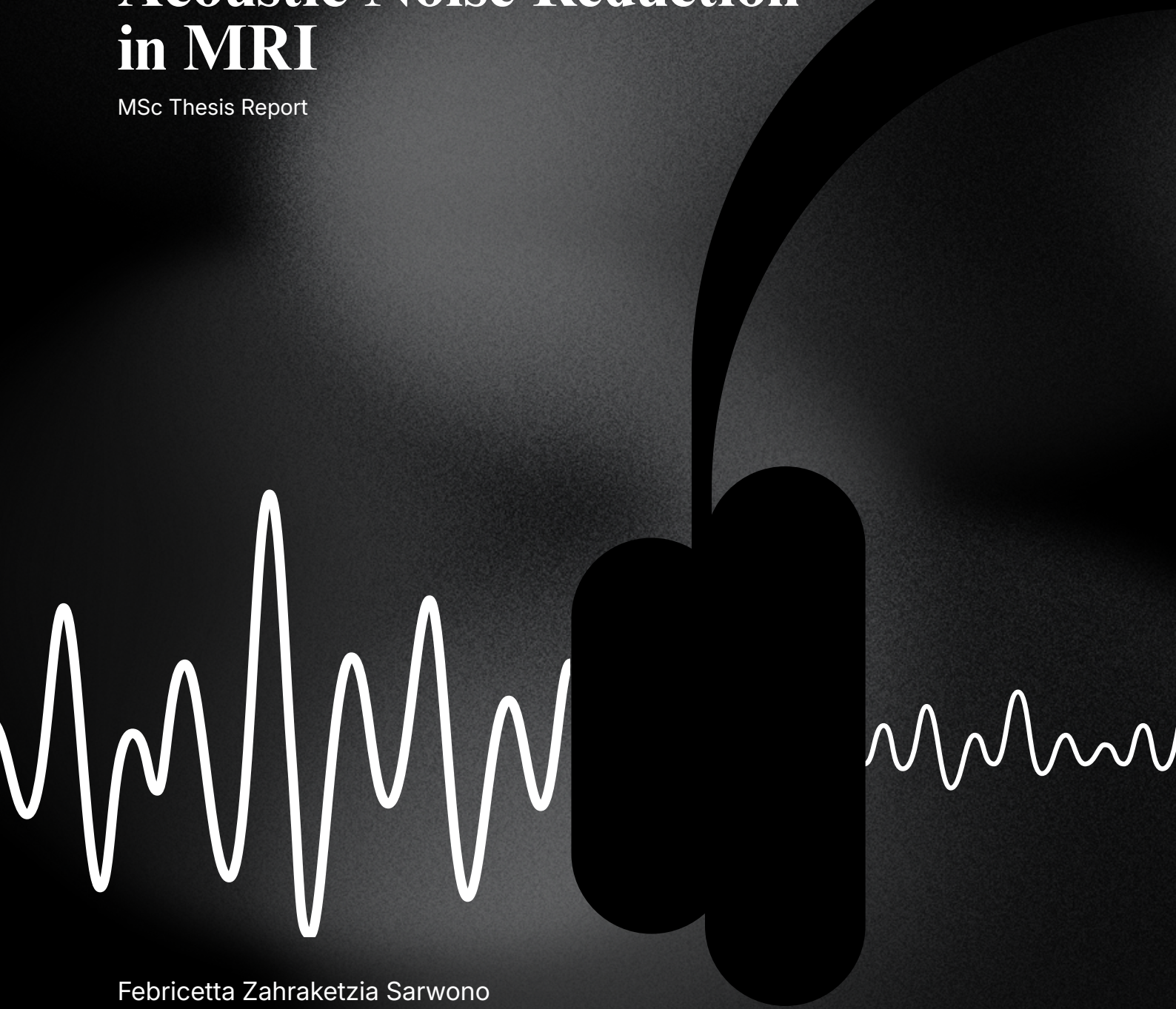
Redevelopment, System Refinement, and Evaluation of Predictive Noise Cancellation for Acoustic Noise Reduction in MRI

MSc Thesis Report



Febricetta Zahraketzia Sarwono
6199852

TU Delft

Redevelopment, System Refinement, and Evaluation of Predictive Noise Cancellation for Acoustic Noise Reduction in MRI

Master Thesis Report

by

Febricetta Zahraketzia Sarwono

Student Number
6199852

in partial fulfilment of the requirements of
Master of Science in Biomedical Engineering (track Medical Physics)
at the Delft University of Technology,
to be defended publicly on Thursday, July 16, 2026, at 09:00.

Project duration:	November 19, 2025 – July 16, 2026
Responsible Supervisor:	Sebastian Weingärtner (TU Delft)
Assessment Chair:	Qian Tao (TU Delft)
Assessment Committee:	Sebastian Weingärtner & Jieun Yang (TU Delft)
Expert Committee:	Kadir Berat Yildirim (TU Delft)
Institutions:	Delft University of Technology
Faculty:	Faculty of Mechanical Engineering, TU Delft

TU Delft

Abstract

Magnetic resonance imaging (MRI) produces high acoustic noise levels that can reduce patient comfort, make the scanning experience intimidating, raise hearing safety concerns, and interfere with auditory functional MRI experiments. Predictive noise cancellation (PNC) aims to reduce this noise by estimating the MRI gradient-to-acoustic transfer function and generating sequence-specific anti-noise. This thesis redeveloped an existing LabVIEW/MATLAB-based PNC pipeline into a Python-based framework and evaluated redesigned calibration strategies for phantom-based MRI acoustic noise reduction.

The Python framework rebuilt the main PNC workflow interface, digital signal processing, calibration, regular sequence processing, audio recording, and arbitrary function generator control. Validation against the original implementation showed that the rebuilt system preserved the main DSP and workflow behavior, with full-workflow errors below 5% for the evaluated blocks except for one flagged regular sequence alignment case. The Python implementation also reduced computation time by approximately 73% and improved workflow usability and stability.

Calibration redesign was evaluated using chirp-based gradients, transfer function stability metrics, live calibration-stage reduction, and phantom measurements using fast field echo (FFE) and echo planar imaging (EPI) sequences. Compared with the original sl014 calibration, the chirp-based gradients provided stronger excitation and generally improved transfer function stability and live reduction. Chirp 1 and chirp 4 were the most reliable broadband calibration candidates across calibration and FFE measurements. For EPI-focused testing, combined chirp+EPI-Y transfer functions improved EPI-band cancellation relative to chirp-only calibration and reduced residual amplitudes near the dominant EPI component.

The achieved cancellation remained limited by channel imbalance, frequency-dependent reduction, low-frequency playback constraints, and acoustic behavior that was not fully described by a fixed linear time-invariant model, especially for the more complex EPI sequence. In vivo evaluation was also not performed, so the results remain limited to phantom-based testing. Overall, this thesis validated a faster and more usable Python-based PNC framework and showed that redesigned broadband and EPI-targeted calibrations can improve sequence-specific MRI acoustic noise reduction.

Keywords: MRI acoustic noise reduction, predictive noise cancellation, acoustic transfer function, calibration gradient design, pipeline redevelopment, phantom validation, fast field echo, echo planar imaging

Preface

This thesis marks the final part of my master's journey in Biomedical Engineering at TU Delft. Looking back, there are many things I wish I had done differently, especially learning to enjoy the thesis process more fully. Even so, I am grateful that this journey led me to MRI. My interest in imaging and acoustics began long before this project, so working on predictive acoustic noise cancellation for MRI felt like an unexpected but fitting combination of both interests. I hope this work can contribute, even in a small way, to improving the MRI experience for patients, especially because the scanning process can be loud, uncomfortable, and intimidating.

I would like to express my sincere gratitude to Sebastian Weingärtner, my responsible supervisor, who excitedly welcomed me as a master's student who, at that time, basically knew nothing about MRI other than the theories he practically taught me inside the classroom. Thanks to his genuine love for research, I felt motivated to express the same effort. I would also like to thank my daily supervisor, Kadir Berat Yildirim. I know that I would not have been able to finish this thesis without him believing in me since the first day, supporting me with technical things that I was not familiar with, and always being supportive.

I am also grateful to Holland Proton Therapy Center, especially Christal van de Steeg-Henzen, for her support and assistance during the experiments. Her help made the MRI scanning sessions possible. I would also like to thank the members of the Mars Lab for their supervision, assistance, and support whenever I needed it. Their contributions were essential to the successful completion of the phantom measurements.

I gratefully acknowledge the financial support provided by Lembaga Pengelola Dana Pendidikan (LPDP), also known as the Indonesia Endowment Fund for Education, for funding my master's studies. This support gave me the opportunity to continue my education at Delft University of Technology and to complete the research presented in this thesis.

Finally, I would like to express my deepest gratitude to my family, especially my father and mother, who have supported me from the very beginning, believed in me, and kept me in their prayers. I am also grateful to my friends in Delft: Agassi, Vadya, Ovi, Kak Grase, Thoriq, Ammar, and Rizqi, as well as to my friends outside Delft who listened to me through both the difficult and happy moments of my thesis and life in Delft: Fanny, Rizqa, Aliya, Axel, and Ai. I would not have made it through this journey without all of them.

Febricetta Zahraketzia Sarwono
Delft, July 2026

Contents

Nomenclature

Abbreviations

AC	Auditory Cortex
AFG	Arbitrary Function Generator
ANC	Active Noise Cancellation
BOLD	Blood Oxygenation Level-Dependent
EPI	Echo Planar Imaging
FFE	Fast Field Echo
fMRI	Functional MRI
FOV	Field of View
I/O	Input/Output
LTI	Linear Time-Invariant
MRI	Magnetic Resonance Imaging
PNC	Predictive Noise Cancellation
RMSE	Root-Mean-Square Error
SER	Signal-to-Error Ratio
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
TE	Echo Time
TF	Transfer Function
TR	Repetition Time
TTL	Transistor-Transistor Logic

1 Introduction

Magnetic resonance imaging (MRI) is a widely used diagnostic imaging modality in modern healthcare. Unlike several other medical imaging techniques, MRI does not use ionizing radiation and is generally considered safer. However, despite this safety advantage, MRI examinations are always accompanied by high levels of acoustic noise, which arise from Lorentz-force-induced vibrations caused by rapid gradient switching. Standard clinical MRI sequences typically produce sound pressure levels (SPLs) in the range of 80–115 dBA [6, 20], while more intense sequences such as echo planar imaging (EPI) can reach 120–130 dB [27, 28]. These high acoustic noise levels can directly contribute to patient discomfort, increase anxiety, and reduce examination tolerability. As a result, some patients report a poorer overall scanning experience than with imaging modalities that pose greater physical risks, such as computed tomography or X-ray [9]. Adding to this concern, these SPLs already exceed public health recommendations [35] and may increase the risk of temporary or permanent hearing threshold shifts [5, 19, 30, 36]. According to the International Electrotechnical Commission, reducing sound level exposure at the eardrum to below 85 dBA is essential. The recommended way to achieve that standard is by utilizing hearing protection when the equipment is capable of producing sound levels above 99 dBA [35]. These considerations show that acoustic noise should not be taken lightly in MRI.

Many strategies have been explored to reduce MRI scanner noise. Each approach offers different advantages and limitations in terms of effectiveness, feasibility, cost, and compatibility with existing scanner hardware. One emerging method is predictive noise cancellation (PNC), which was recently developed within the Mars Lab at TU Delft as an alternative noise control approach. It exploits the deterministic relationship between MRI gradient waveforms and the resulting acoustic noise [39]. This feature distinguishes PNC from conventional real-time active noise cancellation (ANC), as it can avoid some latency-related limitations and be implemented without modifying scanner hardware.

MRI acoustic noise is particularly relevant in auditory functional MRI (fMRI), where the most commonly used sequence for fMRI is EPI. In auditory fMRI, scanner noise directly competes with the presented auditory stimuli, raising the acoustic baseline and potentially confounding interpretation of blood oxygenation level-dependent (BOLD) responses in auditory cortex (AC). Consequently, auditory fMRI studies often require specific design strategies to reduce or avoid scanner noise interference [8].

PNC has shown promising acoustic noise reduction in prior work. This has been demonstrated in phantom experiments, where preliminary evaluations reported acoustic noise reduction of up to 13 dB [39]. However, the original implementation relied on a multi-platform LabVIEW/MATLAB workflow, which limited flexibility, observability, and practical usability during experiments. In addition, the original calibration strategy had not been systematically optimized for different sequence types, and the performance of PNC for more complex sequences remained limited. These limitations motivated the redevelopment of the PNC system into a Python-based framework and the evaluation of optimized calibration strategies in phantom-based MRI measurements.

1.1. Background and Scientific Context

This section provides the background for the present study. It first reviews established approaches for reducing acoustic noise in MRI, followed by the modeling framework used for MRI noise prediction,

and the current state of PNC. It then introduces the relevance of PNC and more complex sequences, such as EPI.

1.1.1. Noise Cancellation in MRI

The acoustic noise generated during MRI examinations primarily originates from the gradient system. Rapidly alternating electrical currents within the gradient coils interact with the scanner's strong magnetic field, producing Lorentz forces that excite mechanical vibrations in the gradient coils and surrounding structures. These vibrations are then radiated into the bore as sound, which is known as MRI acoustic noise. The specific characteristics of this sound depend strongly on the selected MRI sequence protocol, since changes in gradient input directly alter the resulting acoustic profile. Different combinations of parameters, such as slice thickness, field of view (FOV), repetition time (TR), and echo time (TE), can therefore produce different acoustic noise characteristics [26]. As described previously, MRI-generated acoustic noise may reach levels as high as 130 dB, placing it in a very high-noise range, especially during extended examinations [27, 28]. Therefore, minimizing MRI acoustic noise is important for ensuring patient safety and making the imaging experience tolerable.

In the literature, MRI noise reduction approaches can broadly be grouped into passive attenuation, hardware- and sequence-based reduction at the source, and ANC. These methods differ in where they act, how easily they can be implemented, and what compromises they introduce. Passive protection is simple and widely used, while hardware and sequence modifications are more complex as they aim to reduce sound generation at the source. ANC, in contrast, attempts to suppress the acoustic field near the listener's ears by generating anti-noise in real time.

Table 1.1: Representative MRI acoustic noise reduction approaches and reported reductions from the literature

Category	Reported reduction	Main trade-off & limitation
Passive attenuation [10, 31]	Earmuffs: ∼30–37 dB; earplugs: ∼25–28 dB; combined: ∼39–41 dB; helmet+earmuffs+earplugs: ∼55–63 dB (dominant 1–1.4 kHz band)	Residual sound remains due to bone and tissue conduction; bulky head enclosure required for highest attenuation, which is harder for smaller head coils in ultra-high field MRI
Hardware & sequence modification [18, 25, 34]	Quiet-sequence example: 27 dBA peak reduction for quiet MPRAGE; active vibration-control example: ∼3–8 dB vibration reduction for a representative EPI sequence	May require specialized hardware, structural modification, altered sequence timing or contrast, and substantial system integration
Conventional ANC [23, 27]	∼10–35 dB broadband/frequency-band reductions, with peak reductions up to ∼55–60 dB at dominant components	Frequency dependent; constrained by MRI compatibility, bandwidth, latency, and transducer placement
PNC [39]	Mean reduction ∼9–10 dB; up to ∼13 dB; strongest suppression around 0.6–1.2 kHz	Requires calibration, synchronization, and accurate acoustic modeling; setup- and position-dependent

As summarized in Table 1.1, existing approaches require clear trade-offs between acoustic performance and implementation complexity. While passive attenuation is straightforward to implement, bone conduction pathways limit its total effectiveness [31]. Source-based modifications reduce overall sound generation but often alter scanning conditions, while conventional ANC remains constrained by real-time latency and strict MRI compatibility requirements.

Furthermore, conventional electromagnetic audio devices are fundamentally unsuitable for the scanner environment. This limitation necessitates specialized, MR-compatible pneumatic, piezoelectric, or optical transmission systems [23, 24, 39]. These combined physical and algorithmic limitations have motivated interest in predictive approaches, such as PNC, which are discussed in the following subsections.

1.1.2. LTI Model for MRI Noise Prediction

Scanner acoustics can be modeled as a linear time-invariant (LTI) system under calibrated conditions [15, 28, 32, 39, 41]. In this framework, the acoustic response measured at a fixed location is characterized as the output of a transfer function (TF) driven by gradient inputs. Standard implementations

utilize axis-specific TFs, assuming that each physical gradient axis contributes independently to the total acoustic field. This approximation directly maps the deterministic, pre-calculated gradient waveforms to their resulting acoustic output, which makes the noise predictable without requiring real-time anti-noise generation.

For multi-axis prediction, the total acoustic pressure spectrum is expressed as the linear superposition of the contributions from the three physical gradient axes,

$$P(f) = G_x(f)H_x(f) + G_y(f)H_y(f) + G_z(f)H_z(f), \tag{1.1}$$

where $P(f)$ is the predicted pressure spectrum at the target location, $H_i(f)$ represents the acoustic TF for axis i, and $G_i(f)$ denotes the spectrum of the input gradient waveform [17]. For single-axis calibration, the TFs are typically derived using the direct frequency-domain ratio,

$$H_i(f) = \frac{P_i(f)}{G_i(f)}, \tag{1.2}$$

where $P_i(f)$ is measured acoustically while driving axis i with a known calibration sequence.

In practice, accurate multi-axis prediction requires that the LTI model account for the exact physical gradient axes (x, y, and z). The linear superposition assumption holds robustly only when the relative timing, phase relationships, and cross-axis coupling effects are accurately synchronized and calibrated [15, 28, 32, 39, 41]. This LTI framework serves as the algorithmic foundation for the PNC system.

1.1.3. State-of-the-Art in Predictive Noise Cancellation for MRI

The existing PNC implementation provides the technical foundation for the present thesis. It predicts scanner sound in advance from a pre-calibrated acoustic model and plays the corresponding anti-noise in synchrony with the sequence using MRI transistor-transistor logic (TTL) triggering [39].

The existing PNC system was demonstrated on a Philips Ingenia 3T scanner at Holland Proton Therapy Center, Delft, The Netherlands, using a pneumatic headphone-imitating playback chain. In this setup, anti-noise was produced by a speaker outside the scanner room and transmitted into the bore through a hose-and-funnel arrangement, while an arbitrary function generator (AFG) was used to output synchronized anti-noise. Noise prediction was based on the LTI framework described in the previous section, where known gradient derivatives are used to estimate the scanner noise response. In simplified form, the residual sound pressure can be written as

$$p_{\mathrm{res}}(t) = p_{\mathrm{noise}}(t) + p_{\mathrm{anti}}(t), \tag{1.3}$$

so effective cancellation requires the anti-noise to match the scanner noise in amplitude while remaining phase-opposed and time-aligned [39].

