with loudspeaker arrays," *J. Sound Vib.*, vol. 509, p. 116 218, 2021.

[3] J. Hollebon, E. Hamdan, and F. Fazi, "A comparison of the performance of HRTF models in inverse filter design for crosstalk cancellation," in *Proceedings of the Institute of Acoustics*, vol. 41, Curran Associates, Inc., May 2020.

[4] M. Olsen and M. B. Møller, "Sound Zones: On the Effect of Ambient Temperature Variations in Feed-Forward Systems," in *AES Convention 142*, AES, Journal of the Audio Engineering Society, May 2017.

[5] X. Qiu, B. Masiero, and M. Vorländer, "Channel separation of crosstalk cancellation systems with mismatched and misaligned sound sources," *J. Acoust. Soc. Am.*, vol. 126, no. 4, pp. 1796–1806, Oct. 2009.

[6] D. Carvalho, Y. Ushakov, J. Hollebon, and M. Simon, "The role of listener position sensing in dynamic crosstalk cancellation with headrest speakers," in *2026 AES 6th International Automotive Audio Conference*, AES, Journal of the Audio Engineering Society, Jul. 2026.

[7] Y. Li, L. Wang, and J. Reiss, "Impact of mismatched room acoustic modeling on transaural reproduction with loudspeaker arrays," in *AES Convention 155*, AES, Journal of the Audio Engineering Society, Oct. 2023.

[8] S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 7, pp. 2123–2133, 2012.

[9] F. Veronesi, F. Fazi, and J. Hollebon, "Impact of impedance mismatch on crosstalk cancellation with micro-speaker arrays," English, Forum Acusticum Euronoise 2025., 2025.

[10] T. Takeuchi and P. A. Nelson, "Optimal source distribution for binaural synthesis over loudspeakers," *J. Acoust. Soc. Am.*, vol. 112, no. 6, pp. 2786–2797, Dec. 2002.

[11] Y. Li, L. Wang, and J. Reiss, "Neural Regularization for Personal Sound Zones," in *AVARIG 2026: Audio for Virtual and Augmented Reality and Immersive Games*, AES, Journal of the Audio Engineering Society, Jul. 2026.

[12] P. A. Nelson, "A review of some inverse problems in acoustics," *Int. J. Acoust. Vib.*, vol. 6, no. 3, pp. 118–34, 2001.

[13] S. J. Elliott, *Signal Processing for Active Control* (Signal Processing and its Applications). Elsevier Science, 2001.

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[15] C. Van Hoorickx and E. P. B. Reynders, "Numerical realization of diffuse sound pressure fields using prolate spheroidal wave functions," *The Journal of the Acoustical Society of America*, vol. 151, no. 3, pp. 1710–1721, Mar. 2022.

[16] E. Y. Choueiri, "Optimal crosstalk cancellation for binaural audio with two loudspeakers," *Princeton University*, vol. 28, 2008.