

TOWARD OPTIMAL ANC: ESTABLISHING MUTUAL INFORMATION LOWER BOUND

Anonymous authors

Paper under double-blind review

ABSTRACT

Active Noise Cancellation (ANC) algorithms aim to suppress unwanted acoustic disturbances by generating anti-noise signals that destructively interfere with the original noise in real time. Although recent deep learning-based ANC algorithms have set new performance benchmarks, there remains a shortage of theoretical limits to rigorously assess their improvements. To address this, we derive a unified lower bound on cancellation performance composed of two components. The first component is information-theoretic: it links residual error power to the fraction of disturbance entropy captured by the anti-noise signal, thereby quantifying limits imposed by information-processing capacity. The second component is support-based: it measures the irreducible error arising in frequency bands that the cancellation path cannot address, reflecting fundamental physical constraints. By taking the maximum of these two terms, our bound establishes a theoretical ceiling on the Normalized Mean Squared Error (NMSE) attainable by any ANC algorithm. We validate its tightness empirically on the NOISEX dataset under varying reverberation times, demonstrating robustness across diverse acoustic conditions.

1 INTRODUCTION

Active Noise Cancellation algorithms Lueg (1936); Nelson & Elliott (1991); Fuller et al. (1996); Hansen et al. (1997); Kuo & Morgan (1999) aim to improve listening environments by generating anti-noise signals that destructively interfere with unwanted disturbances in real time. Classical adaptive-filtering methods—most notably the Filtered-x Least Mean Squares (FxLMS) algorithm Boucher et al. (1991) and its variants Das & Panda (2004); Patra et al. (1999); Tan & Jiang (2001); Kuo & Wu (2005); Tobias & Seara (2005), have been used to tackle this problem, but they often struggle in non-stationary or highly reverberant settings and offer limited guidance on how close they operate to fundamental performance limits.

In recent years, deep learning-based ANC models such as DeepANC Zhang & Wang (2021), Attentive Recurrent Networks (ARN) Pandey & Wang (2022), and DeepASC Mishaly et al. (2025) learn end-to-end mappings from reference and error microphones to cancellation waveforms, outperforming classical signal processing methods. Variants employing recurrent CNNs Park et al. (2023); Mostafavi & Cha (2023), autoencoders Singh et al. (2024), and fully connected nets Pike & Cheer (2023), together with selective fixed-filter schemes Shi et al. (2020); Park & Park (2023), Kalman-filter hybrids Luo et al. (2023c), and attention modules Zhang et al. (2023a), further enhance robustness and low-latency control. However, no unified theory specifies the minimal achievable error under given acoustic and algorithmic constraints, complicating optimality assessment and targeted architecture design. To fill this gap, we derive a unified lower bound on ANC performance that blends two complementary perspectives:

- An information-theoretic term that links the residual error power to the fraction of disturbance entropy captured by the anti-noise signal—quantifying limits imposed by information-processing capacity.
- A support-based term that measures the irreducible error arising from frequency bands unsupported by the cancellation path—reflecting fundamental physical constraints.

By taking the maximum of these two terms, our bound establishes the first general theoretical ceiling on the NMSE attainable by any ANC algorithm, regardless of model complexity. We stress that this

054 lower bound framework adequate to every possible ANC algorithm, may it be a neural network
 055 trained on dataset, a linear filtering, or adaptive mechanism.

056 We validate our contributions through extensive experiments on the NOISEX benchmark under
 057 varying reverberation times. The primary contributions of this work are summarized as follows:
 058

- 059 • **Theoretical benchmark:** We derive a unified lower bound on ANC performance that ac-
 060 counts for both information-processing and spectral-support constraints.
- 061 • **Empirical validation:** We confirm the tightness of our bound over various noise types and
 062 for different reverberation times. Boundary implementations are released in the supple-
 063 mentary material.
 064

065 2 RELATED WORK

066 The foundational principle of destructive interference was introduced by Lueg, who demonstrated
 067 that appropriately phased signals can cancel unwanted oscillations (Lueg, 1936). Early adaptive
 068 schemes adopted the Least Mean Squares (LMS) criterion to track variations in amplitude and
 069 phase, achieving robust echo cancellation in acoustic feedback paths Burgess (1981); Nelson &
 070 Elliott (1991); Fuller et al. (1996); Hansen et al. (1997); Kuo & Morgan (1999). To account for the
 071 dynamics of the primary and secondary acoustic paths, the FxLMS algorithm was proposed, with
 072 subsequent analysis quantifying the impact of secondary-path inversion errors on convergence and
 073 stability Boucher et al. (1991).
 074

075 Extensions of the LMS framework have focused on mitigating real-world nonlinearities. Functional-
 076 Link Artificial Neural Networks (FLANN) were integrated into LMS filters to compensate for dis-
 077 tortion in the control path Patra et al. (1999). Bilinear filtering improved modeling of mild non-
 078 linearities Kuo & Wu (2005). A tangential-hyperbolic adaptation further captured loudspeaker sat-
 079 uration effects, ensuring stable cancellation under high-amplitude conditions Ghasemi et al. (2016).
 080

081 Data-driven models have more recently transformed ANC by learning end-to-end mappings from
 082 sensor measurements to cancellation signals. Convolutional-LSTM networks estimate both ampli-
 083 tude and phase responses Zhang & Wang (2021), and recurrent convolutional architectures exploit
 084 temporal dependencies for improved active cancellation prediction Park et al. (2023); Mostafavi &
 085 Cha (2023); Cha et al. (2023). Fully connected network also applied to ANC Pike & Cheer (2023)
 086 as well as Autoencoder-based encodings scheme have been employed to extract latent features and
 087 generalize across diverse noise conditions Singh et al. (2024). Selective Fixed-filter ANC (SFANC)
 088 leverages pre-trained filter banks chosen by lightweight neural controllers, enabling rapid adaptation
 089 without full retraining Shi et al. (2020; 2022; 2023a); Luo et al. (2022); Park & Park (2023).

090 The works of Luo et al. (2023b;a; 2024) have focused on synthesizing optimized filter banks for
 091 selective fixed-filter ANC. Multichannel ANC systems exploit spatial diversity to achieve superior
 092 noise attenuation. Deep multichannel controllers learn inter-channel relationships for ANC Zhu et al.
 093 (2021); Zhang & Wang (2023); Antoñanzas et al. (2023); Xiao et al. (2023); Zhang et al. (2023b);
 094 Shi et al. (2023b; 2024). Zhang et al. (2023a) proposed an attention-driven framework for real-time
 095 ANC by integrating the Attentive Recurrent Network of Pandey & Wang (2022), and Zhang et al.
 096 (2022) demonstrated real-time ANC using attentive recurrent architectures for online adaptation. In
 097 addition, metaheuristic strategies have been explored, Zhou et al. (2023) applied a genetic algorithm
 098 to optimize ANC filters, and Ren & Zhang (2022) employed a bee colony optimization scheme to
 099 refine control coefficients.