Calibration was performed using short triangular gradient pulses for 60 seconds to derive TFs of the individual gradient axes. The same calibration stage also estimated the correction terms needed for synchronized anti-noise playback.

The overall working principle of the reported PNC system is illustrated in Figure 1.1. In this framework, an implemented MRI TTL trigger is used to trigger anti-noise playback from AFG, while previously derived TFs are used to predict the acoustic response from the MRI gradient activity. The generated anti-noise is sent from the control computer to the AFG and playback chain, and the resulting sound is delivered through the speaker box system. At the same time, the acoustic response is recorded for calibration and evaluation. This arrangement allows the system to anticipate MRI acoustic noise and generate anti-noise in synchrony with the scanner. The main components of the reported PNC system and their practical limitations are summarized in Table 1.2.

The reported PNC system achieved moderate but consistent in-bore attenuation across ten clinical MRI sequences. Mean noise reduction was reported as 9.21±1.23 dB over the 0.3–4 kHz range and 9.97±1.48 dB over the 0.5–2 kHz range, while individual sequence experiments reached reductions of up to 12.65 dB and 13.55 dB, respectively. The strongest suppression was observed in peak SPL regions around 600–1170 Hz, where SPL peaks were reduced by approximately 55.91–96.76%. Noise reduction was most effective in the mid-frequency range, where MRI acoustic noise is often the strongest. Although this attenuation remains below the strongest ANC or quiet-sequence results in raw magnitude, it is still likely to be meaningful for perceptual benefit, especially when combined with passive hearing protection [39].

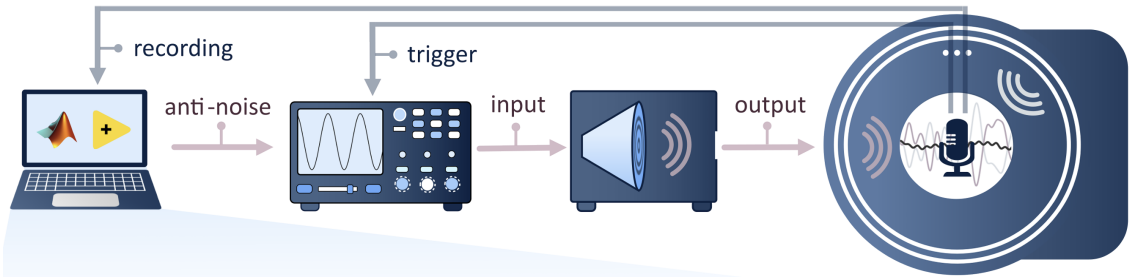


Figure 1.1: PNC setup [39].

Table 1.2: Main components of the reported PNC system and their practical role in live MRI noise cancellation [39].

Component / step	Role in the system	Main practical limitation
Clock drift correction	Prevents cumulative dephasing between recorder and scanner clocks	Requires stable calibration and accurate drift estimation
Equalization (EQ)	Compensates for playback-chain distortion caused by speaker and pneumatic transmission path	Limited by deep transmission nulls and imperfect channel inversion
LTI TFs	Predict scanner noise from gradient inputs for each physical gradient axis	Model error limits cancellation performance
Latency correction	Aligns anti-noise playback with scanner sound in time	Sensitive to timing mismatch and synchronization errors
TTL-triggered AFG output	Synchronizes anti-noise playback with MRI sequence timing	Depends on reliable trigger handling and stable system timing

The main error analysis showed that the largest efficiency loss came from imperfections in the LTI acoustic model. Playback-chain distortion was the next most important contributor, while residual latency error was comparatively small after correction. These findings indicate that further improvement depends mainly on more accurate acoustic modeling and more faithful anti-noise reproduction [39].

PNC also offers practical implementation benefits. Because the anti-noise is pre-computed, the approach reduces reliance on real-time processing and latency-critical feedback control. In principle, this allows it to be implemented without modification to the MRI scanner and used alongside standard clinical sequences. Its performance remains bounded by the stability of the acoustic TF and the accuracy of playback at the listening position. Since air-conducted cancellation cannot fully address sound transmitted through bone and tissue conduction, complete suppression of the perceived acoustic burden is unlikely [39].

In addition to these methodological limitations, the in-house implementation also limited further development and in vivo deployment. The existing pipeline relies on a combination of LabVIEW and MATLAB-based processing, which limits openness, portability, and ease of modification. Moreover, the current implementation is not yet sufficiently robust and computationally efficient for streamlined experimental use. These limitations motivate the redevelopment of the PNC pipeline in Python as part of the present thesis, to create a more accessible, maintainable, and faster framework for future experiments.

1.1.4. Relevance to EPI and Auditory fMRI

EPI is widely used in fMRI because of its high temporal resolution, but it is also one of the acoustically loudest MRI sequence types. Loud acoustic noise directly affects auditory fMRI, where scanner noise can overlap with auditory stimulus presentation, increase the acoustic baseline, and interfere with the measured BOLD response in the auditory cortex. Previous auditory fMRI studies have therefore used acquisition strategies such as sparse or clustered sampling to reduce the overlap between scanner noise and auditory stimulation [1, 2, 8, 11, 12, 14, 16, 21, 22, 29, 33, 38, 40].

Scanner noise reduction has also been investigated as a way to improve auditory fMRI conditions. Previous studies have used approaches such as ANC, quieter imaging sequences, and silent or interleaved acquisition paradigms [4, 7, 13, 29, 37]. These studies show that reducing scanner noise

can influence auditory fMRI measurements, although the observed effect depends on the acquisition scheme, stimulus paradigm, and analysis metric.

Although the present thesis does not include an in vivo auditory fMRI evaluation, auditory fMRI remains an important motivation for improving PNC during EPI. A PNC system that can reduce EPI acoustic noise without modifying the scanner sequence could provide a useful technical basis for future continuous auditory fMRI experiments. Therefore, EPI was included in the phantom validation as a more complex and application-relevant sequence type, allowing the sequence-specific limitations of PNC to be evaluated before in vivo testing.

1.2. Research Questions and Objectives

Based on this scientific and methodological context, this thesis aims to advance PNC for MRI by re-developing the existing framework, refining the system, and evaluating its acoustic performance in phantom experiments. To address this aim, the following research questions are formulated.

1. Can the original PNC pipeline be rebuilt and validated in Python to provide a more open, robust, and computationally efficient implementation?
2. How does calibration optimization affect transfer function stability and live acoustic noise reduction in phantom measurements?
3. Does the optimized calibration approach improve phantom-based acoustic noise reduction during regular MRI sequences compared with the original calibration approach?
4. How does PNC performance vary between generic and more complex sequence types, and what sequence-specific limitations remain?

Based on these research questions, the specific objectives of this thesis are formulated as follows.

1. To redevelop the original PNC workflow into a Python-based framework and validate the rebuilt implementation.
2. To evaluate how calibration optimization affects transfer function stability and live acoustic noise reduction in phantom measurements.
3. To determine the extent to which the optimized calibration approach improves phantom-based acoustic noise reduction during regular MRI sequences compared with the original calibration approach.
4. To assess how PNC performance differs between generic and more complex MRI sequence types, and to identify the sequence-specific limitations that remain.

2

Materials and Methods

This chapter describes the materials and methods used to redevelop, validate, and evaluate the PNC system. It begins by outlining the experimental setup and general PNC workflow, followed by the Python redevelopment, software-level validation, phantom-based evaluation, and data processing methods.

2.1. Experimental Setup

The phantom and in vivo experiments were performed using the same general PNC setup and workflow. This setup combined the MRI scanner, anti-noise playback chain, acoustic recording chain, and synchronization hardware required to generate and evaluate anti-noise during MRI acquisition. The following subsections introduce the hardware configuration, calibration workflow, and regular sequence workflow used throughout the study.

2.1.1. MRI and PNC Hardware Setup

All experiments were performed on a Philips Ingenia 3T MRI scanner located at Holland Proton Therapy Center, Delft, The Netherlands. As shown in Figure 2.1, the current PNC system consisted of the MRI scanner, the anti-noise playback chain, the acoustic recording chain, the TTL synchronization used to time anti-noise playback relative to the MRI sequence, and the PC used for calibration processing, anti-noise generation, and workflow control.

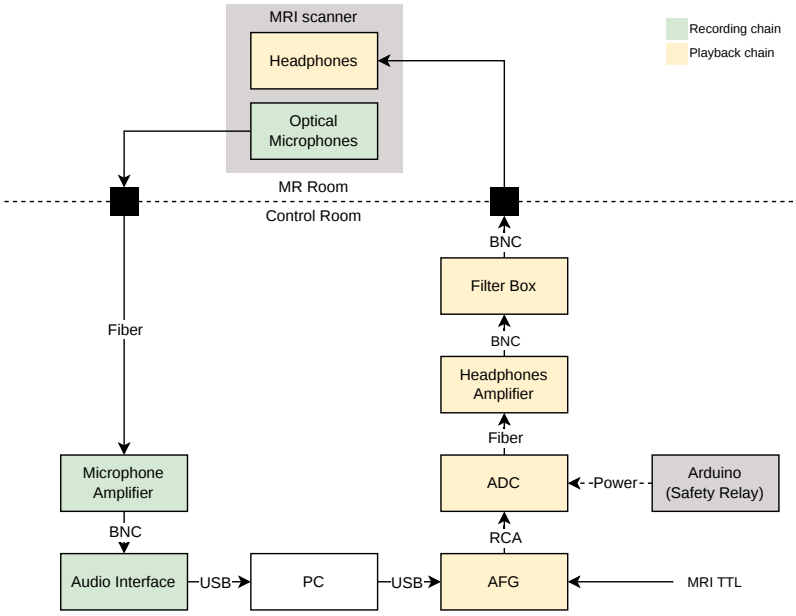


Figure 2.1: PNC setup.

Compared with the original system, which used a single-channel pneumatic speaker-box configuration with hose-and-funnel transmission into the bore, the current system used a two-channel MR-compatible headphone setup. Anti-noise playback was delivered through BOLDfonic MR-compatible headphones connected to an MR Confon amplifier. MR-compatible refers to hardware suitable for use in the MRI environment, with components near the bore designed to avoid ferromagnetic materials and minimize electromagnetic interference. To define the local measurement positions, an optical microphone was positioned on each earcup, as shown in Figure 2.2. These microphones recorded the local acoustic conditions at the left and right playback positions, which also defined the measurement probes for the quiet zone inside the bore. Anti-noise playback was generated using a Tektronix AFG31022 AFG. Acoustic recordings were acquired through a Behringer UMC202HD audio interface at 44100 Hz, the same sampling rate as used in the original PNC program.

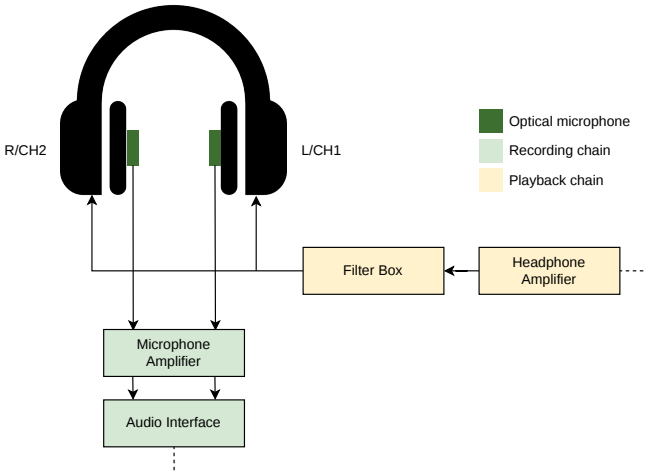


Figure 2.2: Headphone and optical microphone setup.

2.1.2. General PNC Workflow

The PNC workflow consisted of a calibration stage followed by a regular sequence stage. Calibration was used to characterize the acoustic response of the scanner and playback chain at the microphone or measurement positions. The resulting TFs, equalization filters, clock drift estimates, and latency correction were used to generate anti-noise for regular MRI sequences.

The calibration stage consisted of four steps. Step 0 measured the background noise floor, estimated audio recorder clock drift, and derived the equalizer filter for the playback chain. Step 1 recorded the scanner noise produced by known calibration gradients, derived the TFs, generated the predicted anti-noise, and estimated the simulated reduction assuming ideal headphone output and perfect temporal alignment. Step 2 recorded the predicted anti-noise playback alone, estimated the residual latency relative to the step 1 scanner noise recording, and calculated the simulated reduction using the measured headphone output while assuming perfect alignment after latency correction. Step 3 recorded scanner noise and predicted anti-noise simultaneously and quantified the live measured reduction. Meanwhile, the regular sequence stage applied the same step 1–3 logic to selected MRI sequences after calibration. Because the detailed diagrams include redevelopment-specific processing blocks, the corresponding implementation-level workflows are presented in Section 2.2.

2.2. Pipeline and Calibration Redevelopment

The redevelopment work combined software migration with calibration improvement. The original PNC workflow was rebuilt in Python, while new calibration approaches were introduced to improve TF estimation.

2.2.1. Python Pipeline Redevelopment and Optimization

Redevelopment began with a functional mapping of the original LabVIEW/MATLAB workflow in order to identify the processing blocks, instrument-control steps, and workflow dependencies that had to be pre-

served. Based on the mapping shown in Figure 2.3, a Python-based architecture was developed that incorporated DSP, AFG communication, audio recording control, workflow execution, file input/output (I/O), and backend support for the graphical user interface. The user interface was implemented in Flet, while the present work focused primarily on the backend logic that connected the interface to the processing and control stages. A detailed crosswalk of the canonical functions, workflow components, and key data artifacts used in the migration is provided in Appendix A.

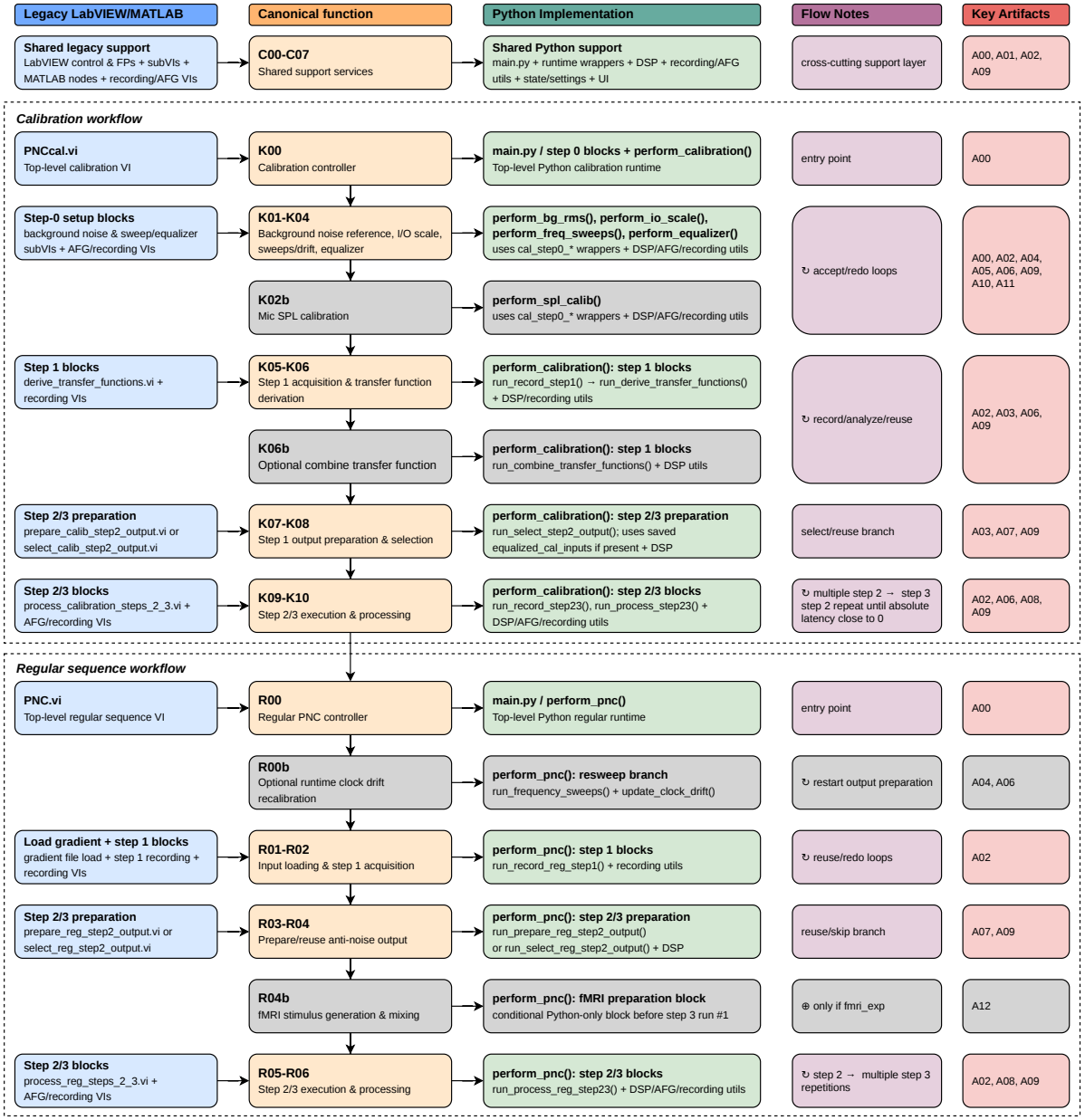


Figure 2.3: High-level crosswalk mapping between the original LabVIEW/MATLAB PNC framework and the rebuilt Python implementation.

The rebuilt pipeline was developed in Python 3.12.8. It included Python implementations of the original DSP functions using NumPy 2.3.0 and SciPy 1.18.0, with additional support from Numba 0.63.1 for accelerated numerical routines where required. AFG communication was implemented through PyVISA 1.16.1 with NI-VISA, while audio recording was handled primarily through PyAudio 0.2.14. An alternative recording backend was also included using sounddevice 0.5.5 to support WASAPI exclusive mode. Audio file I/O was handled using soundfile 0.13.1. The graphical user interface was implemented in Flet 0.80.5, with flet-charts 0.80.5 used for GUI plotting components. In addition, MATLAB integration

was included through the MATLAB 0.1 package, allowing MATLAB-based DSP functions to be called from within Python when required for direct comparison or backward compatibility.