100 3 BACKGROUND - TYPICAL ANC SETUP

101 As we can observe in Figure 1, a typical ANC feedforward setup, the signal captured by the ref-
 102 erence microphone is denoted as $x(n)$ and is transmitted through the primary path $P(z)$, and their
 103 convolution produces the primary signal denoted $d(n)$ and defined as $d(n) = P(z) * x(n)$. The
 104 signal $x(n)$ and the signal captured by the error microphone $e(n)$ are used as inputs of the ANC
 105 system to generate the cancelation signal $y(n)$, which is played by the loudspeaker f_{LS} . The pro-
 106 duced signal $f_{LS}\{y(n)\}$, supposed to eliminate the undesirable noise around the error microphone,
 107 is then transmitted through the secondary path $S(z)$ to generate the anti-signal $a(n)$ defined by,

108 $a(n) = S(z) * f_{LS}\{y(n)\}$ Then, the error signal $e(n)$ is obtained by calculating the difference be-
 109 tween the primary and secondary path's outputs, $e(n) = d(n) - a(n)$. The ANC controller aims
 110 to reduce the error signal $e(n)$ to zero, achieving optimal noise cancellation. In the feedback ANC
 111 method, only $e(n)$ is used to generate the canceling signal, focusing on minimizing the residual
 112 noise captured by the error microphone.

114 4 FUNDAMENTAL LIMITS ON CANCELLATION PERFORMANCE: 115 INFORMATION-THEORETIC BOUNDS 116

117 We address the problem of canceling an undesired distur-
 118 bance signal, denoted as a discrete-time random pro-
 119 cess $d(n)$, which arises from filtering a reference source
 120 signal $x(n)$ through a system characterized by the pri-
 121 mary path transfer function $P(z)$. This scenario is
 122 canonical in applications such as ANC, acoustic echo
 123 cancellation, and adaptive interference mitigation. The
 124 objective is to synthesize a cancellation signal $y(n)$
 125 using an adaptive algorithm or model, potentially involv-
 126 ing non-linear transformations $f(\cdot)$ (e.g., modeling sec-
 127 ondary path dynamics or actuator non-linearities), such
 128 that the residual error signal $e(n) = d(n) - f(y(n))$
 129 is minimized in a suitable sense, typically minimizing its
 130 NMSE or power.

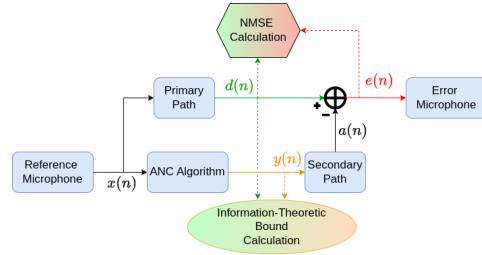


Figure 1: High-level schematic of a feedforward ANC system with lower-bound calculation.

130 To establish fundamental benchmarks against which practical algorithms can be evaluated, we derive
 131 a lower bound on the achievable cancellation performance rooted in information theory. This bound
 132 quantifies the irreducible error stemming from the information processing limitations inherent in
 133 generating $y(n)$ based on available observations.

134 We posit that the reference signal $x(n)$ is a zero-mean Wide-Sense Stationary (WSS) random pro-
 135 cess. Consequently, if the primary path $P(z)$ represents a linear time-invariant (LTI) system, the
 136 disturbance signal $d(n) = P(z)x(n)$ is also a zero-mean WSS process. Furthermore, we assume
 137 the synthesized signal $y(n)$ and the disturbance $d(n)$ are jointly WSS and ergodic. Let $S_{dd}(e^{j\omega})$
 138 denote the power spectral density (PSD) of $d(n)$. The average power of the disturbance is given by
 139 its variance $\sigma_d^2 = \mathbb{E}[|d(n)|^2]$. A cornerstone connecting the time-domain correlation structure of
 140 stationary signals to their frequency-domain power distribution is the Wiener-Khinchin Theorem.

141 [Wiener-Khinchin Theorem for Discrete-Time WSS Processes] Let $v(n)$ be a discrete-time WSS
 142 random process with Auto-Correlation Function (ACF) defined as $R_{vv}(\tau) = \mathbb{E}[v(n)v^*(n - \tau)]$.
 143 Its Power Spectral Density (PSD), $S_{vv}(e^{j\omega})$, is the Discrete-Time Fourier Transform (DTFT) of its
 144 ACF:

$$145 S_{vv}(e^{j\omega}) = \sum_{\tau=-\infty}^{\infty} R_{vv}(\tau)e^{-j\omega\tau}$$

146 Conversely, the ACF can be recovered from the PSD via the inverse DTFT:

$$147 R_{vv}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{vv}(e^{j\omega})e^{j\omega\tau} d\omega$$

148 The average power of the process is the integral of its PSD over the fundamental frequency interval
 149 (or equivalently, the ACF at lag zero):

$$150 \mathbb{E}[|v(n)|^2] = R_{vv}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{vv}(e^{j\omega})d\omega$$

151 This theorem justifies the use of spectral representations for power calculations under the WSS
 152 assumption. For our disturbance signal $d(n) = P(z)x(n)$, its PSD is related to the PSD of the
 153 source $S_{xx}(e^{j\omega})$ by $S_{dd}(e^{j\omega}) = |P(e^{j\omega})|^2 S_{xx}(e^{j\omega})$, assuming $P(z)$ is stable. The average power
 154 of the disturbance is thus:

$$155 \sigma_d^2 = \mathbb{E}[|d(n)|^2] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{dd}(e^{j\omega}) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |P(e^{j\omega})|^2 S_{xx}(e^{j\omega}) d\omega \quad (1)$$

Now, consider the information-theoretic perspective. Let $H(d)$ denote the differential entropy rate of the disturbance process $d(n)$, assuming it exists and is finite. Let $I(y; d)$ denote the mutual information rate between the synthesized signal process $y(n)$ and the disturbance process $d(n)$. Mutual Information measures the information that $d(n)$ and $y(n)$ share and therefore quantifies how much knowing one of these variables reduces uncertainty about the other. Intuitively, effective cancellation requires $y(n)$ (or more precisely, $f(y(n))$), where f is some processing function, to be highly informative about $d(n)$. Building upon rate-distortion theory concepts, a lower bound on the residual error power $\sigma_e^2 = \mathbb{E}[|e(n)|^2]$ can be established:

[Information-Theoretic Bound on Cancellation Error] Under the assumption that $y(n)$ and $d(n)$ are jointly WSS and ergodic, and that $d(n)$ possesses a finite differential entropy rate $H(d)$, the average power of the residual error $e(n) = d(n) - f(y(n))$ is lower bounded by:

$$\sigma_e^2 = \mathbb{E}[|e(n)|^2] \geq \sigma_d^2 \left(1 - \frac{I(y(n); d(n))}{H(d(n))} \right) \quad (2)$$

where $\sigma_d^2 = \mathbb{E}[|d(n)|^2]$ is the average power of the disturbance signal.

Let $d(n)$ and $y(n)$ be jointly wide-sense stationary (WSS) and ergodic random processes. Suppose $d(n)$ has finite variance σ_d^2 and finite differential entropy rate $H(d)$. Define the residual error as:

$$e(n) = d(n) - f(y(n))$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is any deterministic mapping (possibly nonlinear). Let $\hat{d}(n) = f(y(n))$ be the estimate of $d(n)$ based on $y(n)$.

From information theory, the mutual information between $d(n)$ and $y(n)$ satisfies:

$$I(d(n); y(n)) \geq R(D) \quad (3)$$

where $R(D)$ is the rate-distortion function of the source $d(n)$ at distortion level $D = \mathbb{E}[(d(n) - \hat{d}(n))^2] = \sigma_e^2$.

For general stationary ergodic sources with finite differential entropy rate $H(d)$, the Shannon lower bound on the rate-distortion function provides:

$$R(D) \geq H(d) - \frac{1}{2} \log(2\pi e D) \quad (4)$$

Solving equation 4 for D yields:

$$D \geq \frac{1}{2\pi e} \cdot \exp(2(H(d) - I(d(n); y(n)))) \quad (5)$$

However, this expression is difficult to interpret directly. Instead, we derive a more interpretable but looser bound by noting that mutual information is always upper bounded by entropy:

$$I(d(n); y(n)) \leq H(d)$$

Hence, we can define the normalized mutual information ratio:

$$\alpha := \frac{I(d(n); y(n))}{H(d)} \in [0, 1]$$

Recall, that for gaussian noise, the entropy is maximized and the following holds:

$$\sigma_d^2 = \frac{1}{2\pi e} e^{2H(d)}$$

Rearranging terms, of equation 5, and keep in mind that the ratio is always smaller than one, so for negative exponent it is upper bound,

$$D \geq \frac{1}{2\pi e} \cdot \exp(2H(d)) \cdot \exp\left(-\frac{I(d(n); y(n))}{H(d)}\right)$$

expanding the second exponent as Taylor series, and taking the first order yields,

$$\sigma_e^2 \geq \sigma_d^2 \left(1 - \frac{I(d(n); y(n))}{H(d)} \right) \quad (6)$$

Intuitively, this ratio measures the fraction of the information in $d(n)$ that is captured by $y(n)$. If $\alpha = 1$, perfect reconstruction is possible in principle, and $\sigma_e^2 \rightarrow 0$. If $\alpha = 0$, then $y(n)$ carries no information about $d(n)$, and $\sigma_e^2 \geq \sigma_d^2$.

We thus propose a linear lower bound on the distortion based on this ratio:

$$\sigma_e^2 \geq \sigma_d^2(1 - \alpha) = \sigma_d^2 \left(1 - \frac{I(d(n); y(n))}{H(d)} \right) \quad (7)$$

which completes the proof.

This fundamental bound highlights that the efficacy of any cancellation strategy is limited by how much information about the target disturbance $d(n)$ is effectively encoded in the synthesized signal $y(n)$. If $y(n)$ carries no information about $d(n)$ ($I(y; d) = 0$), the bound implies $\sigma_e^2 \geq \sigma_d^2$, meaning no cancellation is achieved on average. Conversely, if $y(n)$ allows for perfect reconstruction such that $f(y(n)) = d(n)$, then $I(y; d)$ must approach $H(d)$ (under suitable conditions, e.g., Gaussianity), and the bound approaches zero. However, achieving $I(y; d) = H(d)$ may be impossible in practice due to several factors not explicitly captured in the bound’s derivation but reflected in the properties of $y(n)$ generated by real-world algorithms. These include:

- **Causality:** Practical algorithms operate causally, potentially requiring a significant observation window (context) to accurately estimate latent variables (e.g., channel coefficients, internal states) needed to generate an effective $y(n)$. During initialization or adaptation phases, $I(y; d)$ might be substantially lower than its asymptotic potential.
- **Model Mismatch:** The assumed structure relating $x(n)$ to $d(n)$ (given by $P(z)$) or the model used to generate $y(n)$ (including $f(\cdot)$) might not perfectly match reality.
- **Computational Constraints:** The complexity of the algorithm generating $y(n)$ might limit its ability to extract and encode all relevant information about $d(n)$.

Therefore, the bound equation 2 represents an idealized limit conditioned on the statistical properties (specifically, the joint distribution encoded in $I(y; d)$) of the output $y(n)$ produced by a given cancellation system. We can express this bound in terms of the NMSE, often reported in decibels (dB), defined as $\text{NMSE} = \mathbb{E}[|e(n)|^2] / \mathbb{E}[|d(n)|^2]$. Substituting equation 1 and converting to dB:

$$\text{NMSE}_{\text{dB}} = 10 \log_{10} \left(\frac{\mathbb{E}[|e(n)|^2]}{\mathbb{E}[|d(n)|^2]} \right) \geq 10 \log_{10} \left(1 - \frac{I(y; d)}{H(d)} \right) + 10 \log_{10} (\|P(z)\|_2^2) \quad (8)$$

This information-theoretic bound is inherently dependent on the specific algorithm (as it determines the statistics of $y(n)$ and thus $I(y; d)$). It serves as a valuable theoretical ceiling, quantifying performance limitations arising purely from the representational capacity and information-processing capabilities of the cancellation architecture.

5 SYSTEM SUPPORT CONSTRAINTS AND MODEL-INDEPENDENT LIMITS

Beyond information-theoretic limitations, cancellation performance can be constrained by fundamental physical or structural properties of the system, irrespective of the algorithm’s sophistication. A critical constraint arises from the frequency-domain support of the system components, particularly when the cancellation mechanism $f(y(n))$ cannot generate signals covering the entire frequency spectrum of the disturbance $d(n)$.