PNC migration also included restructuring the workflow to improve performance and maintainability. These changes included modularization of the processing blocks, clearer separation between hardware-control and DSP components, simplification of data flow between workflow stages, and reduction of unnecessary overhead during calibration and regular sequence processing.

In addition to reproducing the original workflow structure, the Python redevelopment introduced targeted extensions to improve flexibility, timing stability, level control, and experimental usability. These extensions included additional calibration procedures for equalizer initial I/O scale estimation, playback-level scaling, expanded clock drift handling during calibration and regular sequence measurements, support for combining multiple TF estimates of the same LTI system, and improved workflow controls for selecting MRI sequence files, adjusting calibration settings, and monitoring experiment status. Equalizer derivation was also implemented more efficiently in Python, reducing processing overhead during calibration.

The implemented calibration workflow and its default settings are shown in Figure 2.4 and Table 2.1. The corresponding regular sequence workflow and default settings are shown in Figure 2.5 and Table 2.2. Newly implemented or modified features are shaded in gray. For regular sequence experiments used in future auditory fMRI, stimulus mixing with anti-noise and stimulus delivery were added to support EPI-based fMRI measurements. In the workflow diagram, this fMRI-specific block is indicated by a dashed outline.

The main algorithmic additions introduced in the Python workflow are described below. The I/O scale and equalized I/O scale procedures define the level-setting steps used before and after equalizer application, while the robust clock drift procedure defines the timing-correction method used to maintain alignment between playback and recorded signal.

2.2.1.1. Input/Output Scale Estimation

An I/O scale estimation was implemented to estimate the relationship between the commanded playback amplitude and the voltage recording by the microphones. This step was used to set an approximate playback scaling before equalizer derivation. The calibration signal consisted of a 1 kHz sine tone with a linear-in-dB amplitude ramp from -40 dB to 0 dB. A short fade was applied at the beginning and end of the ramp to avoid abrupt waveform discontinuities.

For each recording channel, the ramp recording was first bandpass filtered around the test frequency to isolate the 1 kHz component from the background noise and unrelated spectral components. The ramp section was then cropped after the initial low-amplitude portion, and a sliding-window RMS estimate was calculated over the remaining ramp. The corresponding commanded level, L_{cmd}, was obtained from the generated ramp envelope. A linear model was fitted between the commanded level and recorded RMS level in decibels,

$$20\log_{10}\left(V_{\mathrm{rms}}\right) = BL_{\mathrm{cmd}} + A, \tag{2.1}$$

where V_{rms} is the measured microphone RMS voltage, B is the fitted slope, and A is the fitted intercept. Only samples with sufficient signal-to-noise ratio (SNR) and below the clipping/safety threshold were included in the fit. The low-amplitude region was therefore excluded from the fitted range when it approached the background noise floor, while the high-amplitude region was excluded when it approached the selected playback safety limit.

The fitted intercept was used to estimate the recorded RMS voltage corresponding to the 0 dB commanded level,

$$V_{\mathrm{rms,0dB}} = 10^{A/20}. \tag{2.2}$$

The I/O scale for each channel was then calculated as

$$\mathrm{I/O}_{\mathrm{scale}} = \frac{V_{\mathrm{target}}}{V_{\mathrm{rms,0dB}}}, \tag{2.3}$$

where V_{target} is the desired target RMS level used for the playback test. The resulting scale factor was stored as a channel-specific scalar used for subsequent playback amplitude setting during the calibration workflow.

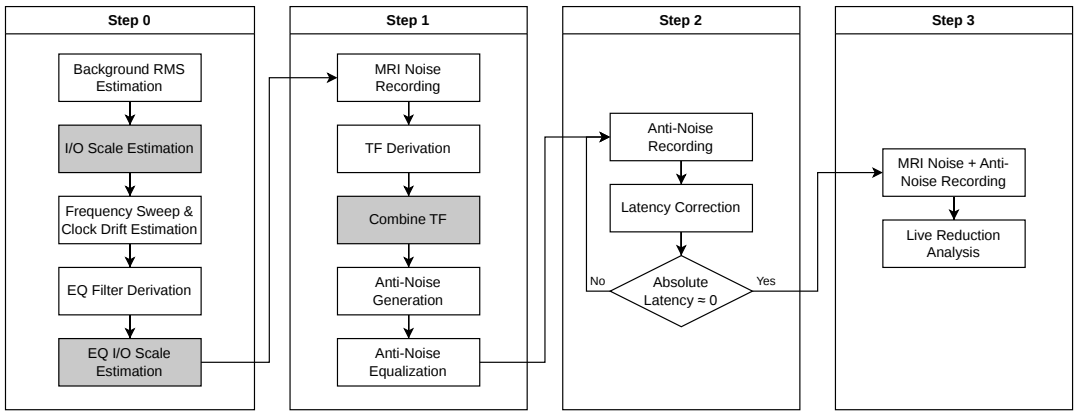


Figure 2.4: Python-PNC calibration workflow.

Table 2.1: Default parameters for the Python-PNC calibration stage.

Parameter	Value	Description / Mode
Sequence	chirp4_0608	Default calibration gradient waveform.
Averages	5	Number of repetitions for MRI noise gradient.
Background Mode / Len	calibrate / 3.0 s	Estimation of ambient noise floor.
Sweep Mode / Count / Len	on / 5 / 1.0 s	Frequency sweeps enabled for clock drift estimation.
Clock Drift Mode	from sweeps	Clock drift estimation mode.
EQ Deriv. Mode	first	Equalizer filter derivation starting order.
EQ Input	1s sweep	Reference signal for equalizer derivation.
IO Test Threshold	0.045 V	10% safety margin for signal levels.

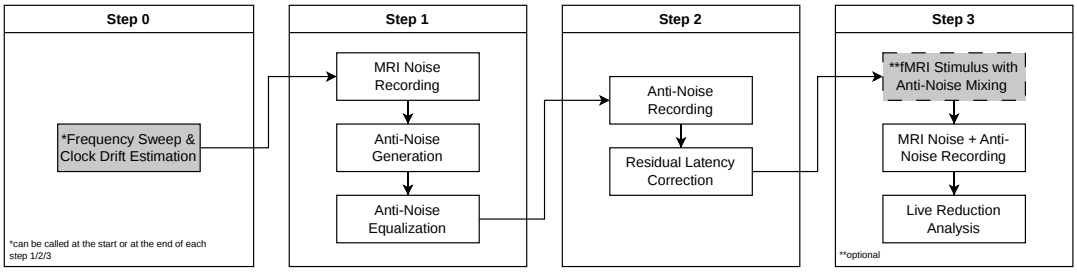


Figure 2.5: Python-PNC regular sequence workflow.

Table 2.2: Default parameters for the Python-PNC regular sequence stage.

Parameter	Value	Description / Mode
Repetitions	5	Number of iterative Step 3 processing loops.
fMRI Stimulus Level	95.0 dB	Target sound pressure level used when fMRI stimulus playback was enabled.

2.2.1.2. Equalized Input/Output Scale Estimation

An additional I/O scale estimation was implemented for the equalized playback path. This step used the same 1 kHz linear-in-dB ramp procedure described in 2.2.1.1, but the ramp was played after applying the derived equalization filter. The purpose was to estimate the effective playback level after equalization, since the equalizer can change the output amplitude even when the commanded digital input level is unchanged.

The same sliding-RMS fitting procedure was therefore applied to the equalized ramp recording, producing an equalized 0 dB RMS estimate, $V^{\mathrm{EQ}}_{\mathrm{rms,0dB}}$. The equalized I/O scale was then calculated as

$$\mathrm{I/O}^{EQ}_{scale} = \frac{V_{\mathrm{target}}}{V^{\mathrm{EQ}}_{\mathrm{rms,0dB}}}. \tag{2.4}$$

This scale was used for playback with estimated SPL after equalizer application, especially for future fMRI experiments. Separating the original and equalized I/O scales ensured that the calibration accounted for level changes introduced by the equalized playback path.

2.2.1.3. Robust Clock Drift Estimation

Clock drift correction was redeveloped to improve the stability of alignment between playback and recorded audio. Small differences between the playback and recording sampling clocks can accumulate over the duration of a measurement, causing temporal misalignment between the scanner-noise recording and the generated anti-noise waveform. In the Python workflow, clock drift estimation was applied before TF derivation, anti-noise generation, or residual estimation when drift correction was enabled.

Clock drift was estimated using repeated frequency sweep recordings. A series of chirp sweeps with increasing amplitude was generated and recorded. The first sweep was used as the timing reference, and the remaining sweeps were compared against it using sub-sample delay estimation. For each sweep i, the estimated delay d_i was associated with the corresponding sweep time t_i. Under a stable sampling-clock offset, the delay was modeled as approximately linear with time,

$$d_i \approx d_0 + \beta t_i, \tag{2.5}$$

where d_0 is the initial offset and β is the clock drift estimate.

To make the estimate robust to noisy recordings and occasional delay outliers, an initial drift estimate was calculated from the median of pairwise adjacent slopes,

$$\hat{\beta}_0 = \mathrm{median}_i \left(\frac{d_{i+1} - d_i}{t_{i+1} - t_i} \right). \tag{2.6}$$

An initial intercept was then estimated from the median residual,

$$\hat{d}_0 = \mathrm{median}_i \left(d_i - \hat{\beta}_0 t_i \right). \tag{2.7}$$

Sweeps were treated as valid if their delay estimates were close to the robust initial line,

$$\left| d_i - \left(\hat{\beta}_0 t_i + \hat{d}_0 \right) \right| < 10 \text{ samples}. \tag{2.8}$$

When enough valid sweeps were available, a first-order polynomial fit was applied to the valid delay estimates to obtain the final drift estimate,

$$\hat{\beta}_{\mathrm{dfs}} = \mathrm{polyfit}_1(t_i, d_i) \quad \text{for valid sweeps}. \tag{2.9}$$

If too few valid sweeps were available, the median slope estimate was retained as the fallback drift estimate. The recorded signal was then resampled using the corrected effective sampling rate,

$$f_{s,\mathrm{corr}} = f_s - \hat{\beta}_{\mathrm{dfs}}, \tag{2.10}$$

where f_s is the recording sampling rate. The resampled signal was used for the subsequent DSP stages.

2.2.2. Calibration Gradient Redesign

In addition to the software improvements, the redevelopment process included the redesign of alternative calibration gradients to improve PNC performance. The original triangular blip calibration, referred to as sl014, was retained as the reference method, while chirp-based gradients were introduced to increase acoustic excitation and improve spectral coverage for TF estimation. Three chirp variants were investigated. Chirp 1 used a Tukey-windowed linear chirp, chirp 2 used a Tukey-windowed chirp with amplitude adjustment to generate a flatter spectrum strength, and chirp 4 used the same amplitude-adjusted formulation as chirp 2 but with an extended excitation band. These gradients were evaluated against the original triangular blip calibration.

Chirp 1 and chirp 2 were defined over a duration $T_{\mathrm{dur}} = 20$ ms, with start and end frequencies of $f_0 = 150$ Hz and $f_1 = 4500$ Hz. The selected frequency band ensures that the chirp covers the full range of MRI acoustic noise evaluated in prior PNC work (300–4000 Hz) [39]. The instantaneous frequency was defined as

$$f_{\mathrm{inst}}(t) = f_0 + kt, \qquad k = \frac{f_1 - f_0}{T_{\mathrm{dur}}}, \tag{2.11}$$

for $0 \leq t \leq T_{\mathrm{dur}}$. The corresponding phase was

$$\phi(t) = 2\pi\left(f_0 t + \frac{1}{2}kt^2\right). \tag{2.12}$$

To reduce sharp onset and offset transitions, both chirps were multiplied by a Tukey-type window with $r = 0.1$. Chirp 1 used the windowed linear chirp,

$$g_{\mathrm{chirp}_1}(t) = A_1\, w(t) \sin\phi(t), \tag{2.13}$$

with a gradient scaling of 6 mT/m. Chirp 2 used the same chirp phase and window, but included frequency-dependent amplitude adjustment. The pre-emphasis term was defined as

$$p(t) = \frac{f_0}{f_{\mathrm{inst}}(t)}, \tag{2.14}$$

which produced the amplitude-adjusted chirp

$$g_{\mathrm{chirp}_2}(t) = A_2\, w(t)\, p(t) \sin\phi(t), \tag{2.15}$$

with a gradient scaling of 20 mT/m. Because this amplitude adjustment reduces the waveform amplitude at higher frequencies, chirp 2 was normalized before application so that its effective output level remained comparable to the intended gradient scaling.

Chirp 4 was introduced as an extended-band version of chirp 2. It used the same amplitude-adjusted chirp formulation, windowing, and duration, but with the frequency range extended to 100–5000 Hz. Therefore, for chirp 4, the instantaneous frequency and phase were calculated using $f_0 = 100$ Hz and $f_1 = 5000$ Hz in Equation 2.12 and Equation 2.14. The amplitude-adjusted chirp 4 waveform was then defined as

$$g_{\mathrm{chirp}_4}(t) = A_4\, w(t)\, p(t) \sin\phi(t), \tag{2.16}$$

with a gradient scaling of 12 mT/m.

A higher gradient amplitude produces stronger acoustic excitation and can increase the effective SNR of the calibration measurement. However, the selected gradient scaling was constrained by the scanner slew-rate limit. For the chirp-based calibration gradients, the applied scaling was therefore set to the highest allowable value for the corresponding waveform.

2.2.3. EPI-Focused Combined Transfer Function Estimation

The chirp-based calibration was introduced to provide broadband excitation with higher SNR for TF estimation. However, some regular MRI sequences, especially EPI, do not produce a broadly distributed acoustic spectrum. Instead, the acoustic response can be dominated by a small number of strong spectral peaks. A broadband calibration may not provide the most accurate estimate at the frequencies that dominate the noise. To improve prediction for EPI-like sequences, an additional EPI-Y

targeted calibration waveform was combined with a broadband calibration gradient. This mathematical combination was used because directly playing a combined chirp and EPI-Y calibration waveform was not feasible within the scanner slew rate limits.

The EPI-Y calibration waveform was designed from the frequency content of the EPI gradient waveform. The gradient was extracted from the EPI sequence file. Each of the X, Y, and Z gradient axes spectrum was inspected in frequency domain, as shown in Figure 2.6. The Y-axis gradient showed two dominant spectral components, approximately at 825 Hz and 2460 Hz. These frequencies were then used to construct a short targeted multisine calibration waveform for the Y axis.

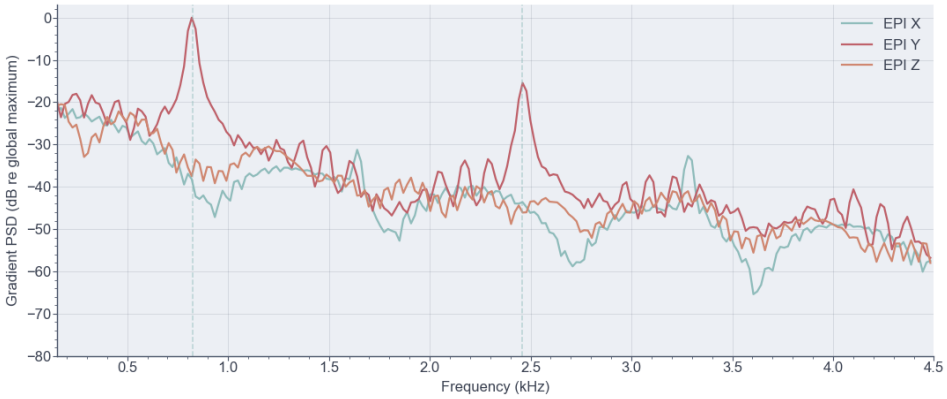


Figure 2.6: Gradient spectrum of the EPI sequence used to design the targeted EPI-Y calibration waveform.

The EPI-Y calibration waveform was implemented as a 20 ms windowed two-tone multisine,

$$g_{\mathrm{EPI-Y}}(t) = A_{EPI-Y} w(t) \left[A_0 \sin(2\pi f_0 t + \phi_0) + A_1 \sin(2\pi f_1 t + \phi_1)\right], \tag{2.17}$$

where $w(t)$ is a tapering window, $f_0 = 824.93$ Hz, $f_1 = 2459.88$ Hz, $A_0 = 1.0$, $A_1 = 0.1735$, with a gradient scaling of 19 mT/m. The waveform was normalized after generation to remain within the allowed gradient amplitude range. This EPI-Y calibration was implemented to provide a new TF from the same LTI system.

As described in Section 1.1.2, the PNC model assumes that the measured acoustic response can be predicted from the gradient input and the corresponding gradient-to-acoustic TF. In the present implementation, the TF was derived from the differentiated gradient waveform, consistent with the original PNC processing pipeline. For multiple calibration waveforms, each calibration measurement was treated as an observation of the same LTI system. The combined TF was estimated using a frequency-domain least-squares estimator, consistent with standard cross-spectral TF estimation methods [3]. At each frequency, the estimator minimized the prediction error between the measured acoustic response and the response predicted from the differentiated gradient input. For each gradient axis and microphone channel, this can be written as

$$\hat{H}_{\mathrm{comb}}(f) = \arg\min_{H} \left[\sum_k \left| P_k(f) - G_k(f) H(f) \right|^2 + \lambda |H(f)|^2 \right], \tag{2.18}$$

where $G_k(f)$ is the spectrum of the differentiated gradient waveform for calibration k, $P_k(f)$ is the corresponding measured acoustic response, and λ is a small regularization term used to avoid unstable estimates at frequencies with low input energy. Solving this regularized least-squares problem gives

$$\hat{H}_{\mathrm{comb}}(f) = \frac{\sum_k G_k^*(f) P_k(f)}{\sum_k |G_k(f)|^2 + \lambda}, \tag{2.19}$$

where $G_k^*(f)$ denotes the complex conjugate of $G_k(f)$.