Let us specialize to the common case where the cancellation mechanism is an LTI system, $f(y(n)) = S(z)y(n)$, representing the secondary path in ANC or the echo path model in echo cancellation. Let $P(e^{j\omega})$ and $S(e^{j\omega})$ be the frequency responses of the primary and secondary paths, respectively. We define the frequency support of a system as the set of frequencies where its response is non-zero. Let

$$\begin{aligned} \text{supp}(P) &= \{\omega \in [-\pi, \pi] \mid P(e^{j\omega}) \neq 0\} \\ \text{supp}(S) &= \{\omega \in [-\pi, \pi] \mid S(e^{j\omega}) \neq 0\} \end{aligned}$$

If there exists a frequency band where the primary path transmits energy (i.e., $\omega \in \text{supp}(P)$) but the secondary path has no response (i.e., $\omega \notin \text{supp}(S)$), then the component of the disturbance $d(n)$ at

frequency ω cannot be cancelled by any signal $y(n)$ filtered through $S(z)$. The cancellation signal $f(y(n)) = S(z)y(n)$ will have zero energy at such frequencies.

Let $\Omega_{\text{uncancelable}} = \text{supp}(P) \setminus \text{supp}(S)$ denote the set of frequencies where the disturbance exists but the cancellation path has no gain. The portion of the disturbance power residing in this uncancelable frequency region constitutes an irreducible component of the error power. The power of $d(n)$ restricted to this frequency set is given by:

$$\sigma_{d,\text{uncancelable}}^2 = \frac{1}{2\pi} \int_{\Omega_{\text{uncancelable}}} S_{dd}(e^{j\omega}) d\omega = \frac{1}{2\pi} \int_{\Omega_{\text{uncancelable}}} |P(e^{j\omega})|^2 S_{xx}(e^{j\omega}) d\omega \quad (9)$$

Since the cancellation signal $f(y(n))$ has zero energy in $\Omega_{\text{uncancelable}}$, the residual error $e(n) = d(n) - f(y(n))$ must contain at least this amount of power. This yields a model-independent lower bound on the error power:

[Support-Based Bound on Cancellation Error] Assuming $d(n) = P(z)x(n)$ and $f(y(n)) = S(z)y(n)$, where $x(n)$ is WSS and $P(z), S(z)$ are LTI systems, the average power of the residual error $e(n)$ is lower bounded by the power of the disturbance in the frequency region where the primary path has support but the secondary path does not:

$$\sigma_e^2 = \mathbb{E}[|e(n)|^2] \geq \sigma_{d,\text{uncancelable}}^2 = \frac{1}{2\pi} \int_{\text{supp}(P) \setminus \text{supp}(S)} S_{dd}(e^{j\omega}) d\omega \quad (10)$$

This bound holds regardless of the algorithm used to generate $y(n)$, even if it possessed infinite computational power and perfect knowledge of the system, as it stems from the inherent inability of the system $S(z)$ to counteract certain frequency components of $d(n)$. The corresponding NMSE lower bound is obtained by normalizing this irreducible error power by the total disturbance power σ_d^2 :

$$\text{NMSE} = \frac{\mathbb{E}[|e(n)|^2]}{\sigma_d^2} \geq \frac{\sigma_{d,\text{uncancelable}}^2}{\sigma_d^2} = \frac{\int_{\text{supp}(P) \setminus \text{supp}(S)} S_{dd}(e^{j\omega}) d\omega}{\int_{-\pi}^{\pi} S_{dd}(e^{j\omega}) d\omega} \quad (11)$$

In decibels:

$$\text{NMSE}_{\text{dB}} \geq 10 \log_{10} \left(\frac{\int_{\text{supp}(P) \setminus \text{supp}(S)} |P(e^{j\omega})|^2 S_{xx}(e^{j\omega}) d\omega}{\int_{-\pi}^{\pi} |P(e^{j\omega})|^2 S_{xx}(e^{j\omega}) d\omega} \right) \quad (12)$$

This expression quantifies the fraction of the total disturbance power that lies in the frequency bands structurally unreachable by the cancellation path $S(z)$. It represents a fundamental performance ceiling imposed by the physical setup or inherent characteristics of the primary and secondary paths.

6 UNIFIED PERFORMANCE BOUND

Combining the insights from the information-theoretic and support-based analyses, we arrive at a unified lower bound on the achievable cancellation performance. The actual residual error power must satisfy both constraints simultaneously. Therefore, the NMSE is lower bounded by the maximum of the two individual bounds:

[Unified Lower Bound on NMSE] Under the assumptions of joint WSS for $d(n), y(n)$, finite entropy rate $H(d)$, LTI primary path $P(z)$, and LTI cancellation path $S(z)$, the Normalized Mean Squared Error (NMSE) in decibels is lower bounded by:

$$\text{NMSE}_{\text{dB}} \geq \max \left\{ 10 \log_{10} \left(1 - \frac{I(y; d)}{H(d)} \right), 10 \log_{10} \left(\frac{\int_{\text{supp}(P) \setminus \text{supp}(S)} S_{dd}(e^{j\omega}) d\omega}{\int_{-\pi}^{\pi} S_{dd}(e^{j\omega}) d\omega} \right) \right\} \quad (13)$$

This unified bound encapsulates the dual nature of performance limitations in cancellation systems. Performance can be bottlenecked either by the algorithm's inability to capture and utilize sufficient information about the disturbance (the $I(y; d)/H(d)$ term) or by the physical inability of the cancellation path to address certain frequency components of the disturbance (the spectral support mismatch term). This provides a comprehensive theoretical tool for evaluating the fundamental limits of cancellation systems, guiding algorithm design and system configuration. Understanding which bound is dominant in a particular scenario can inform whether efforts should focus on improving the information processing capabilities of the algorithm or on modifying the physical system (e.g., actuator placement, speaker characteristics) to improve spectral coverage.

7 EXPERIMENTS

Evaluation data: We report performances on the customary NOISEX dataset Varga & Steeneken (1993), which includes a collection of real-world noise recordings commonly encountered in speech and audio processing tasks, including babble, factory, and engine noises. To align with our preprocessing pipeline, all samples were standardized to 3 seconds and resampled at 16 kHz. **Simulator:** A rectangular room of dimensions [3, 4, 2] meters (width, length, height) was simulated. Room impulse responses (RIRs) were generated (Allen & Berkley, 1979) via the Python *rir_generator* package, with a 512-tap length and high-pass filtering enabled. Microphone and speaker positions were set at [1.5, 3, 1] m (error mic), [1.5, 1, 1] m (reference mic), and [1.5, 2.5, 1] m (cancellation speaker). Customary reverberation times were sampled from $\{0.15, 0.175, 0.2, 0.225, 0.25\}$ s in the boundary calculations. Although the Scaled Error Function (SEF) Tobias & Seara (2006) is commonly used to model loudspeaker nonlinearity, in this work we consider only the linear case, corresponding to $\eta^2 \rightarrow \infty$. **Hyperparameters:** In order to reproduce the performances presented in DeepASC Mishaly et al. (2025) we reproduced the training based on their hyperparameters description. **Baseline Methods:** We confronted our bound derived above with the state-of-the-art deep learning-based ANC methods named DeepANC Zhang & Wang (2021), Attentive Recurent Network (ARN) Zhang & Wang (2023) and, DeepASC Mishaly et al. (2025). To evaluate the effectiveness of each method under varying noise conditions, we employed pretrained models as a standardized benchmark. This approach ensured consistent performance assessment across different noise types. **Ressources:** All experiments were conducted using NVIDIA RTX 3090 GPUs.