For the EPI-focused implementation, the broadband chirp 1 or chirp 4 calibration and the EPI-Y-targeted calibration were combined. The combined estimate was therefore

$$\hat{H}_{\mathrm{comb}}(f) = \frac{G_{\mathrm{chirp1/4}}^*(f) P_{\mathrm{chirp1/4}}(f) + G_{\mathrm{EPI\text{-}Y}}^*(f) P_{\mathrm{EPI\text{-}Y}}(f)}{|G_{\mathrm{chirp1/4}}(f)|^2 + |G_{\mathrm{EPI\text{-}Y}}(f)|^2 + \lambda}. \tag{2.20}$$

This formulation does not average the individual TF magnitudes directly. Instead, each calibration measurement constrains the combined estimate according to the differentiated-gradient input energy available at that frequency. Frequency components that are strongly excited by the broadband chirp are mainly determined by the chirp measurement, while frequency components that are strongly excited by the EPI-Y-targeted waveform receive additional weighting from the EPI-Y measurement.

The calculation was performed independently for each gradient axis and microphone channel. The resulting frequency-domain TFs were converted back to the time domain and used for acoustic prediction and anti-noise generation. The combined TF was therefore intended to preserve broadband information from the chirp calibration while improving the estimate near the EPI-Y frequencies that were explicitly excited by the targeted waveform.

2.3. Software-Level Validation

After redevelopment, the rebuilt Python pipeline was validated against the original implementation at the software level. This validation focused on the direct comparison of corresponding DSP functions in the MATLAB and Python implementations, with an additional full PNC workflow experiment in a phantom. The purpose of this validation was to verify that the rebuilt pipeline reproduced the numerical behavior of the original framework within an acceptable tolerance while also improving computational performance.

2.3.1. DSP Function Validation

DSP validation was performed to verify that the Python implementation reproduced the numerical behavior of the MATLAB reference while reducing computation time. Most DSP functions were evaluated using matched-input comparisons, where identical input data were provided to the MATLAB and Python implementations and the corresponding outputs were compared directly.

For functions producing time-domain waveforms, such as filtering, resampling, fractional shifting, signal generation, equalization, and anti-noise generation, the Python outputs were compared with the MATLAB reference using root-mean-square error (RMSE) and signal-to-error ratio (SER). RMSE was used to quantify the absolute sample-wise deviation.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_{\mathrm{py}}[i] - x_{\mathrm{mat}}[i]\right)^2} \tag{2.21}$$

where N is the number of samples, $x_{\mathrm{py}}[i]$ is the Python output, and $x_{\mathrm{mat}}[i]$ is the MATLAB reference. To evaluate the relative error, SER was calculated as:

$$\mathrm{SER}(dB) = 10\log_{10}\left(\frac{\sum_{i=1}^{N}\left(x_{\mathrm{mat}}[i]\right)^2}{\sum_{i=1}^{N}\left(x_{\mathrm{py}}[i] - x_{\mathrm{mat}}[i]\right)^2}\right) \tag{2.22}$$

For functions that returned scalar metrics or indices, including delay estimates, RMS values, and clipping indices, validation was based on the maximum absolute difference. Output lag was also checked for waveform comparisons to ensure that numerical agreement was not affected by a residual sample shift.

Because calibration anti-noise generation depends on several consecutive DSP operations, a chained calibration validation was also performed. In this validation, the outputs of the main calibration processing stages were propagated through subsequent calibration steps and compared between MATLAB and Python. This made it possible to determine whether numerical differences observed at the function level affected TF derivation, predicted noise generation, anti-noise equalization, anti-noise recording analysis, latency correction, and the resulting reduction metrics.

A function or processing stage was considered to show strong numerical agreement when scalar outputs matched exactly or when waveform outputs achieved an SER above 60 dB. SER was used as the main numerical criterion for comparing DSP outputs. When SER was below 60 dB, the interpretation also inspected to identify whether the difference arose from implementation-sensitive operations. The downstream effect of the discrepancy was then evaluated, especially its influence on predicted or estimated noise reduction. If the downstream effect could not be fully assessed at the individual

DSP-function level, it was evaluated later through the integrated workflow validation. Execution time was recorded for all evaluated functions in order to quantify the computational speed-up achieved by the Python implementation.

2.3.2. Integrated Workflow Validation

Complementing direct DSP validation, the rebuilt pipeline was evaluated under real experimental conditions. This comparison was intended to verify that the migrated system reproduced the practical behavior of the original workflow at the level of integrated calibration and live anti-noise application.

The real-time comparison included the full calibration workflow using the original triangular calibration gradient and a full step 1–3 comparison on a short regular MRI sequence (TR 15 ms, TE 8.0 ms, FOV 230x230x1 mm, VOX 1x1x1 mm, coronal plane with 0°/0°/0° angulation). The compared calibration outputs included background RMS estimation, clock drift estimation, equalizer derivation, and calibration step 1–3 processing. The short sequence comparison extended this validation to the full experimental workflow and allowed direct comparison of the resulting anti-noise performance under actual acquisition conditions. In addition, the total runtime of the calibration and regular sequence workflow was measured to quantify the practical speed of improvement of the Python implementation.

2.4. Phantom-Based Evaluation

This section describes the phantom-based experiments used to evaluate the acoustic performance of the PNC system through calibration comparison and regular MRI sequence evaluations.

2.4.1. MRI Phantom

Phantom experiments were performed using a wooden head phantom with silicone ears and a 3 L liquid-bottle phantom. The head phantom was used to approximate normal headphone placement during MRI measurements. During acquisition, the headphones were positioned on the head phantom in the same way as in a normal experimental setup. The phantom was then placed at the isocenter inside the MRI scanner.

2.4.2. Calibration Comparison

Calibration comparison was performed to evaluate differences in TF derivation and live acoustic reduction between the original calibration gradient and newly introduced calibration gradients. The comparison was carried out in two measurement sessions. The first session included the original sl014 calibration, chirp 1, chirp 2, and the mathematically combined chirp 1+EPI-Y derived TF. The second session included sl014, chirp 1, chirp 4, the combined chirp 1+EPI-Y derived TF, and the combined chirp 4+EPI-Y derived TF. Chirp 4 was included to evaluate the effect of using a wider excitation band compared with chirp 2.

For each calibration approach, the evaluated outputs included TF stability metrics and calibration step 3 acoustic reduction results. TF stability was assessed using magnitude variability, complex TF variability, coherence, and leave-one-out cross-validation SER. Live calibration performance was evaluated using step 3 reduction, including the mean reduction across gradient axes and recording channels and the worst axis-channel reduction. Axis- and channel-specific step 3 reductions were also compared to assess whether the calibration performance was consistent across playback and recording paths. Each calibration gradient measurement was repeated three times to assess repeatability and performance robustness.

2.4.3. Regular MRI Sequences Evaluation

Short regular MRI sequences were used to evaluate acoustic noise reduction performance under controlled phantom conditions. These measurements were used to assess the achieved reduction across different sequence types, to evaluate the effect of calibration gradient choice, and to examine whether sequence-specific spectral characteristics influenced PNC performance.

Two groups of sequences were evaluated. The first group consisted of short gradient echo sequence variants implemented on the Philips system as fast field echo (FFE) sequences. The same ten representative sequences as the previous study were used to allow comparison with earlier PNC measurements [39]. These sequences were used to evaluate PNC performance across a range of sequence parameters, including TR, TE, FOV, voxel size, orientation, and angulation.

For the main FFE evaluation, four broadband calibration approaches were compared: sl014, chirp 1, chirp 2, and chirp 4. This allowed the general effect of calibration-gradient choice to be assessed across broadband-dominated regular sequence conditions. The FFE comparison was first performed in an initial measurement session, and the same ten FFE variants were then repeated in two additional sessions using the same four calibration approaches. Within each session, five repeated live-reduction recordings were acquired for each FFE sequence-calibration condition. These additional sessions were included to assess session-to-session variability in the achieved reduction and to examine whether the relative ranking of the calibration gradients was stable across independent measurement sessions. Because setup and equalizer quality could differ between sessions, the session results were summarized separately.

During the initial FFE session, the phantom and headphone setup were manually repositioned using guide points on the phantom before the acquisition of sequence 7 and then restored using the available guide points. This procedural note was recorded because small changes in phantom position, headphone fit, or microphone placement can affect the acoustic transfer path and downstream reduction.

A separate comparison between chirp 4 and the combined chirp 4+EPI-Y derived TF was performed to assess whether adding EPI-Y information affected the broadband reduction capability of chirp 4 during FFE measurements. This comparison was treated separately from the three-session broadband calibration comparison because it was designed specifically to test whether the EPI-targeted TF changed regular FFE performance. The tested FFE variants are listed in Table 2.3.

Table 2.3: FFE sequence variants used for phantom-based acoustic evaluation.

Sequence	TR / TE (ms)	FOV (mm)	VOX (mm)	Plane/Angulation (AP/RL/FH)
1	15 / 8.0	230 × 230 × 1	1 × 1 × 1	Coronal, 0°/0°/0°
2	7.6 / 3.7	230 × 230 × 10	1 × 1 × 10	Transverse, 14°/-27°/12°
3	9.7 / 5.7	330 × 330 × 1	1 × 1 × 1	Transverse, 30°/25°/10°
4	5.1 / 3.4	220 × 220 × 5	1 × 1 × 5	Sagittal, 0°/0°/0°
5	44 / 29	230 × 230 × 10	1 × 1 × 10	Coronal, 0°/0°/0°
6	25 / 11	212 × 212 × 1	2 × 2 × 1	Coronal, 5°/29.38°/10°
7	20 / 3.7	230 × 230 × 20	1 × 1 × 20	Sagittal, 45°/2°/34°
8	14 / 1.1	330 × 330 × 5	3 × 3 × 5	Coronal, 30°/25°/16.83°
9	27 / 1.1	330 × 330 × 5	3 × 3 × 5	Coronal, 30°/25°/10°
10	18 / 3.0	200 × 200 × 10	0.5 × 0.5 × 10	Sagittal, 30°/25°/10°

The second group consisted of continuous EPI acquisitions. These measurements were included because EPI produced a more concentrated acoustic spectrum with strong spectral peaks, making it a useful test case for the EPI-focused combined TF approach. The EPI sequence designed for a future in vivo auditory fMRI experiment was applied, as summarized in Table 2.4. For the phantom EPI evaluation, chirp 4, the combined chirp 1+EPI-Y derived TF, and the combined chirp 4+EPI-Y derived TF were compared. Chirp 4 was used as the chirp-only broadband reference, while the two combined TFs were used to test whether adding EPI-targeted calibration information improved reduction for EPI-like spectral content. The inclusion of both chirp 1+EPI-Y and chirp 4+EPI-Y allowed the effect of the EPI-Y component to be evaluated with different broadband chirp baselines.

For the repeated EPI live-reduction evaluation, each EPI condition was measured over five repetitions. Additional repeated noise-only and anti-noise-only recordings were acquired for the EPI 10-dynamics condition to assess whether changes in reduction were associated with scanner-noise variation, anti-noise playback variation, or time-varying mismatch between the two.

2.5. Data Processing and Analysis

Acoustic recordings were processed to quantify PNC performance during calibration and phantom-based regular sequence measurements. The analysis included signal alignment, bandpass filtering, spectral analysis, and calculation of reduction metrics from the scanner-only and residual recordings. The calculations are identical to those of the previous study, with the implementation rebuilt in Python [39]. Unless stated otherwise, acoustic reduction was evaluated over the 500–4000 Hz band and

Table 2.4: Acquisition parameters for EPI.

Parameter	Value
TR / TE	3000 / 35 ms
Flip angle	83.04°
Acquisition plane	Transverse
FOV / VOX	210 × 210 × 150 mm / 2.38 × 2.40 × 3.50 mm
Matrix size / slices	88 × 88 / 43
Slice gap	0 mm
Number of dynamics	10 & 33

expressed in decibels. For FFE regular sequence analysis, an additional narrowband evaluation over the 500–2000 Hz band was included to assess reduction in the lower-to-mid frequency range.

For each reduction measurement, the reference signal was defined as the scanner-noise recording without anti-noise playback, and the residual signal was defined as the corresponding recording obtained during scanner noise with anti-noise playback. For the main live reduction analyses, step 3 was used as the primary reduction metric for both calibration and regular sequence evaluation, unless otherwise stated. Step 2 results were used only for intermediate workflow validation for redeveloped pipeline validation. Broadband reduction was calculated from the RMS ratio between these two signals,

$$R_{\mathrm{dB}} = 20\log_{10}\left(\frac{p_{rms,noise}}{p_{rms,res}}\right), \tag{2.23}$$

where p_{noise} is the scanner-only recording and p_{res} is the residual recording after PNC. The same calculation was also applied to selected frequency bands when narrower-band behavior was relevant.

For calibration evaluation, all reduction was calculated for each gradient axis and channel. Step 3 indicates the live reduction recorded via the recording chain. These values were used to compare the original sl014 calibration, the chirp-based calibration gradients, the combined chirp 1+EPI-Y derived TF, and the combined chirp 4+EPI-Y derived TF. In addition to mean reduction, the worst-axis-channel reduction was reported to assess whether a calibration approach performed consistently across gradient axes and playback channels.

For regular sequence evaluation, reduction metrics were calculated for each recording channel and repeated step 3 measurement. Step 3 was used as the main live reduction metric. For FFE sequences, the analysis focused on broadband reduction over 500-4000 Hz, narrowband reduction over 500-2000 Hz, improvement relative to the original sl014 calibration, and temporal and spectral residual behavior. Broadband and narrowband reductions were evaluated both to quantify overall acoustic reduction and to assess performance in the lower-to-mid frequency range where dominant FFE components were observed. For EPI, additional analyses were performed because the residual was more time-varying and peak-dominated. These analyses included tracking broadband reduction across repetitions, measuring the residual peak near the dominant EPI component, and comparing repeated noise-only and anti-noise-only spectra to assess amplitude stability.

Spectral analysis was used to identify which frequency components were effectively suppressed and which components remained in the residual. Welch-based spectra were used for noise and residual comparisons, while gradient spectra were inspected to relate the acoustic response to the input gradient waveform.

3

Results

This chapter presents the results of the redevelopment, validation, and evaluation of the PNC system. The rebuilt Python-based pipeline was assessed at the software level by comparing its DSP outputs and workflow-level performance with those of the original LabVIEW/MATLAB implementation. Then, phantom experiments were used to evaluate acoustic performance, including the comparison of optimized calibration approaches and noise reduction across regular MRI sequence variants.

3.1. Python Pipeline Redevelopment

The original PNC workflow was rebuilt as a Python-based framework for calibration and regular sequence experiments. The implemented pipeline included DSP processing, audio recording, AFG communication, file handling, workflow execution, and backend support for the user interface. The main calibration and regular sequence workflow structure of the original PNC program was retained. The complete source code of the program is available in Appendix B.1.

A graphical user interface was implemented for system settings, calibration, regular sequence processing, and experiment monitoring. A representative calibration page of the rebuilt interface is shown in Figure 3.1, with the complete interface available in Appendix B.2.

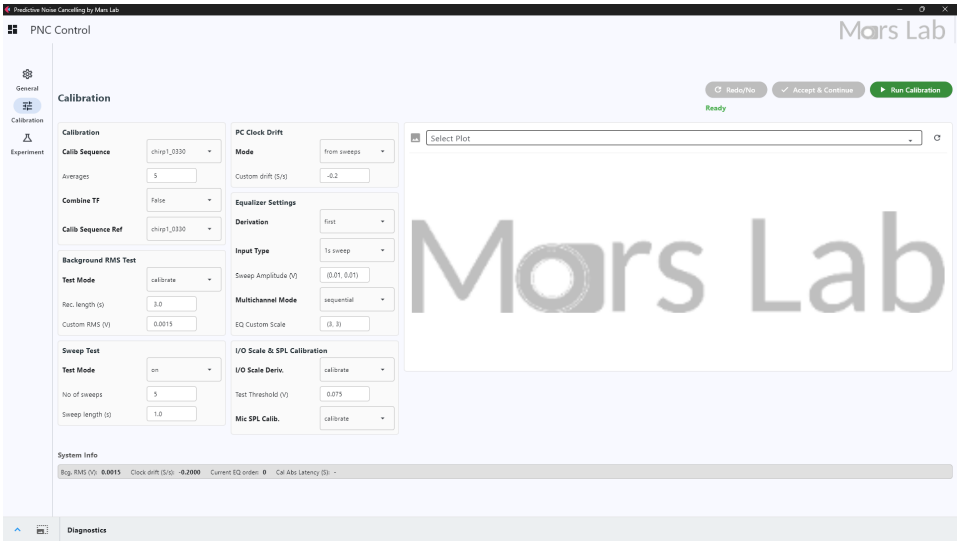


Figure 3.1: Representative view of the rebuilt Python-based PNC interface.

Compared with the original workflow, the rebuilt Python pipeline added support for clock drift handling, easier reuse of calibration outputs during regular sequence experiments, SPL-related calibration information, flexible audio recording options, and mathematical combination of multiple TF estimates.

The combined-TF functionality allowed broadband and sequence-specific calibration measurements to be merged for anti-noise generation.

3.2. Validation of the Rebuilt Pipeline

The rebuilt Python-based PNC pipeline was first validated against the original implementation. Validation was performed by comparing the individual DSP functions against the MATLAB reference and the integrated workflow outputs under calibration and regular sequence conditions.

3.2.1. DSP Validation

At the function level, most deterministic DSP functions showed high numerical agreement with the MATLAB reference. As shown in Figure 3.2 (a), preprocessing and timing functions, frequency sweep processing, regular sequence prediction and equalization, and the final latency-shifted output all reached SER values above 60 dB. Lower SER values were observed mainly in equalizer design and application, TF derivation, and predicted calibration acoustic noise.

The most pronounced measured improvement of the Python implementation was reduced computation time. As shown in Figure 3.2 (b), Python was faster across all evaluated DSP stages. The largest speed-up was observed for preprocessing and equalizer-related processing, while the smallest speed-up was observed for regular sequence anti-noise equalization. In addition to the numerical agreement metrics, all waveform comparisons showed zero sample lag between the Python and MATLAB outputs. This result shows that the Python implementation did not introduce a residual time shift relative to the MATLAB reference. Overall, the validation showed that the Python implementation preserved the main DSP behavior of the MATLAB reference while significantly reducing its computation time.