8 RESULTS AND DISCUSSIONS

8.1 BOUNDARIES HOLD NMSE PERFORMANCES OVER VARIOUS REVERBERATION TIMES

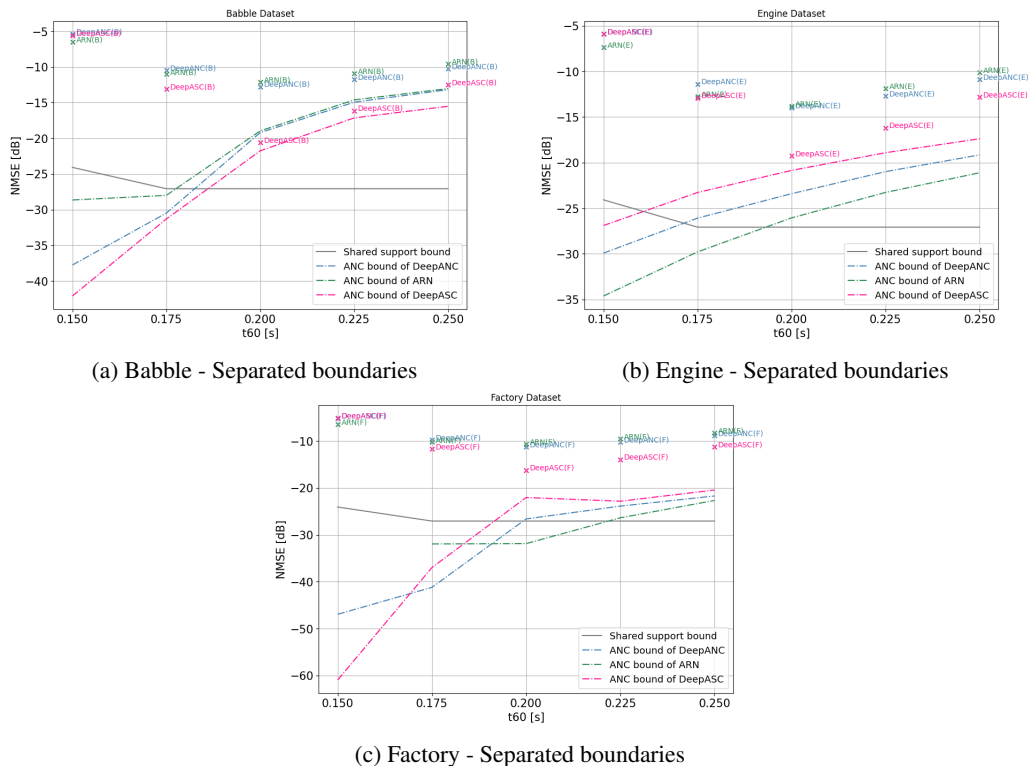


Figure 2: Model performances for various reverberation times with associated boundaries

Figure 2 presents the performances of three baseline methods on three noise types from the NoiseX dataset - Babble, Engine, and Factory noises - and the corresponding boundaries derived in the previous sections for various reverberation times sampled from $\{0.15, 0.175, 0.2, 0.225, 0.25\}$. We

show the information-theoretic bound separated from the bound based on support constraints from the ANC method in order to highlight the impact of each term of the unified boundary Equation 13.

Separated Boundaries Firstly, we notice that the bound based on support constraints (Eq. 12) is model and dataset independent. Indeed, the subfigures from Figure 2 share the same support-based boundary (in gray). This bound holds for all the reported performances obtained from the three different methods. We also observe that the Information-Theoretic bound (ANC bound - Eq. 8) is increasing with reverberation times. As reverberation time increases, the signal becomes more complex since the physical translation is the creation of more bounces on the simulator’s virtual walls. This leads to longer, more intricate room impulse responses, making both the primary and secondary paths temporally extended and harder to model accurately. This complexity is further amplified by the dense overlap of reflected waves, which blurs the distinction between direct and reflected components, making path estimation more challenging. Consequently, this growing complexity contributes to the steady increase in the ANC information-theoretic boundaries.

Audio Noise Cancellation Performances Figure 2 shows the performances of three models (DeepASC, DeepANC and ARN) measured using NMSE over three different noise types. Contrary to what is often observed, the performances are not only calculated and reported for one reverberation time (usually $t_{60} = 0.2s$ Zhang & Wang (2021; 2023); Mishaly et al. (2025)) but for all the reverberation times used in the training process ($t_{60} \in \{0.15, 0.175, 0.2, 0.225, 0.25\}$). In Figure 2 each subfigure corresponds to a different type of noise (respectively Babble, Engine, and Factory noises). From left to right, we observe lower NMSE performances than the previous one. Based on these observations, Babble noise appears to be the easiest to cancel, followed by Engine noise, while Factory noise proves to be the most challenging. Figure 3 provides visual confirmation of these observations by displaying the spectrograms of the noisy signals. The Engine noise exhibits temporally repetitive frequency structures, which enhance its predictability and facilitate its attenuation using adaptive filtering methods. In contrast, the Factory noise shows a more broadband and statistically uniform distribution, lacking distinct temporal patterns, which increases its resistance to cancellation. Furthermore, Engine contains more high-frequency energy than Babble, which may reduce the effectiveness of noise suppression techniques due to the higher spectral variability and lower correlation with past signal frames.

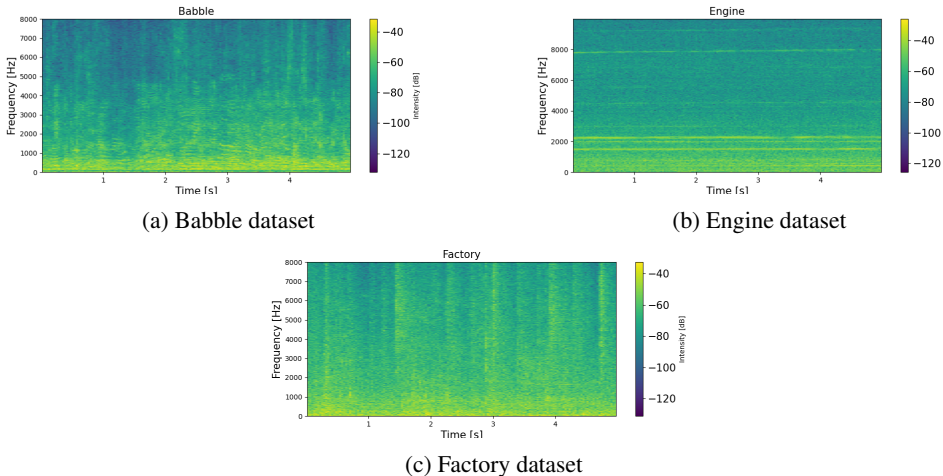


Figure 3: Time-Frequency analysis of the different noise

Unified Boundaries By applying the max operator between the support bound and the ANC bound in Figure 2 the unified boundaries are obtained . It demonstrates that the unified bound derived from Equation 13 is consistently satisfied across the three methods and types of noise. According to previous analysis, we observe that the unified bound necessitates the two terms because the support bound leads for low reverberation times and the Information-Theoretic bound for high reverberation times. Another observation is the increasing of the gap between the performances and the boundaries as noise complexity increases (Figure 2). As previously discussed and illustrated in Figure 3, these observations indicate an increasing level of cancellation difficulty: Babble, followed by Engine,

432 and finally Factory, which presents the greatest challenge—reflected in a larger gap between actual
433 performance and theoretical boundaries.

434 Therefore the gap between the performance and the boundary changes through noise types. Also,
435 Figure 4 in Appendix A.3 is a more general observation of the performances that can be used in
436 various fields. Indeed, performance values and deviations from a theoretical boundary can lead to
437 a better understanding of data or algorithms. This work proposes theoretically derived performance
438 boundaries that offer a principled benchmark for analyzing and comparing active noise control al-
439 gorithms. By unifying information-theoretic and support-based criteria, the framework provides
440 insight into the fundamental limitations of system performance under varying noise conditions, and
441 can inform the design of more effective adaptive filtering strategies in the future.

442 9 LIMITATIONS AND BROADER IMPACT

443 **Limitations:** While deriving a lower bound for various ANC methods provides valuable insight into
444 their capabilities, such bound estimations come with inherent limitations. In particular, computing
445 the Information-Theoretic bound necessitates a large number of samples to accurately estimate the
446 underlying probability density functions of the signals. Insufficient sampling leads to poor approx-
447 imation of the true distributions, which in turn distorts the bound calculations, making them unreli-
448 able or overly optimistic. Consequently, to ensure meaningful and accurate Information-Theoretic
449 evaluations in ANC scenarios, a substantial volume of diverse audio data is essential. **Broader**
450 **Impact:** This work helps identify performance limitations and guides algorithm design. Improved
451 ANC technology can positively impact society by enhancing comfort and health in noisy environ-
452 ments, with applications in consumer electronics, healthcare, and public infrastructure. While gen-
453 erally positive, potential concerns include over-reliance on ANC instead of addressing noise sources
454 directly, and the need for careful integration to ensure fairness and transparency.

455 10 CONCLUSIONS

456 We presented a unified bound that captures the two fundamental limitations in cancellation sys-
457 tems: the algorithm’s capacity to extract information about the disturbance ($I(y; d)/H(d)$) and the
458 physical system’s ability to counter specific frequency components. It serves as a theoretical tool
459 for assessing ANC limits and guiding design decisions—helping determine whether improvements
460 should target algorithmic information processing or physical system setup. The unified bound was
461 evaluated across various noise types using baseline methods and consistently provided a lower bound
462 on their observed performance. Analysis of the bound provides valuable insights into the charac-
463 teristics of different noise types. Specifically, we observe that the bound increases with the complex-
464 ity of the noise, indicating greater theoretical difficulty in achieving effective cancellation. Addition-
465 ally, the widening gap between the model performance and the bound for more challenging noises
466 suggests limitations in the models’ ability to fully exploit the available information, highlighting
467 areas for potential algorithmic improvement.

468 11 REPRODUCIBILITY STATEMENT

469 In this paper we introduce a new boundary that aims to help all stakeholders in the field improve their
470 understanding of ANC algorithms. This understanding could lead to more efficient algorithms. To
471 do so, We have made efforts to make our boundary reproducible and easily applicable to algorithms
472 developed by future readers. We provide the code used to calculate the various terms of the boundary
473 and an explanation of the methods used during implementation in the appendix.

474 12 ETHICAL STATEMENT

475 This work helps identify performance limitations and guides algorithm design. Improved ANC tech-
476 nology can positively impact society by enhancing comfort and health in noisy environments, with
477 applications in consumer electronics, healthcare, and public infrastructure. While generally positive,
478 potential concerns include over-reliance on ANC instead of addressing noise sources directly, and
479 the need for careful integration to ensure fairness and transparency.