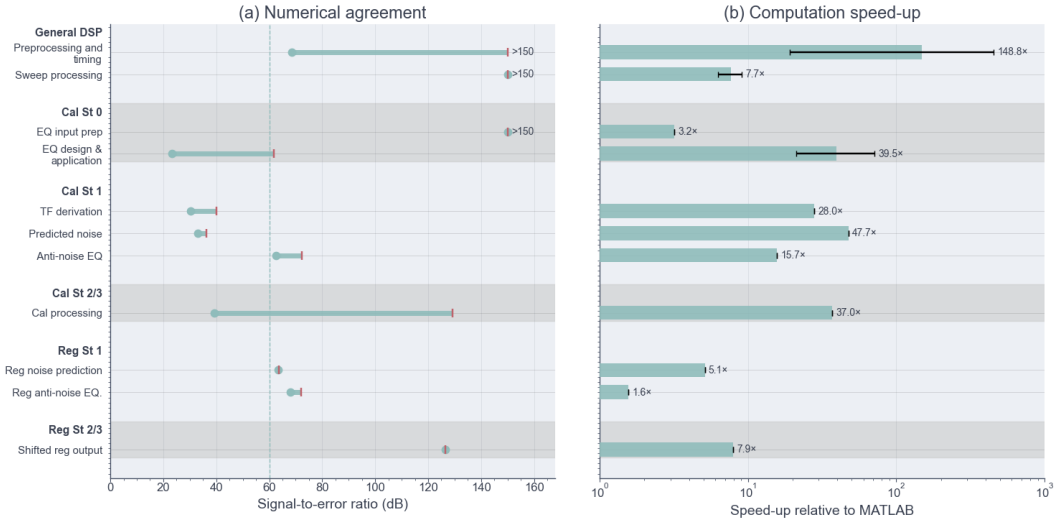


Figure 3.2: DSP validation of the rebuilt Python implementation according to its (a) numerical agreement and (b) computational speed-up.

Although lower SER values were observed in some functions, especially in intermediate calibration-chain outputs, the downstream reduction metrics remained closely matched between MATLAB and Python, as summarized in Figure 3.3. The differences in predicted and aligned reduction values were small, with Python producing slightly higher values in several cases and slightly lower values in others.

3.2.2. Integrated Workflow Validation

The close agreement observed at the DSP level was largely preserved when the full calibration and regular sequence workflows were compared in the same scanning session. The mean workflow-level errors stayed under 5%, as summarized in Figure 3.4 (a). Excluding the flagged regular sequence step 1 result, the median workflow-block error was 2.98%, and the maximum workflow-block error was 4.39%.

The representative regular sequence comparison further showed that the Python implementation

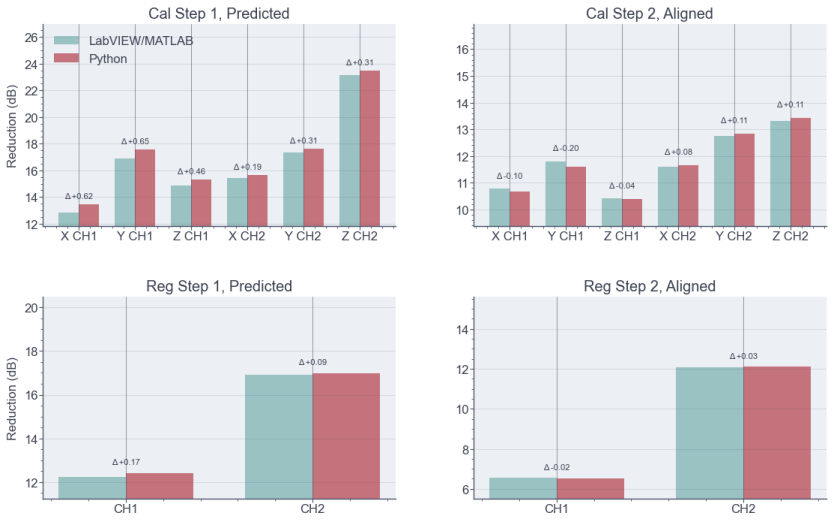


Figure 3.3: Downstream reduction metric agreement between MATLAB and Python DSP outputs

reproduced the intended anti-noise behavior of the original workflow. For the regular sequence, step 2 and step 3 reduction values were closely matched, if not improved, between LabVIEW/MATLAB and Python, as shown in Figure 3.4 (b). The regular step 3 values showed close agreement between the two implementations, with a mean workflow-block error of 2.99%. In addition to preserving the acoustic reduction behavior, the Python implementation substantially reduced the full workflow execution time by approximately 73.6%, as shown in Figure 3.4 (c).

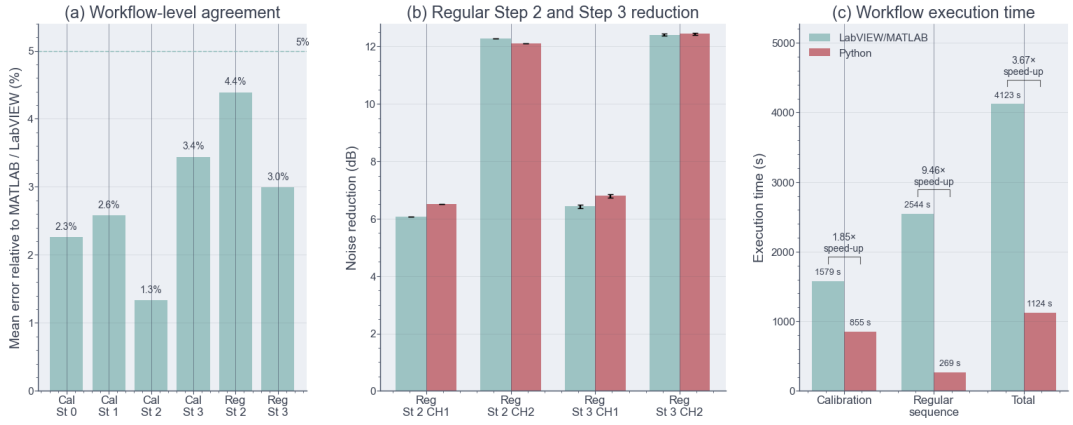


Figure 3.4: Integrated workflow validation, shown by (a) mean workflow-level error relative to the LabVIEW/MATLAB reference, (b) direct comparison of ffe2 step 2/3 reduction values between implementations, and (c) full workflow execution time.

The main exception was the step 1 result for the regular sequence. The original implementation produced negative reduction values, while the Python implementation produced positive reduction values. This discrepancy was limited to step 1 and was not propagated to step 2 or step 3.

3.3. Phantom-Based Acoustic Evaluation

Phantom experiments were used to evaluate the acoustic performance of the rebuilt PNC system under controlled MRI conditions. Results are presented for both calibration and regular MRI sequence measurements.

3.3.1. Calibration Gradient Evaluation

Phantom-based calibration measurements were used to compare the original triangular blip calibration, sl014, with the newly introduced chirp-based gradients. The comparison included sl014, chirp 1, chirp

2, chirp 4, the combined chirp 1+EPI-Y TF, and the combined chirp 4+EPI-Y TF. The physical calibration waveforms are shown in Figure 3.5, using the same time and amplitude scale for comparison. For the combined-TF conditions, chirp 1 or chirp 4 was used as the physical calibration gradient during step 1–3 measurements, while the combined chirp+EPI-Y TF was used for anti-noise generation.

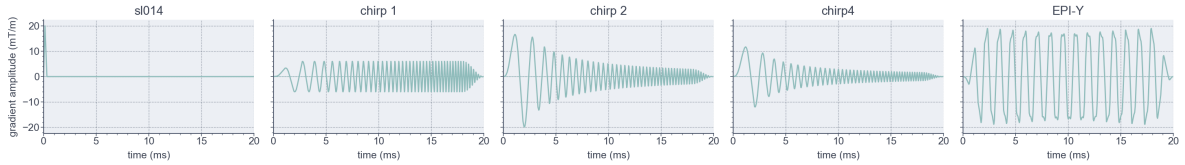


Figure 3.5: Calibration gradient waveforms.

The measured step 1 calibration responses differed between the physically played calibration gradients. The absolute peak of the averaged recorded calibration response, $p_{\mathrm{av,xyz}}$, was used to compare the recorded acoustic level. The differentiated gradient spectrum and the Welch spectrum of $p_{\mathrm{av,xyz}}$ were used to compare the effective input excitation and the recorded acoustic response across the calibration band. Compared with sl014, the chirp-based gradients produced higher recorded acoustic levels and higher acoustic excitation. Chirp 1, chirp 2, and chirp 4 showed stronger broadband acoustic responses than sl014, although their spectral distributions differed. Chirp 2 showed stronger excitation in some lower-frequency regions and axis-channel combinations, while EPI-Y produced a more sequence-specific excitation pattern. The recorded level and spectral coverage of the tested physical calibration gradients are summarized in Figure 3.6. Roughly, the recorded acoustic spectrum has a similar magnitude response distribution to each respective gradient spectrum. The combined TF is not shown as a separate excitation because it was estimated from separate chirp 1 or chirp 4 and EPI-Y measurements.

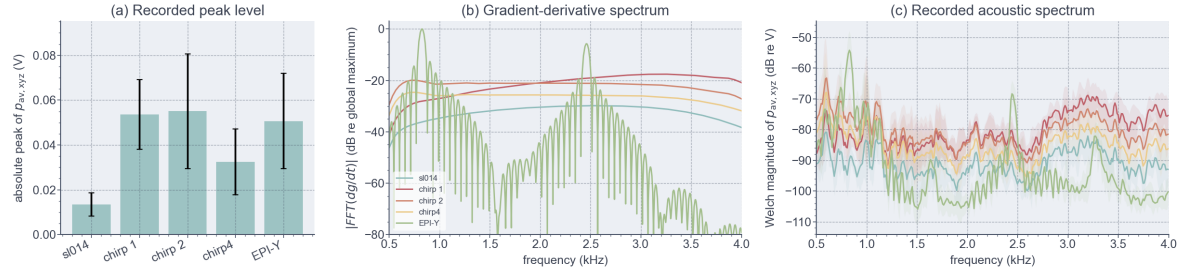
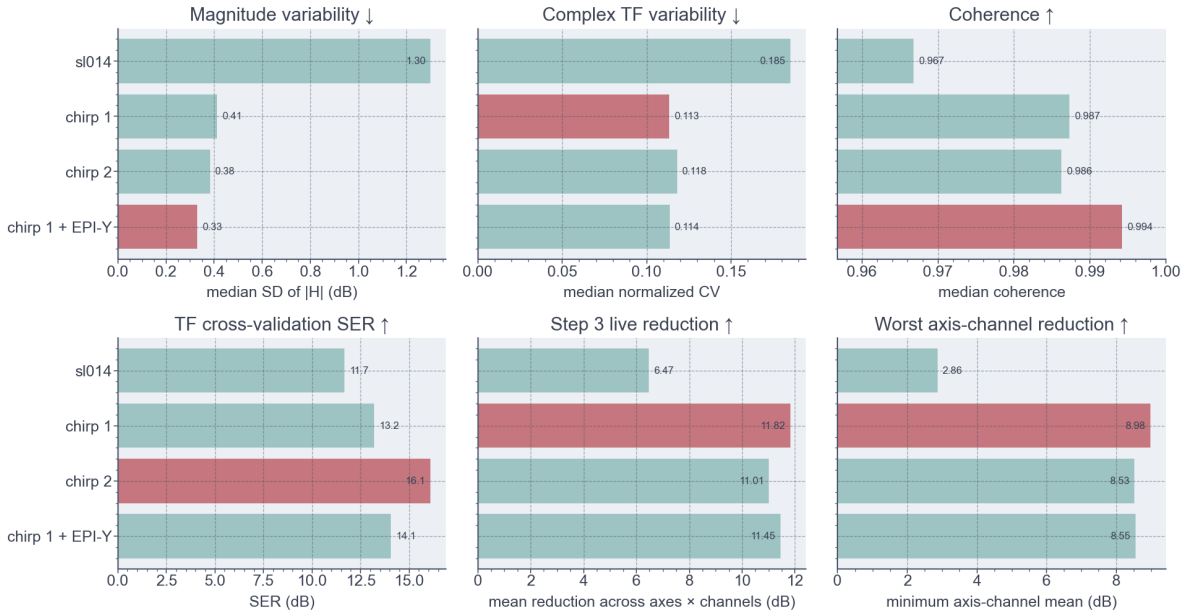


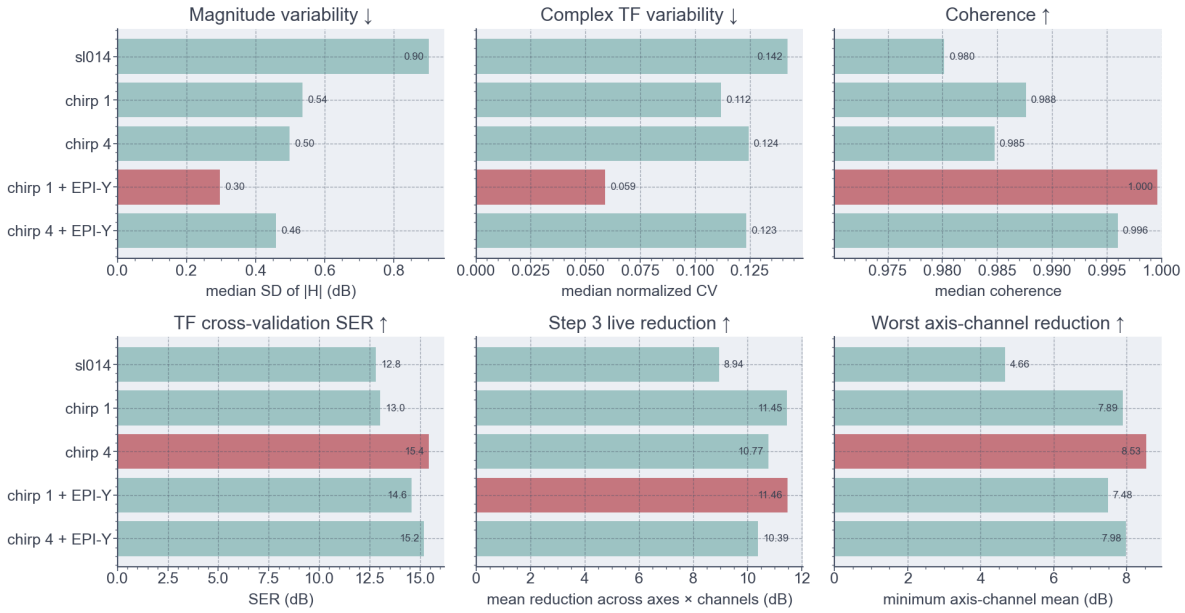
Figure 3.6: Calibration gradient excitation and recorded acoustic response.

The calibration gradient comparison was conducted across two independent measurement sessions. The first session compared sl014, chirp 1, chirp 2, and the combined chirp 1+EPI-Y derived TF, while the second extended the comparison to include chirp 4 and the combined chirp 4+EPI-Y derived TF in place of chirp 2. As the two sessions were conducted on different days under different measurement conditions, the results are not directly comparable across sessions, but they provide a consistent picture of relative calibration performance. The quantitative TF stability metrics and calibration step 3 performances from both sessions are summarized in Figure 3.7.

Across both sessions, sl014 consistently showed the highest TF magnitude variability and the lowest step 3 mean and worst axis-channel reduction. The chirp-based approaches showed lower TF variability than sl014 in both sessions. In the first session, chirp 2 produced the highest TF cross-validation SER, while the combined chirp 1+EPI-Y derived TF showed the highest median coherence and the lowest TF magnitude variability. Despite this, chirp 1 produced the highest step 3 live reduction and the highest worst axis-channel reduction. In the second session, the combined chirp 1+EPI-Y derived TF again showed the lowest TF magnitude and complex variability and the highest median coherence. Chirp 4 produced the highest TF cross-validation SER, but like chirp 2 in the first session, this did not translate to the highest step 3 reduction. Chirp 1 and the combined chirp 1+EPI-Y derived TF produced the highest mean step 3 reductions, while chirp 4 produced the highest worst axis-channel reduction.



(a) First calibration session: sl014, chirp 1, chirp 2, and combined chirp 1+EPI-Y.



(b) Second calibration session: sl014, chirp 1, chirp 4, combined chirp 1+EPI-Y, and combined chirp 4+EPI-Y.

Figure 3.7: TF stability metrics and calibration step 3 performance across two independent calibration sessions. Highlighted bars indicates the best performing variant for each metric.

The axis- and channel-specific step 3 reductions from the second session are shown in Figure 3.8. The chirp-based and combined TF approaches produced higher reductions than sl014 in most axis-channel combinations, with the exception of X CH1 and Y CH1, where sl014 performed comparably to or better than the chirp-based approaches. The most pronounced underperformance of sl014 was observed in Z CH1, where it produced substantially lower reduction than all chirp-based approaches. Chirp 1+EPI-Y produced the highest reduction for X CH2 and Y CH2, while chirp 1 and chirp 1+EPI-Y produced the highest reductions for Z CH1 and Z CH2 respectively. Larger error spreads were observed for chirp 1 in X CH2 and Y CH2, while the other chirp-based approaches showed smaller variation across repeated measurements. Across all gradient axes, CH2 consistently showed higher step 3 reductions than CH1.

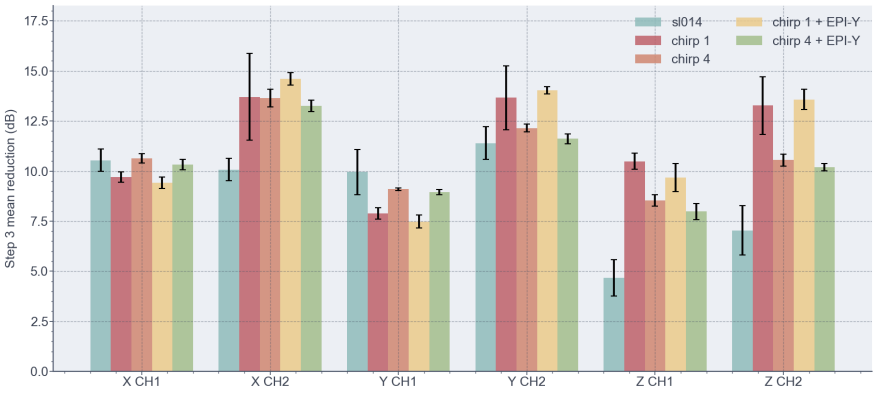


Figure 3.8: Calibration step 3 mean live reduction by gradient axis and recording channel in the second calibration session.

As channel imbalance was observed in the calibration results, a diagnostic test was performed by comparing the better channel playback of the normal setup with the phantom head flipped upside down setup and the headphone flipped right-left setup. The live reduction of the calibration is shown in Figure 3.9. Differences in reduction were observed, but the discrepancies are not as extreme as the imbalance between the two channels observed in Figure 3.8.