REFERENCES

- 486
487
488 Jont B Allen and David A Berkley. Image method for efficiently simulating small-room acoustics.
489 *The Journal of the Acoustical Society of America*, 65(4):943–950, 1979.
- 490
491 Christian Antoñanzas, Miguel Ferrer, Maria De Diego, and Alberto Gonzalez. Remote microphone
492 technique for active noise control over distributed networks. *IEEE/ACM Transactions on Audio,
493 Speech, and Language Processing*, 31:1522–1535, 2023.
- 494
495 CC Boucher, SJ Elliott, and PA Nelson. Effect of errors in the plant model on the performance of al-
496 gorithms for adaptive feedforward control. In *IEE Proceedings F (Radar and Signal Processing)*,
497 volume 138, pp. 313–319. IET, 1991.
- 498
499 John C Burgess. Active adaptive sound control in a duct: A computer simulation. *The Journal of
500 the Acoustical Society of America*, 70(3):715–726, 1981.
- 501
502 Young-Jin Cha, Alireza Mostafavi, and Sukhpreet S Benipal. Dnoisenet: Deep learning-based feed-
503 back active noise control in various noisy environments. *Engineering Applications of Artificial
504 Intelligence*, 121:105971, 2023.
- 505
506 Debi Prasad Das and Ganapati Panda. Active mitigation of nonlinear noise processes using a novel
507 filtered-s lms algorithm. *IEEE Transactions on Speech and Audio Processing*, 12(3):313–322,
508 2004.
- 509
510 Christopher C Fuller, Sharon Elliott, and Philip Arthur Nelson. *Active control of vibration*. Aca-
511 demic press, 1996.
- 512
513 Sepehr Ghasemi, Raja Kamil, and Mohammad Hamiruce Marhaban. Nonlinear thf-fxlms algorithm
514 for active noise control with loudspeaker nonlinearity. *Asian Journal of Control*, 18(2):502–513,
515 2016.
- 516
517 Colin H Hansen, Scott D Snyder, Xiaojun Qiu, Laura A Brooks, and Danielle J Moreau. *Active
518 control of noise and vibration*. E & Fn Spon London, 1997.
- 519
520 Sen M Kuo and Dennis R Morgan. Active noise control: a tutorial review. *Proceedings of the IEEE*,
521 87(6):943–973, 1999.
- 522
523 Sen M Kuo and Hsien-Tsai Wu. Nonlinear adaptive bilinear filters for active noise control systems.
524 *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(3):617–624, 2005.
- 525
526 Paul Lueg. Process of silencing sound oscillations. *US patent 2043416*, 1936.
- 527
528 Zhengding Luo, Dongyuan Shi, and Woon-Seng Gan. A hybrid sfanc-fxn lms algorithm for active
529 noise control based on deep learning. *IEEE Signal Processing Letters*, 29:1102–1106, 2022.
- 530
531 Zhengding Luo, Dongyuan Shi, Woon-Seng Gan, and Qirui Huang. Delayless generative fixed-filter
532 active noise control based on deep learning and bayesian filter. *IEEE/ACM Transactions on Audio,
533 Speech, and Language Processing*, 2023a.
- 534
535 Zhengding Luo, Dongyuan Shi, Xiaoyi Shen, Junwei Ji, and Woon-Seng Gan. Deep generative
536 fixed-filter active noise control. In *ICASSP*, pp. 1–5. IEEE, 2023b.
- 537
538 Zhengding Luo, Dongyuan Shi, Xiaoyi Shen, Junwei Ji, and Woon-Seng Gan. Gfanc-kalman: Gen-
539 erative fixed-filter active noise control with cnn-kalman filtering. *IEEE Signal Processing Letters*,
2023c.
- 536
537 Zhengding Luo, Dongyuan Shi, Xiaoyi Shen, and Woon-Seng Gan. Unsupervised learning based
538 end-to-end delayless generative fixed-filter active noise control. In *ICASSP*, pp. 441–445. IEEE,
539 2024.
- 536
537 Yehuda Mishaly, Lior Wolf, and Eliya Nachmani. Deep active speech cancellation with multi-band
538 mamba network, 2025. URL <https://arxiv.org/abs/2502.01185>.
- 538
539 Alireza Mostafavi and Young-Jin Cha. Deep learning-based active noise control on construction
sites. *Automation in Construction*, 151:104885, 2023.

- 540 Philip Arthur Nelson and Stephen J Elliott. *Active control of sound*. Academic press, 1991.
- 541
- 542 Ashutosh Pandey and DeLiang Wang. Self-attending rnn for speech enhancement to improve cross-
543 corpus generalization. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30:
544 1374–1385, 2022.
- 545 JungPhil Park, Jeong-Hwan Choi, Yungyeo Kim, and Joon-Hyuk Chang. Had-anc: A hybrid sys-
546 tem comprising an adaptive filter and deep neural networks for active noise control. In *INTER-*
547 *SPEECH*, volume 2023, pp. 2513–2517. International Speech Communication Association, 2023.
- 548
- 549 Seunghyun Park and Daejin Park. Integrated 3d active noise cancellation simulation and synthesis
550 platform using tcl. In *2023 IEEE 16th International Symposium on Embedded MCoS*, pp. 111–
551 116. IEEE, 2023.
- 552 Jagdish Chandra Patra, Ranendra N Pal, BN Chatterji, and Ganapati Panda. Identification of non-
553 linear dynamic systems using functional link artificial neural networks. *IEEE transactions on*
554 *systems, man, and cybernetics, part b (cybernetics)*, 29(2):254–262, 1999.
- 555
- 556 Alexander Pike and Jordan Cheer. Generalized performance of neural network controllers for feed-
557 forward active control of nonlinear systems. 2023.
- 558
- 559 Xing Ren and Hongwei Zhang. An improved artificial bee colony algorithm for model-free ac-
560 tive noise control: algorithm and implementation. *IEEE Transactions on Instrumentation and*
Measurement, 71:1–11, 2022.
- 561
- 562 Dongyuan Shi, Woon-Seng Gan, Bhan Lam, and Shulin Wen. Feedforward selective fixed-filter
563 active noise control: Algorithm and implementation. *IEEE/ACM Transactions on Audio, Speech,*
564 *and Language Processing*, 28:1479–1492, 2020.
- 565
- 566 Dongyuan Shi, Bhan Lam, Kenneth Ooi, Xiaoyi Shen, and Woon-Seng Gan. Selective fixed-filter
567 active noise control based on convolutional neural network. *Signal Processing*, 190:108317, 2022.
- 568
- 569 Dongyuan Shi, Woon-Seng Gan, Bhan Lam, Zhengding Luo, and Xiaoyi Shen. Transferable latent
570 of cnn-based selective fixed-filter active noise control. *IEEE/ACM Transactions on Audio, Speech,*
and Language Processing, 31:2910–2921, 2023a.
- 571
- 572 Dongyuan Shi, Bhan Lam, Xiaoyi Shen, and Woon-Seng Gan. Multichannel two-gradient direction
573 filtered reference least mean square algorithm for output-constrained multichannel active noise
control. *Signal Processing*, 207:108938, 2023b.
- 574
- 575 Dongyuan Shi, Woon-seng Gan, Xiaoyi Shen, Zhengding Luo, and Junwei Ji. What is behind the
576 meta-learning initialization of adaptive filter?—a naive method for accelerating convergence of
577 adaptive multichannel active noise control. *Neural Networks*, 172:106145, 2024.
- 578
- 579 Deepali Singh, Rinki Gupta, Arun Kumar, and Rajendar Bahl. Enhancing active noise control
580 through stacked autoencoders: Training strategies, comparative analysis, and evaluation with
practical setup. *Engineering Applications of Artificial Intelligence*, 135:108811, 2024.
- 581
- 582 Li Tan and Jean Jiang. Adaptive volterra filters for active control of nonlinear noise processes. *IEEE*
Transactions on signal processing, 49(8):1667–1676, 2001.
- 583
- 584 Orlando José Tobias and Rui Seara. Leaky-fxlms algorithm: Stochastic analysis for gaussian data
585 and secondary path modeling error. *IEEE Transactions on speech and audio processing*, 13(6):
586 1217–1230, 2005.
- 587
- 588 Orlando José Tobias and Rui Seara. On the lms algorithm with constant and variable leakage factor
in a nonlinear environment. *IEEE transactions on signal processing*, 54(9):3448–3458, 2006.
- 589
- 590 Andrew Varga and Herman JM Steeneken. Assessment for automatic speech recognition: Ii. noise-
591 92: A database and an experiment to study the effect of additive noise on speech recognition
592 systems. *Speech communication*, 12(3):247–251, 1993.
- 593
- Tong Xiao, Buye Xu, and Chuming Zhao. Spatially selective active noise control systems. *The*
Journal of the Acoustical Society of America, 153(5):2733–2733, 2023.