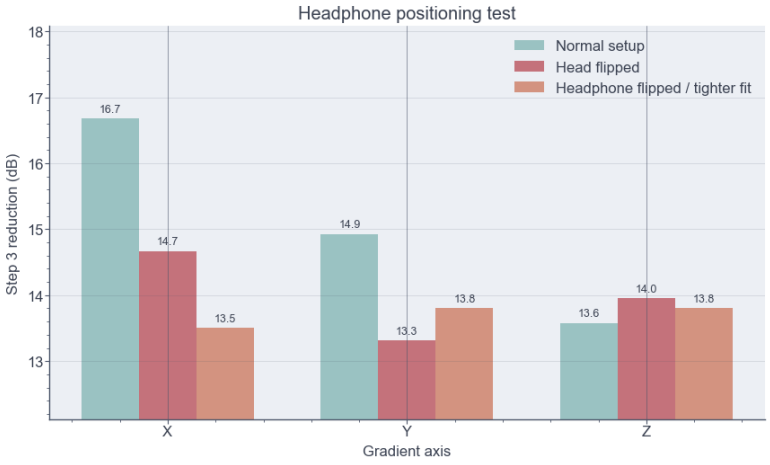


Figure 3.9: Channel imbalance test with the better channel playback.

3.3.2. Regular MRI Sequence Evaluation

Following calibration, live PNC performance was evaluated during short phantom-based MRI acquisitions. The regular sequence evaluation was divided into FFE sequence variants, which were used to assess general phantom reduction performance, and EPI variants, which were used to assess performance for a more peak-dominated and temporally demanding sequence type.

3.3.2.1. Representative FFE Sequence Evaluation

For the FFE evaluation, ten sequence variants were measured and compared across the evaluated calibration approaches. Live reduction was summarized using broadband reduction, narrowband reduction, improvement relative to the original sl014 calibration, temporal repeatability across repeated step 3 measurements, frequency-dependent spectral reduction, and session-to-session differences in the equalizer filter reduction quality metric.

For the FFE sequence variants, the mean reduction across CH1 and CH2 is summarized in Table 3.1. The values are reported descriptively, because repetitions from the same sequence, session, and playback setup are not fully independent. The initial-session mean reductions are visualized across individual sequences in Figure 3.10. The initial measurement session showed higher mean reductions for the chirp-based calibration approaches than for the original sl014 calibration in both the broadband and narrowband analyses. In this initial session, chirp 4 produced the highest average reduction across all FFE sequences.

Table 3.1: Mean FFE reduction across CH1 and CH2 over three measurement sessions. Each calibration/session/band includes 10 FFE sequence variants, five repeated playbacks, and two microphone channels ($n = 100$ sequence-repetition-channel measurements). BB = 500–4000 Hz and NB = 500–2000 Hz. Values are mean $\pm$ standard deviation in dB; bold and underline indicate the highest and second-highest values within each session/band.

Calibration	Initial session		Repeat session 1		Repeat session 2	
	BB	NB	BB	NB	BB	NB
sl014	8.76 $\pm$ 2.53	9.51 $\pm$ 3.21	9.24 $\pm$ 2.06	10.11 $\pm$ 2.23	9.24 $\pm$ 1.65	10.17 $\pm$ 1.78
chirp 1	<u>10.79 $\pm$ 1.71</u>	<u>11.69 $\pm$ 2.07</u>	**9.94 $\pm$ 1.61**	**11.09 $\pm$ 1.55**	**9.84 $\pm$ 1.47**	**10.63 $\pm$ 1.53**
chirp 2	10.30 $\pm$ 1.46	10.99 $\pm$ 1.40	8.76 $\pm$ 0.80	9.43 $\pm$ 0.78	8.45 $\pm$ 0.98	9.01 $\pm$ 1.07
chirp 4	**10.89 $\pm$ 1.69**	**12.12 $\pm$ 2.23**	<u>9.66 $\pm$ 1.60</u>	<u>10.85 $\pm$ 1.80</u>	<u>9.64 $\pm$ 1.17</u>	<u>10.34 $\pm$ 1.24</u>

The FFE evaluation was repeated in two additional measurement sessions to summarize session-to-session variations in the mean reduction values. The repeated sessions did not show a similar low sl014 reduction for sequence 7. At the same time, the overall reduction values were lower. In repeat sessions 1 and 2, chirp 1 produced the highest mean reduction in the broadband and narrowband analyses.

Across the three sessions, differences were observed in both the absolute reduction values and the ranking of the calibration approaches. In the initial session, chirp 4 produced the highest mean reduction in both analysis bands, with chirp 1 producing the second-highest values. In both repeated sessions, chirp 1 produced the highest mean reduction, while chirp 4 produced the second-highest values. Chirp 1 and chirp 4 showed higher mean reductions than sl014 in both analysis bands across all three sessions. In contrast, chirp 2 showed higher mean reductions than sl014 in the initial session, but lower mean reductions than sl014 in both repeated sessions and in both analysis bands. The narrowband reduction was consistently higher than the broadband reduction across all calibration approaches and sessions.

The equalizer filter reduction quality metric for the chirp calibration also differed between measurement sessions, as summarized in Table 3.2. The initial session had the highest channel-averaged equalizer filter reduction metric, with 27.35 dB, compared with 24.75 dB in repeat session 1 and 24.55 dB in repeat session 2. Imperfect residual equalizer filter components were observed near the analysis-band boundaries, especially around 500 Hz and 4000 Hz, with more prominent residual components in the repeat-session measurements.

Table 3.2: Equalizer filter reduction quality metric for the chirp calibration across measurement sessions. Values are shown for CH1 and CH2, together with the channel-averaged value.

Session	CH1 (dB)	CH2 (dB)	Mean (dB)
Initial session	28.2	26.5	27.35
Repeat session 1	24.0	25.5	24.75
Repeat session 2	25.4	23.7	24.55

During the initial session, sequence 7 showed a lower sl014 reduction than the other FFE variants, which decreased the overall mean and increased the standard deviation. This scan was acquired after

manual repositioning of the phantom and headphone setup, which may have changed the acoustic transfer path despite using guide points to restore the position. Sequence 7 was included in the primary analysis, but an additional sensitivity analysis excluding it was performed to assess its impact, as shown in Table 3.3. Removing this variant increased the sl014 mean in both broadband and narrowband evaluations. Chirp 4 remained the highest-performing calibration in this initial-session sensitivity analysis, followed by chirp 1.

Table 3.3: Initial-session FFE mean reduction across CH1 and CH2 with sequence 7 excluded. Each calibration/band includes nine FFE sequence variants, five repeated playbacks, and two microphone channels ($n = 90$ sequence-repetition-channel measurements). BB = 500–4000 Hz and NB = 500–2000 Hz. Values are mean $\pm$ standard deviation in dB; bold and underline indicate the highest and second-highest values in each column.

Calibration	BB	NB
sl014	9.37 $\pm$ 1.77	10.24 $\pm$ 2.47
chirp 1	<u>11.16 $\pm$ 1.37</u>	<u>12.15 $\pm$ 1.63</u>
chirp 2	10.25 $\pm$ 1.34	10.93 $\pm$ 1.34
chirp 4	**11.29 $\pm$ 1.24**	**12.64 $\pm$ 1.65**

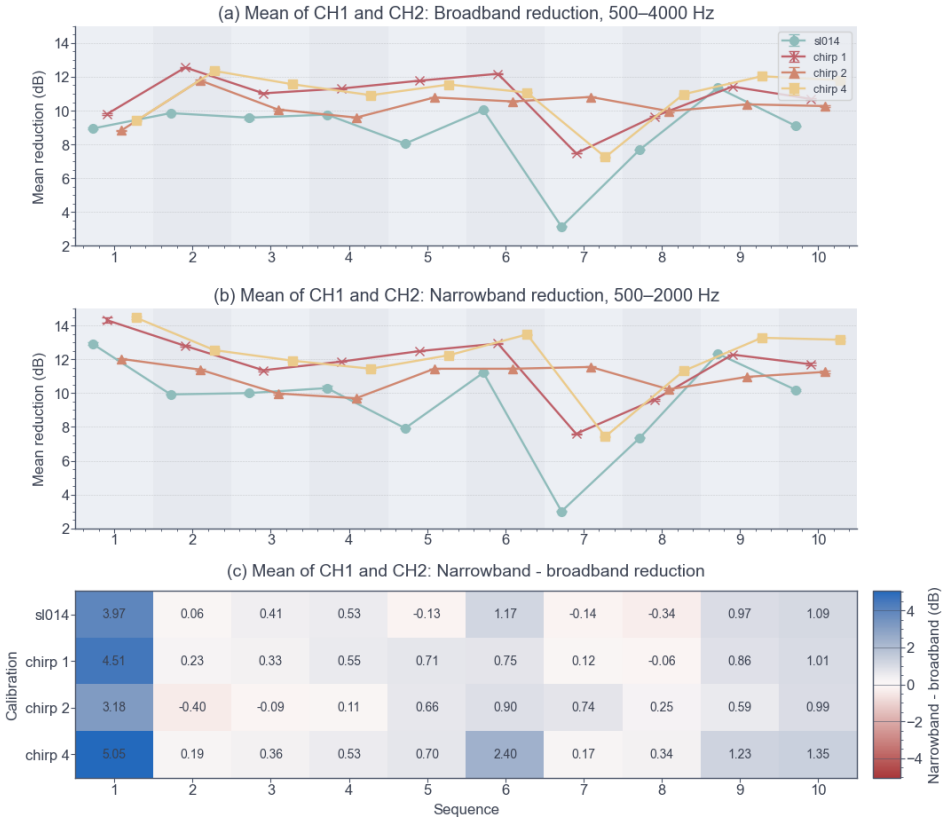


Figure 3.10: FFE mean reduction for (a) broadband, (b) narrowband, and (c) narrowband - broadband over all calibration gradients.

The improvements relative to sl014 are illustrated in Figure 3.11. During the initial session, chirp-based calibrations enhanced most FFE sequences in the channel-averaged broadband analysis. Given sequence 7's lower reduction discussed previously, its specific improvements should be interpreted alongside the repeat-session results (Table 3.1). Excluding this outlier, sequence 5 with chirp 1 yielded the largest improvements in both the channel-averaged broadband and narrowband analyses, increasing reductions by 3.72 dB and 4.56 dB, respectively.

The channel-specific heatmaps showed larger calibration-dependent changes in CH1 than in CH2. Excluding sequence 7, CH1 still showed improvements above 5 dB for sequence 5 across all chirp calibrations. In CH2, the improvements were generally smaller and more consistent across sequences.

Negative differences relative to sl014 were observed in selected cases. In CH1, negative broadband differences occurred for sequence 9 across all chirp calibrations, with the largest decrease for chirp 2. In the channel-averaged broadband analysis, negative differences were limited to chirp 2 for sequences 1, 4, and 9.

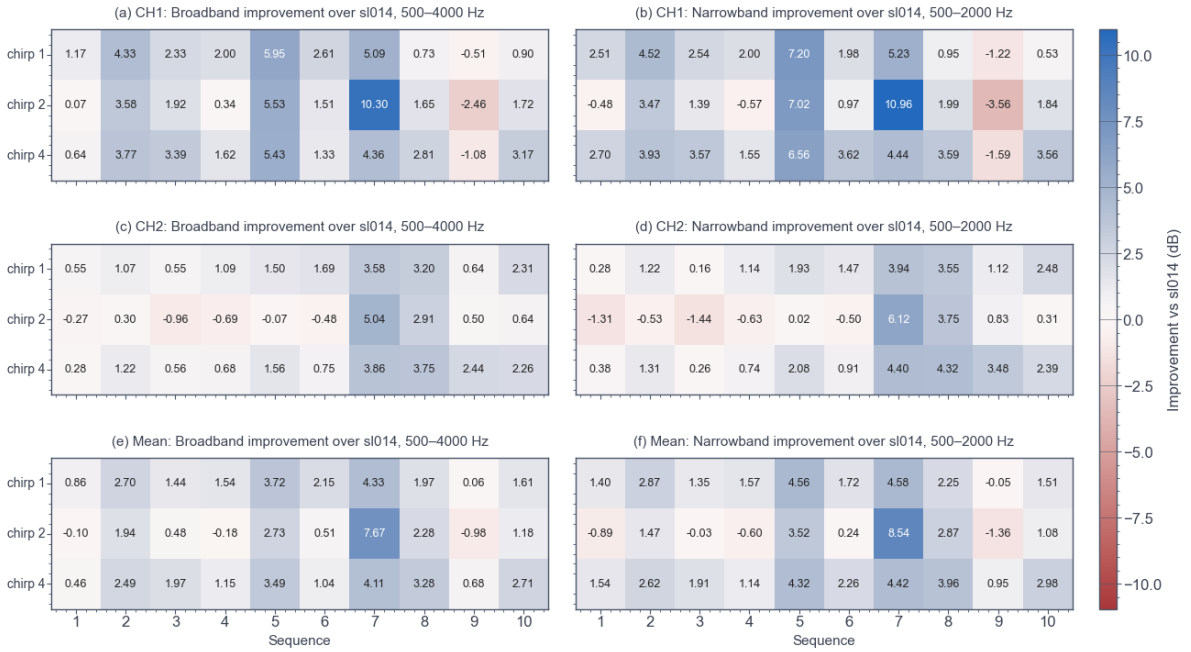


Figure 3.11: FFE mean improvement over sl014 for (a) broadband and (b) narrowband reduction.

To examine whether adding the EPI-Y component affected regular sequence performance, chirp 4 and the combined chirp 4+EPI-Y derived TF were compared directly in a separate session across the ten FFE sequence variants. As shown in Figure 3.12, the two approaches produced similar reduction values across sequences. Across all sequences, the mean broadband reduction was 9.26±1.13 dB for chirp 4 and 9.16±1.25 dB for chirp 4+EPI-Y, while the mean narrowband reduction was 9.94±1.19 dB and 9.78±1.19 dB, respectively. The combined TF produced slightly higher reduction in some sequences and slightly lower reduction in others, but the average differences were small and did not show a consistent direction across sequences.

The temporal and spectral reduction results across the FFE sequence variants for the chirp 4 calibration are shown in Figure 3.13. Chirp 4 was selected for this representative example because it produced the highest mean reduction and overall equalizer filter metrics in the initial full calibration comparison. Visible reduction of the scanner acoustic waveform was observed across the tested sequences. In the temporal plots, the residual waveform amplitude and envelope differed between sequences and channels. The corresponding spectra showed frequency-dependent residual components after cancellation. For several sequences, the strongest visible suppression occurred around the dominant spectral components below approximately 1.5 kHz.

The residual patterns differed across sequences. Some sequences showed residual components distributed across a broader frequency range, while others showed residuals concentrated near the lower-frequency boundary of the analysis band, especially around 0.5–0.6 kHz. These lower-boundary residuals were particularly visible in sequences 1, 7, and 9. In sequence 6, residual components were more prominent toward the upper part of the broadband analysis range.

3.3.2.2. EPI Sequence Evaluation

The EPI evaluation focused on repeated measurements of the 10- and 33-dynamics variants. Compared with the FFE measurements, the EPI residual spectra showed more prominent narrow spectral peaks. The repeated EPI live reduction results are summarized in Figure 3.14. Across both EPI variants, the combined TFs generally produced higher broadband reduction than chirp 4.

For the 10-dynamics EPI, the mean broadband reduction across CH1 and CH2 started at 8.40 dB

Figure 3.12: FFE mean reduction comparison for (a) broadband, (b) narrowband, and (c) narrowband - broadband chirp 4 vs. chirp 4 + EPI-Y calibration gradient.

for chirp 1+EPI-Y, 8.59 dB for chirp 4+EPI-Y, and 7.56 dB for chirp 4. Across the five repetitions, all three approaches showed a gradual decrease. By the fifth repetition, the mean broadband reduction was 7.31 dB for chirp 1+EPI-Y, 7.35 dB for chirp 4+EPI-Y, and 6.53 dB for chirp 4. The two combined TFs showed similar reduction levels across repetitions, while chirp 4 remained lower.

For the 33-dynamics EPI, the difference between chirp 4 and the combined TFs was larger. The mean broadband reduction started at 7.43 dB for chirp 1+EPI-Y, 7.57 dB for chirp 4+EPI-Y, and 6.95 dB for chirp 4. By the fifth repetition, these values decreased to 5.80 dB, 5.76 dB, and 5.14 dB, respectively. Again, chirp 1+EPI-Y and chirp 4+EPI-Y showed similar reduction trends, while chirp 4 gave the lowest reduction.

The residual level in the highlighted EPI-dominant band between 750 and 950 Hz is also shown in Figure 3.14. For both EPI variants, the residual peak increased across repetitions for all calibration approaches. However, the combined TFs maintained lower residual peaks than chirp 4. This difference was especially visible in CH2. In the 33-dynamics EPI CH2 measurement, for example, the residual peak increased from 3.78 mV to 5.17 mV for chirp 4+EPI-Y and from 3.87 mV to 5.30 mV for chirp 1+EPI-Y, while chirp 4 increased from 5.57 mV to 7.47 mV. A similar but smaller separation was also present in the 10-dynamics measurement.

To further examine the EPI reduction behavior, repeated noise-only and anti-noise-only recordings were inspected for the EPI 10-dynamics condition, as shown in Figure 3.15. To contextualize the repetition trend, runs 1–5 were acquired consecutively without an intentional gap, followed by an approximately 3-minute gap before run 6, a 5-minute gap before run 7, and a 10-minute gap before run 8. The broadband and narrowband RMS values remained relatively stable throughout the repetitions for both recordings. Across all runs, the anti-noise-only recording had lower broadband and narrowband RMS values than the noise-only recording. The estimated reduction, calculated from the noise-only and anti-noise-only recordings, decreased across the consecutive runs 1–5 and increased again after the pause before run 6. By run 8, the estimated reduction had returned close to its initial value at run 1.

The EPI strong-band peak magnitude showed a different pattern from the RMS-based comparison. Although the anti-noise-only recording had lower overall RMS values, it showed a consistently higher peak magnitude near the EPI-dominant band than the noise-only recording. Within this EPI-dominant band, the noise-only peak magnitude showed larger variation across repetitions than the anti-noise-only peak.

Figure 3.13: FFE temporal and spectral broadband reduction for chirp 4 gradient derived TF.

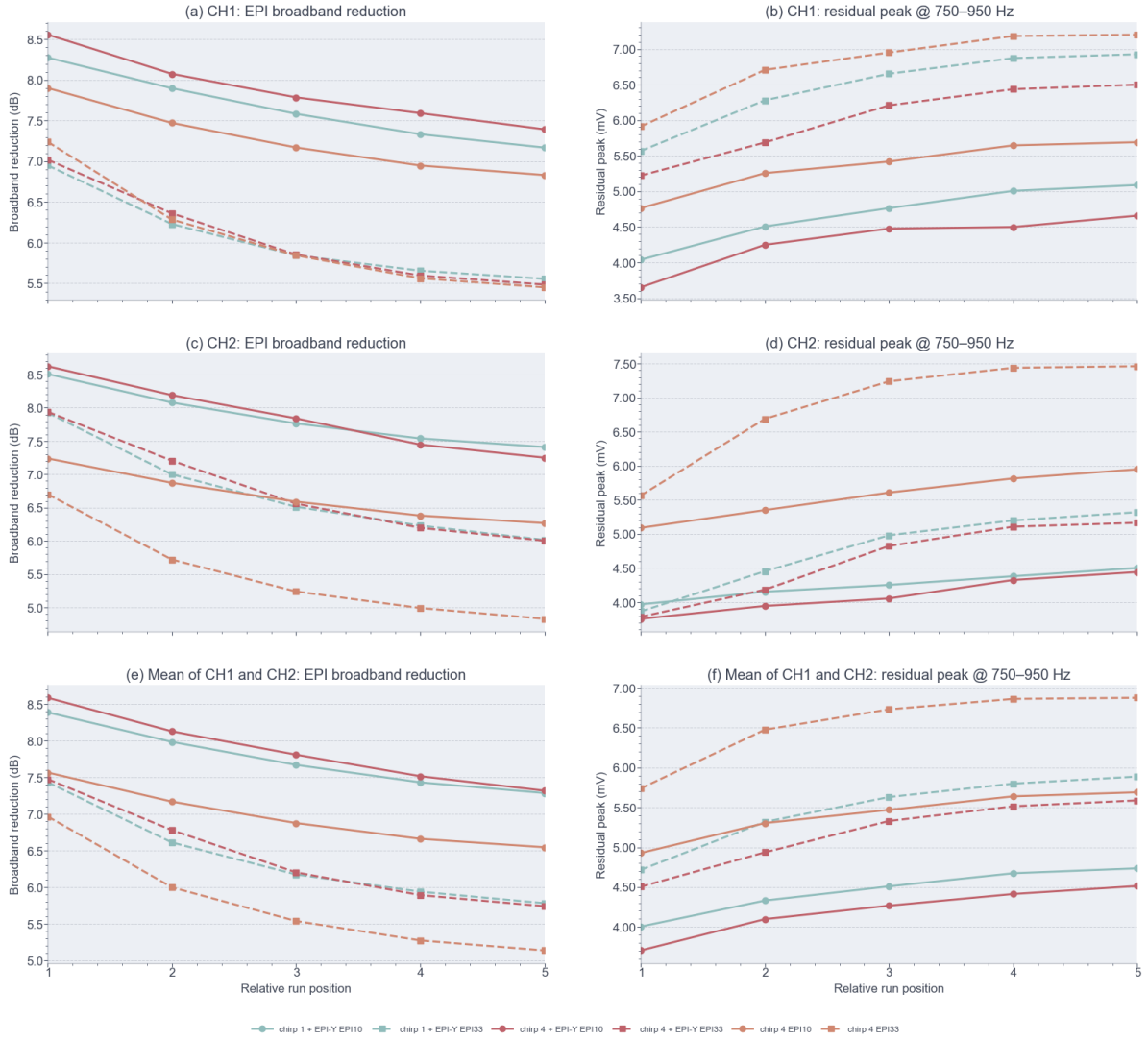


Figure 3.14: Repeated EPI live reduction for chirp 1+EPI-Y, chirp 4+EPI-Y, and chirp 4 TFs. Broadband reduction is shown for CH1, CH2, and the channel mean in panels (a), (c), and (e), while the corresponding residual amplitudes near the dominant EPI spectral component around 825 Hz are shown in panels (b), (d), and (f). Solid and dashed lines indicate EPI10 and EPI33, respectively.

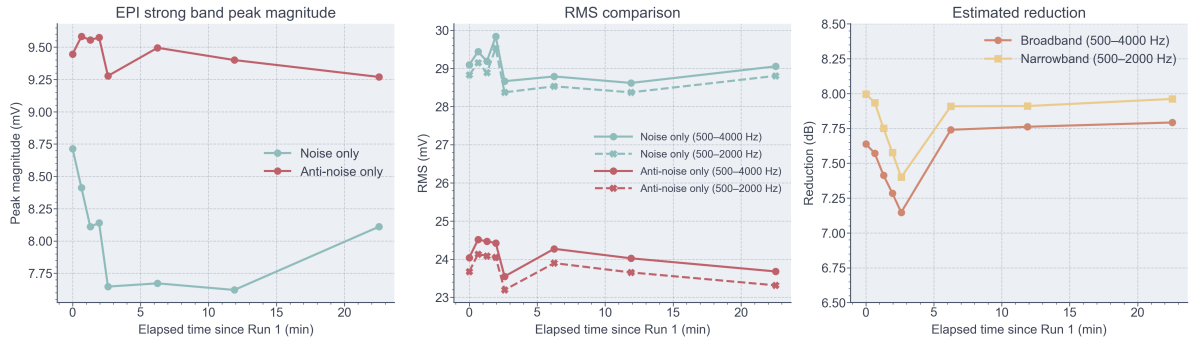
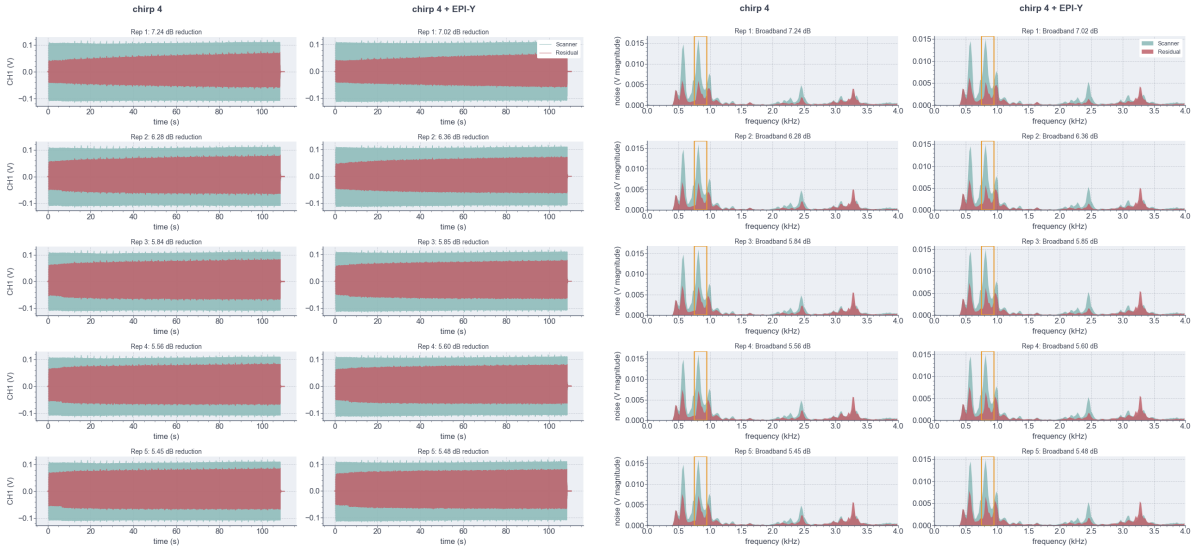


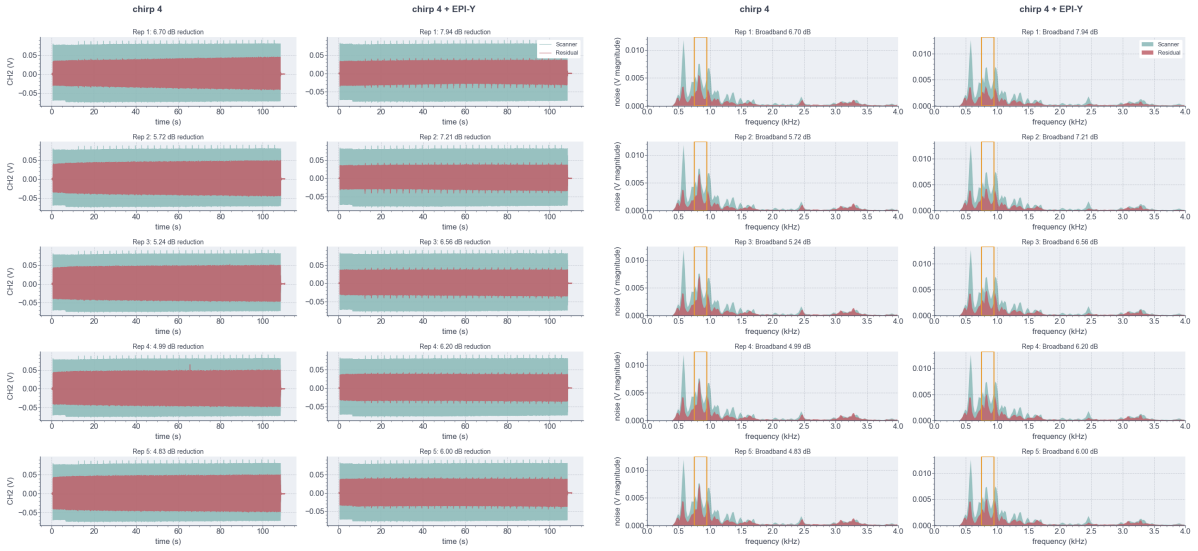
Figure 3.15: Repeated EPI noise-only and anti-noise-only metrics.

The temporal and spectral EPI results for the 33-dynamics measurement are shown in Figure 3.16 for CH1 and Figure 3.17 for CH2. Visible residual structure remained after cancellation for both chirp 4 and chirp 4+EPI-Y. In CH2, the temporal comparison showed a clearer difference between the two TFs. With chirp 4 alone, the residual amplitude became larger across time and repetitions, while the chirp 4+EPI-Y residual appeared more stable over time. This difference was less visible in CH1. The spectral plots showed a similar pattern, with chirp 4+EPI-Y producing lower residual magnitude near the highlighted EPI-dominant compared to chirp 4, especially in CH2. Additional residual peaks were observed around 1.65 kHz, approximately twice the frequency of the dominant EPI spectral component.



(a) Temporal EPI reduction. **(b)** Spectral EPI reduction.

Figure 3.16: EPI 33 CH1 dynamics live reduction over five repetitions with chirp 4 and chirp 4+EPI-Y derived TFs.



(a) Temporal EPI reduction. **(b)** Spectral EPI reduction.

Figure 3.17: EPI 33 CH2 dynamics live reduction over five repetitions with chirp 4 and chirp 4+EPI-Y derived TFs.

4

Discussion

This thesis investigated the redevelopment, validation, and evaluation of a PNC system for MRI. The results showed that the original LabVIEW/MATLAB-based PNC workflow could be successfully redeveloped in Python, while preserving the main DSP and workflow behavior of the original implementation and improving computational efficiency. The phantom acoustic experiments further showed that PNC performance was strongly affected by calibration quality, sequence-specific acoustic behavior, playback stability, timing accuracy, and channel-specific acoustic coupling. This chapter discusses the findings in relation to the software redevelopment, acoustic performance, and remaining limitations of the system for reliable noise cancellation during MRI phantom experiments.

4.1. Pipeline Redevelopment and Validation

Although the original PNC system was functional, it was difficult to maintain and extend because of its restrictive architecture. The workflow relied on LabVIEW for experiment control, audio recording, and AFG communication, while MATLAB was used for DSP. This dependency reduced openness and portability, and made debugging less straightforward because the logic was distributed across two separate proprietary environments. The LabVIEW workflow was also prone to instability due to high memory use and buffer-related issues when audio recording, workflow execution, and hardware communication were handled concurrently. This concurrent execution appeared to contribute to race-condition-like behavior, including hard crashes during audio recording and computationally intensive function calls. This was a critical practical limitation because it interrupted experiments and wasted allocated scanning time. In addition, MATLAB functions called from LabVIEW were slow, particularly when processing longer audio arrays.

The Python redevelopment addressed these limitations by consolidating the full PNC workflow into a single, modular framework. Calibration and regular sequence experiments were made accessible from the same interface, with monitoring information available during experiment execution. This integration improved workflow continuity because calibration outputs could be reused directly in regular sequence experiments, sequence settings were easier to inspect and modify, and experiment execution became less dependent on manual switching between tools. The single-language implementation also simplified debugging and maintenance, while the modular structure enabled targeted improvements such as robust clock drift estimation, combined TF generation, and sequence-specific calibration extensions.

Calibration updates improved robustness.
Several targeted calibration updates were added to improve robustness and level control. An important change was the use of a more robust clock drift estimation during calibration. Instead of relying on ordinary linear fitting, which could be affected by outlier delay estimates from noisy or shifted correlation peaks, the Python workflow used a median-based approach that was less sensitive to occasional poor delay estimates. This was important because clock drift affects resampling, anti-noise timing, downstream alignment, and ultimately the achieved cancellation. For longer experimental durations, recalibrating clock drift during regular sequence experiments may therefore be needed to maintain accurate anti-noise generation.

The added I/O scale derivations also improved calibration reliability. The initial I/O scale provided a channel-specific mapping between the commanded playback level and recorded microphone level, helping the first equalization iteration start at a level high enough for a reliable response while remaining safely below the playback limit. After equalization, the equalized I/O scale accounted for level changes introduced by the EQ filter and provided a more meaningful anchor between digital playback level and measured sound level at the microphone position. This information should make it easier in future fMRI experiments to present sounds at a desired level in the scanner environment.

Python preserved DSP behavior while reducing runtime.
The rebuilt Python system reproduced the main behavior of the original implementation while substantially reducing computation time. At the DSP level, Python outputs closely matched the MATLAB reference, particularly for deterministic operations such as filtering, signal shifting, resampling, and drift correction. These waveform comparisons also confirmed that the Python implementation did not introduce an additional sample lag relative to the MATLAB reference. This timing agreement is important for PNC because any implementation-induced delay in the generated anti-noise would appear during playback as a phase mismatch relative to the scanner noise, reducing cancellation performance. The preprocessing and filtering functions showed the largest speed-up gains, while more complex calibration and regular sequence processing showed more modest improvements because these stages combine several DSP operations.

Larger numerical discrepancies were observed for equalizer design and TF derivation, where some SER values were lower than in other processing stages. This was expected because these stages involve filter coefficients, interpolation, spectral-domain operations, regularization, and boundary handling, which are more sensitive to small implementation differences than simpler deterministic operations.

The lower SER values in intermediate stages did not necessarily indicate poor agreement. In the PNC workflow, intermediate parameters such as averaged calibration responses and derived TFs are used to generate anti-noise, and the relevance of their numerical differences depends on whether they affect the downstream anti-noise signal and reduction metrics. The calibration-chain validation showed that later calibration outputs remained close to the MATLAB reference, even when some intermediate outputs were not numerically identical. Estimated and simulated reduction metrics based on these DSP outputs were also similar to those obtained with MATLAB. This suggests that the Python implementation preserved the functional behavior of the DSP chain sufficiently for anti-noise generation. In several cases, Python produced slightly higher reduction metrics than MATLAB, but the differences were small and should be interpreted as comparable performance.

The computational speed-up has direct experimental relevance. DSP functions are repeatedly called throughout the PNC workflow, so function-level improvements also reduce the overall execution time. This is especially beneficial during MRI experiments, where scanner time is constrained by patient comfort and safety, allocated MRI schedules, and the need for repeated calibration, troubleshooting, and parameter testing. The reduced processing time allowed additional calibration attempts within the same scanning session, providing more opportunities to optimize parameters or recover from recording issues without extending scan duration. At the integrated-workflow level, the Python implementation reproduced the calibration step 1–3 results and the regular sequence outputs while reducing total execution time by approximately 73%. The rebuilt workflow also improved operational reliability during audio recording, especially when recording was performed together with AFG communication, where the original LabVIEW implementation was prone to crashes.

Workflow validation supported functional equivalence.
Beyond the matched-input DSP comparison, the integrated workflow validation tested whether the rebuilt Python implementation reproduced the behavior of the original PNC workflow during calibration-chain and regular sequence processing. The matched-input DSP validation isolated implementation differences between MATLAB and Python, while the integrated workflow validation tested whether these differences affected the full PNC workflow. Across the evaluated workflow blocks, the errors remained below 5%, except for the flagged regular sequence step 1 case. The downstream step 2 and step 3 results remained closely matched between implementations, consistent with the downstream agreement observed in the DSP validation. These findings show that the rebuilt Python pipeline was sufficiently accurate for the tested PNC workflows.

However, workflow equivalence does not guarantee acoustic robustness across all scanner conditions. Final acoustic reduction also depends on playback-chain stability, triggering, clock drift, equalization quality, and headphone or microphone coupling. The workflow validation therefore provides evidence that the rebuilt Python implementation reproduced the original workflow for the tested regular sequence, but it was not intended to cover every possible MRI sequence. Sequences with different acoustic spectra, timing structure, or dominant residual peaks may still produce different PNC behavior, as observed later in the differences between FFE and EPI measurements.

The main exception in the workflow comparison was the regular sequence step 1 result. The original implementation produced a negative reduction, while the Python implementation produced a positive reduction that was more consistent with the subsequent step 2 and step 3 behavior. This discrepancy likely reflects sensitivity in the original step 1 calculation to alignment, window selection, or scaling. Because step 1 estimates idealized predicted reduction by aligning the predicted and recorded acoustic trigger responses, it can become unreliable when the trigger prediction has low SNR or when the alignment peak is ambiguous.

This discrepancy did not extend to later processing steps, which are more relevant for live reduction because they include residual latency correction and live anti-noise evaluation. The result therefore suggests that the original step 1 alignment procedure for the regular sequence was not fully robust in this validation case, while the Python implementation produced a step 1 result more consistent with later workflow output. However, this conclusion is based on a single representative validation example and should not be interpreted as evidence of greater Python robustness across all sequence types.