594 Hao Zhang and DeLiang Wang. Deep anc: A deep learning approach to active noise control. *Neural*
595 *Networks*, 141:1–10, 2021.

597 Hao Zhang and DeLiang Wang. Deep mcanc: A deep learning approach to multi-channel active
598 noise control. *Neural Networks*, 158:318–327, 2023.

600 Hao Zhang, Ashutosh Pandey, and DeLiang Wang. Attentive recurrent network for low-latency
601 active noise control. In *INTERSPEECH*, pp. 956–960, 2022.

602 Hao Zhang, Ashutosh Pandey, and DeLiang Wang. Low-latency active noise control using attentive
603 recurrent network. *IEEE/ACM transactions on audio, speech, and language processing*, 31:1114–
604 1123, 2023a.

606 Huawei Zhang, Jihui Zhang, Fei Ma, Prasanga N Samarasinghe, and Huiyuan Sun. A time-domain
607 multi-channel directional active noise control system. In *2023 31st EUSIPCO*, pp. 376–380.
608 IEEE, 2023b.

610 Yang Zhou, Haiquan Zhao, and Dongxu Liu. Genetic algorithm-based adaptive active noise control
611 without secondary path identification. *IEEE Transactions on Instrumentation and Measurement*,
612 2023.

614 Wenzhao Zhu, Bo Xu, Zong Meng, and Lei Luo. A new dropout leaky control strategy for multi-
615 channel narrowband active noise cancellation in irregular reverberation room. In *2021 7th Inter-*
616 *national Conference on Computer and Communications (ICCC)*, pp. 1773–1777. IEEE, 2021.

618 A APPENDIX

619 A.1 NUMERICAL CALCULATION OF THE INFORMATION-THEORETIC BOUNDARY (EQ. 8)

620 The numerical estimation of Equation 8 requires the estimation of the marginal and the joint en-
621 tropies of the signals.

622 To do so, we calculated the marginal Probability Density Functions (PDFs) associated to the primary
623 signal $d(n)$ and the output of the ANC algorithm $y(n)$. To estimate the marginal and joint PDFs of
624 two time series variables, we developed a kernel-based approach that aggregates density estimates
625 across paired sequences. For each pair of subsequences in the datasets, we first apply Gaussian
626 Kernel Density Estimation (KDE) to compute the marginal PDFs and joint PDF over fixed evaluation
627 grids. Bandwidth is set inversely proportional to the number of histogram bins (`bin_count`) to control
628 the smoothness of the estimates. The joint and marginal PDFs are normalized using numerical
629 integration to ensure valid probability distributions. These estimates are then averaged across all
630 subsequence pairs to yield the final marginal and joint PDFs.

631 We compute an information-theoretic bound by estimating the mutual information (MI) between
632 predicted and observed sequences using entropy terms derived from their joint and marginal prob-
633 ability distributions. Entropies are calculated via normalized histograms obtained from KDE-based
634 PDF estimates. The MI is then used to derive a coefficient that scales the energy of the direct path
635 in the room impulse response, yielding an upper-bound estimate and its decibel-scaled version.

639 A.2 NUMERICAL CALCULATION OF THE SYSTEM SUPPORT-BASED BOUNDARY (EQUATION 640 12)

641 To assess the spectral separation between primary and secondary acoustic paths, we compute a
642 support-based boundary metric in the frequency domain. The method performs a Fast Fourier Trans-
643 form (FFT) on the impulse responses of both paths and identifies the spectral support of the primary
644 path as the set of frequency bins that exceed a predefined magnitude threshold (e.g. 45 dB). A
645 boundary score is then calculated as the proportion of these primary-support bins that do not overlap
646 with the secondary path’s support (defined as bins below the threshold). This ratio quantifies the
647 distinctness of the primary path and is reported both linearly and in decibels.

A.3 BOUNDARY OVERVIEW FOR A FIXED REVERBERATION TIME ($t_{60} = 0.2s$)

In this section, we provide a consolidated overview of the boundaries across different noise types to follow standard machine learning conventions and offer a comprehensive view of the results. Indeed, Figure 4 shows the performances and the discussed boundaries but this time for $t_{60} = 0.2s$. As previously mentioned, the unified boundary—defined as the maximum of the Information-Theoretic and Support-based boundaries—encloses the performance achieved by the different methods across all noise types. Figure 4 highlights the decline in performance and the increasing gap between performances and theoretical boundaries as the noises become more complex (from left to right along the x-axis in Figure 4). A key point is that, for a given algorithm, the lower bound depends on the noise characteristics as it changes the equivalent channel capacity.

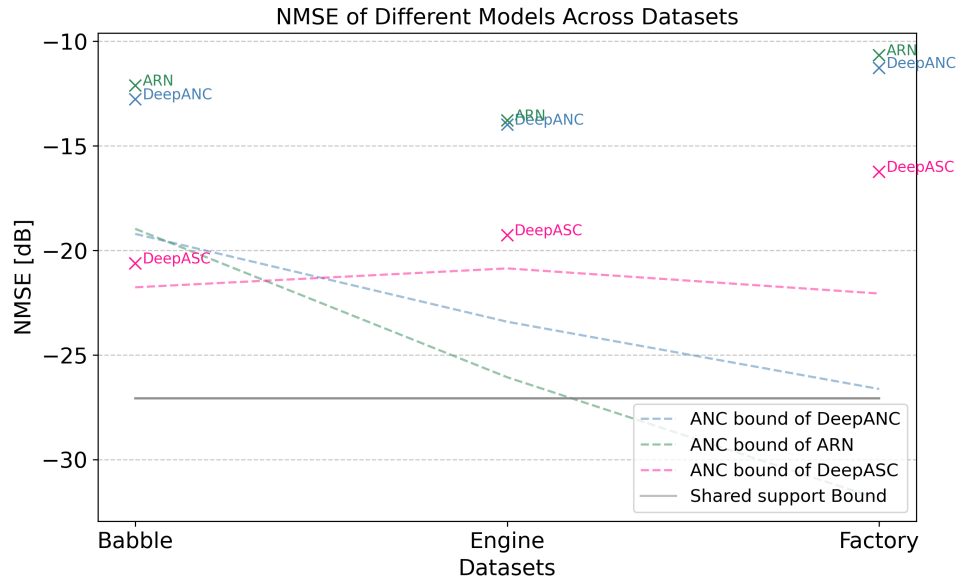


Figure 4: The unified lower bound and NMSE comparison over different noise datasets for $t_{60} = 0.2s$

B USAGE OF LARGE LANGUAGE MODELS

In this paper Large Language Models (LLM) were used to perform grammatical checks and assist with LaTeX layout (creation and editing of tables and figures).