Channel imbalance reflected hardware-path limitations.
The validation further identified slightly greater deviations in CH1 compared to CH2 for several parameters, particularly in I/O scaling and regular sequence processing. While these differences remained within acceptable limits, they suggest that the two channels may exhibit varying sensitivity to implementation details or hardware characteristics. This channel-specific behavior became more apparent during later acoustic evaluation, where CH1 and CH2 showed distinct performance patterns.

The channel imbalance observed during calibration is unlikely to be explained primarily by headphone fit or positioning. Changing the head phantom orientation or headphone placement affected the reduction, but did not reproduce the large CH1–CH2 difference observed in the main calibration results. Thus, the weaker CH1 performance more likely originated from a channel-specific hardware or acoustic-transfer limitation, such as the headphone driver, playback path, cable connection, or channel-specific equalization. Based on the available observations, the headphone driver is a plausible contributor, although this was not isolated experimentally.

Accurate amplitude and phase matching in each playback path is critical for PNC performance. If one channel performs poorly, it can limit live cancellation even when the software implementation and TF estimation are working as intended. Accordingly, the CH1 behavior should be regarded as a technical limitation of the current hardware setup, and not a problem with the rebuilt Python pipeline itself. This interpretation is supported by the observation that a similar channel imbalance was also present in the original PNC program during the integrated workflow comparison.

Overall, the redevelopment successfully reproduced the main behavior of the original PNC workflow while providing substantial improvements in speed, stability, and workflow traceability. These improvements made the system more suitable for repeated MRI scanner experiments and future development. However, the remaining discrepancies show that software equivalence by itself is not sufficient to guarantee acoustic performance, because final PNC reduction also depends on alignment, timing accuracy, recording stability, playback stability, and channel-specific hardware behavior.

4.2. Calibration Gradient Redesign Improves PNC Performance

The comparison of calibration approaches demonstrated that the original sl014 calibration was less robust than the redesigned chirp-based approaches, while the combined-TF approaches provided additional sequence-targeted calibration options. The sl014 calibration exhibited weaker performance across multiple calibration quality metrics, including TF variability, coherence, live step 3 reduction, and worst-axis-channel reduction.

Broadband chirps improved calibration reliability.
The improvement achieved with the broadband chirps can be partly explained by the step 1 excitation analysis. The sl014 calibration produced lower recorded acoustic levels and weaker spectral coverage across the calibration band. This likely decreased the effective SNR of the TF estimation. Under these conditions, the estimated TF becomes more susceptible to background noise, spectral leakage, alignment uncertainty, and regularization effects. This may explain the higher TF variability and weaker live reduction observed with sl014. In contrast, the chirp-based gradients produced stronger recorded excitation, which generally improved the stability of the estimated TF and led to stronger live acoustic reduction.

Although the broadband chirps improved calibration reliability overall, stronger excitation was not the sole determinant of calibration performance. The results show that gradient scaling by itself was not sufficient to predict calibration quality. In the first calibration session, chirp 2 achieved the highest TF cross-validation SER, indicating strong repeatability of the estimated TF across repeated calibration measurements. However, this did not translate into the strongest live step 3 reduction or worst-axis-channel reduction, where chirp 1 performed better. Chirp 1 also showed lower complex TF variability and higher coherence than chirp 2 in the first session, suggesting that the best calibration setting depends on which aspect of TF quality is being evaluated.

Chirp 4 was then designed to broaden spectral coverage from 150–4500 Hz to 100–5000 Hz. In the second calibration session, it achieved the highest TF cross-validation SER and strongest worst-axis-channel reduction among the single-gradient candidates, consistent with improved TF repeatability and worst-case robustness. However, chirp 4 did not outperform chirp 1 across all metrics, indicating that broader excitation provided selective benefits across the PNC pipeline.

TF metrics did not fully predict live reduction.
A key outcome from the calibration evaluation was that TF-domain metrics did not consistently predict live cancellation performance. In the initial calibration comparison, chirp 2 achieved the highest TF cross-validation SER, but chirp 1 achieved stronger live step 3 reduction. Similarly, in the second calibration comparison, chirp 4 achieved the highest TF cross-validation SER and strongest worst-axis-channel reduction, while chirp 1 still produced slightly stronger live step 3 reduction.

This mismatch likely occurs because step 3 evaluates the full PNC chain, not just TF repeatability. Live reduction also depends on anti-noise generation, equalization, playback through the AFG and amplifier, headphone output, latency correction, acoustic coupling, and microphone recording. Calibration quality should therefore be assessed using both model-level repeatability and live acoustic reduction, so that the selected setting reflects the best overall system behavior.

For this reason, chirp 1 and chirp 4 were both retained as useful calibration candidates. Chirp 1 showed the strongest live reduction among the single-gradient candidates in both calibration evaluations, making it a reasonable choice when the main priority is end-to-end acoustic cancellation. Chirp 4, by contrast, achieved strong TF-domain repeatability and the strongest worst-axis-channel reduction in the second comparison, supporting it as a robust general-purpose calibration option when consistency across axes and channels is important. These results suggest that chirp 1 is preferable when live reduction performance is the primary criterion, while chirp 4 may be preferable when TF repeatability and worst-case axis-channel robustness are prioritized.

Combined TFs added targeted EPI information.
Unlike the broadband chirp gradients, the combined TFs were generated by combining information from separate calibration measurements instead of using a single excitation waveform. The EPI-Y component was designed as a targeted, sequence-specific excitation, not as a general broadband calibration waveform. In the calibration evaluation, the combined TFs were assessed using the corresponding chirp reference recordings, because the combined estimates were derived mathematically from separate calibration measurements. Therefore, the step 3 reduction for a combined TF reflects how the combined estimate performed relative to its chirp reference condition, rather than the response to a physically combined chirp+EPI-Y excitation.

The main purpose of adding EPI-Y was to increase TF information around EPI-relevant spectral components. The EPI-Y calibration response showed that the recorded acoustic spectrum was more complex than the differentiated gradient input spectrum. Although the waveform targeted the dominant EPI-related gradient frequencies, the measured response also contained harmonic-related com-

ponents. This indicates that the recorded acoustic response was not a one-to-one reflection of the differentiated gradient spectrum. These harmonic-related components may be visible in the EPI-Y-only acoustic response, but they may be underrepresented in the combined TF because the regularized least-squares estimator weights each calibration measurement by its differentiated-gradient input energy. In particular, response components that are strong in $P_{\text{EPI-Y}}(f)$ but weak in $G_{\text{EPI-Y}}(f)$ may contribute less to the combined estimate after regularization. This may explain why the combined-TF calibration improved TF information near the explicitly excited EPI-relevant frequencies, while not necessarily preserving all harmonic-related acoustic components observed in the EPI-Y response.

The combined chirp 1+EPI-Y derived TF showed particularly strong TF-domain stability, with low magnitude variability, low complex TF variability, and high coherence. Its step 3 live reduction remained comparable to chirp 1, indicating that adding the targeted EPI-Y component did not substantially compromise calibration-stage reduction. For this reason, chirp 1+EPI-Y remained an important EPI-targeted combined-TF condition for later sequence-specific evaluation.

The chirp 4+EPI-Y combined TF did not demonstrate the same calibration-stage advantage over its broadband chirp baseline as chirp 1+EPI-Y. Although chirp 4+EPI-Y maintained strong TF-domain performance, it did not improve mean step 3 reduction or worst-axis-channel reduction compared with chirp 4. However, chirp 4+EPI-Y was still retained for later evaluation because its TF-domain metrics remained strong and because it represented the combination of one of the most robust general-purpose chirp candidates with targeted EPI-Y information. This allowed the later regular-sequence and EPI evaluations to test whether EPI-specific TF information could provide sequence-dependent benefits even when calibration-stage live reduction was not improved.

Channel robustness affected calibration selection.
The channel imbalance observed during calibration shaped how the results should be interpreted. Axis-channel analyses showed that some calibration approaches improved one channel more than the other. This meant that channel-averaged mean reduction alone was insufficient, because it could mask uneven performance across channels. For a more complete picture, mean step 3 reduction, worst-axis-channel reduction, and axis-channel-specific reduction were evaluated collectively.

Overall, the calibration evaluation did not identify a single approach that was superior across all metrics. Instead, it supported retaining complementary candidates for later testing. Chirp 1 was retained as the primary live-reduction candidate, while chirp 4 was retained as the more robust broadband candidate. The chirp 1+EPI-Y and chirp 4+EPI-Y combined TFs were retained to test whether targeted EPI-Y information could provide sequence-dependent benefits beyond the broadband chirp calibrations. These findings indicate that calibration gradient design affects both TF stability and live PNC performance, requiring final calibration selection to balance spectral coverage, TF repeatability, channel robustness, and end-to-end acoustic reduction.

4.3. Regular MRI Sequence Evaluation

The regular sequence evaluation was used to test whether the calibration improvements translated into acoustic reduction during MRI sequence playback. Two sequence groups were considered. Short FFE variants were used as representative regular MRI sequences with relatively broadband acoustic behavior, while EPI was used as a peak-dominated and temporally demanding sequence relevant for future auditory fMRI applications.

4.3.1. Representative FFE Performance with Redesigned Broadband Calibration

The short FFE phantom evaluation showed that the redesigned broadband chirp calibrations improved regular sequence PNC performance compared with the original sl014 calibration. This improvement was observed in both broadband and narrowband analyses, although the magnitude and consistency of the improvement varied across calibration designs and measurement sessions. Across the complete FFE dataset, the redesigned chirp calibrations generally outperformed sl014, while chirp 2 showed less consistent behavior across sessions. The following analysis therefore focuses on why chirp 1 and chirp 4 emerged as the most reliable broadband calibration candidates for representative FFE sequences.

Chirp 1 and chirp 4 were the strongest FFE candidates.
Chirp 1 and chirp 4 provided the most reliable FFE reductions, although neither chirp showed a session-independent advantage over the other. In the initial session, chirp 4 produced the highest mean reduction, while chirp 1 produced the second-highest values. In both repeated sessions, this pattern was reversed, with chirp 1 producing the highest mean reduction and chirp 4 producing the second-highest values. However, the difference between chirp 1 and chirp 4 was small in all sessions compared with the larger separation from sl014 and chirp 2. This suggests that both chirps provided similarly reliable FFE performance, with the preferred option depending on session conditions.

One initial-session measurement required separate consideration. Sequence 7 showed especially low reduction with sl014 and therefore contributed to the lower overall sl014 average in that session. This low value may reflect a position-sensitive stress case, because the phantom had been repositioned before the scan despite attempts to reproduce the same position using guide points. To check whether the overall improvement depended mainly on this case, the averages were recalculated with sequence 7 excluded. The redesigned calibrations still performed better than sl014, supporting that the observed improvement was not driven solely by this low-reduction measurement.

Session conditions affected the preferred chirp.
The reversal between chirp 1 and chirp 4 likely reflected session-dependent playback-chain conditions, especially headphone coupling and equalizer filter quality. The initial session had the highest equalizer filter reduction quality metric and showed less prominent residual components near the analysis-band boundaries. Under these conditions, chirp 4 produced the highest FFE reduction. Because much of the FFE acoustic energy was concentrated toward the lower end of the analysis band, accurate anti-noise reproduction near the 500 Hz boundary was likely important for the overall broadband and narrowband reduction. A better equalized playback path may therefore have allowed chirp 4 to take greater advantage of its flatter and wider excitation profile, producing more balanced TF estimates across the full analysis band, including the low-frequency region that strongly contributed to the measured reduction.

Chirp 2 followed the same session-dependent pattern. It performed competitively in the initial session, where equalizer quality was highest, but its reduction decreased substantially in the repeated sessions. Although chirp 2 had higher excitation amplitude than chirp 4, its narrower spectral coverage may have made it more vulnerable when the equalizer response was less well controlled near the analysis-band boundaries. The repeated-session results indicate that excitation strength was insufficient without robust coverage of the full relevant frequency range. Both excitation amplitude and spectral coverage likely contributed to TF quality, with their relative importance depending on the playback-chain conditions of each session.

In the repeated sessions, where the equalizer filter reduction quality metric was lower and boundary residuals near 500 Hz were more prominent, chirp 1 produced the highest mean reduction. Its stronger excitation at higher frequencies within the analysis band may have provided more reliable TF estimation and cancellation in regions less affected by the low-frequency boundary. This does not imply that chirp 1 reduced the low-frequency residual more effectively. More likely, chirp 1 achieved stronger suppression in other reproducible spectral components, particularly at higher frequencies where its excitation energy was greater, which compensated for the reduced cancellation near 500 Hz. This compensation may explain why chirp 1 performed better when the low-frequency boundary was less clearly equalized.

The complementary behavior of chirp 1 and chirp 4 reflects a trade-off between excitation strength and spectral uniformity. Chirp 4 has lower raw excitation strength, but provides a flatter and more extended calibration spectrum. This appears advantageous when the playback system is well equalized and can reproduce anti-noise accurately across the full analysis band. Chirp 1 provides stronger excitation, but with less spectral uniformity. This may make it more robust when equalization quality is reduced or when the playback path is less reliable near the lower-frequency boundary. For broadband PNC, spectral flatness and excitation strength therefore appear to act as complementary calibration properties. The best calibration choice depends on the calibration gradient together with the headphone fitting, equalizer quality, and frequency-dependent reproducibility of the playback chain in a given session. Thus, chirp 4 may be preferred under well-controlled playback conditions, whereas chirp 1 may be more practical when equalization quality is less uniform.

Residuals were frequency- and sequence-dependent.
The difference between broadband and narrowband performance further shows that PNC performance

depended on the analysis band. Narrowband reduction was usually higher than broadband reduction, indicating that the strongest cancellation occurred in the frequency range where the dominant FFE acoustic components were concentrated. This interpretation agrees with the spectral plots, which showed the strongest visible suppression below approximately 1.5 kHz. However, sequence-dependent behavior was still present. Some sequences responded better to specific calibration TFs, suggesting that TF accuracy at individual spectral peaks, phase response, timing mismatch, and channel-dependent acoustic coupling also contributed to the final reduction.

The sequence-wise results showed that the residual pattern differed between FFE variants. In the chirp 4 temporal and spectral examples, chirp 4 generally reduced the waveform amplitude and suppressed the dominant spectral peaks, but the residual energy was not distributed in the same way for every sequence. Some sequences retained broader residual components, while others showed residuals concentrated near the lower-frequency boundary. Sequence 6 showed a different pattern, with more residual energy in the upper part of the broadband analysis range. These differences show that the final reduction still depended on where each sequence's acoustic energy fell within the analysis band and how accurately those frequency regions were characterized during calibration and reproduced by the playback path.

Sequence 6 provides a useful example of this calibration-dependent behavior. Because its residual energy was concentrated at higher frequencies, the higher reduction obtained with chirp 1 for this sequence is consistent with the stronger high-frequency excitation of that calibration gradient. The two chirps therefore appear to have achieved strong FFE reduction through different mechanisms. Chirp 4 relied on broader and flatter spectral characterization, while chirp 1 relied more on stronger excitation in frequency regions that may remain reproducible when low-frequency equalization is less reliable.

This behavior is consistent with the physical requirement of PNC, where the anti-noise must match the scanner noise in both amplitude and phase at the dominant spectral components of each sequence. Because these spectral components differ across FFE variants, no broadband calibration TF is expected to be uniformly optimal for every sequence. Persistent low-frequency residuals around 0.5–0.6 kHz likely limited the achievable reduction in several FFE measurements. This frequency region is near the lower edge of the analysis band and close to the low-frequency limit of the headphone equalization and anti-noise playback. Even when the calibration TF estimated the scanner response in this region, the playback system may not have reproduced the required anti-noise amplitude and phase accurately near the boundary. As a result, some remaining residuals likely reflect low-frequency playback-chain and control limitations in addition to TF-estimation differences.

EPI-Y information did not compromise FFE reduction.
The separate chirp 4 versus chirp 4+EPI-Y comparison showed that adding the EPI-Y component did not substantially change FFE reduction. The two approaches produced similar reduction values across the FFE variants, with chirp 4+EPI-Y performing slightly better in some sequences and slightly worse in others. Because this comparison was acquired in a separate measurement context from the main FFE calibration comparison, the absolute values are not directly interchangeable with the main chirp-gradient results. Nevertheless, the result indicates that adding targeted EPI-Y information did not introduce an obvious broadband-reduction penalty for the tested chirp 4-based TF. This supports carrying the combined-TF approach forward for later EPI-focused evaluation, including both chirp 4+EPI-Y and chirp 1+EPI-Y, while retaining the corresponding broadband chirps as reference conditions.

Channel imbalance limited mean FFE reduction.
The FFE results also showed clear differences between channels. Calibration-related changes were more noticeable in CH1 than in CH2. CH1 showed larger reductions in performance with sl014 and larger improvements with the redesigned calibration gradients, while CH2 remained more stable across methods. This pattern is consistent with the channel imbalance observed during calibration and the headphone-positioning checks. Since PNC requires the anti-noise signal to match the scanner noise in both amplitude and phase at the measurement point, any channel-specific difference in the headphone path, acoustic coupling, or microphone response can limit the achievable cancellation, even when the TF estimate itself is stable.

Overall, the FFE evaluation supports chirp 1 and chirp 4 as the preferred broadband calibration options for representative FFE sequences. Both provided more reliable reduction than sl014 and chirp

2 across the three sessions, although their relative performance changed with session-specific equalization quality. The sequence-wise and channel-wise residuals show that calibration design was only one part of the achieved reduction. Residual low-frequency components, spectral overlap between the sequence noise and calibration response, channel-dependent acoustic coupling, and setup sensitivity all contributed to the final cancellation. Thus, the redesigned broadband calibrations improved regular-sequence PNC performance, but the playback chain and channel-specific acoustic path still limited the maximum reduction.

4.3.2. EPI Performance with Sequence-Specific Combined TFs

The EPI results showed a different limitation than the FFE measurements. While the FFE sequences were generally improved by redesigned broadband chirp calibration, EPI performance benefited from adding targeted EPI-Y information to the broadband TF. Both chirp 1+EPI-Y and chirp 4+EPI-Y produced higher broadband reduction and lower residual amplitudes near the dominant EPI spectral component than chirp-only calibration. This agrees with the motivation for the combined TF approach. Because EPI contains strong sequence-specific spectral peaks, accurate cancellation requires the anti-noise to match both the amplitude and phase of those peaks.

EPI-Y information drove the combined-TF improvement.
The similar reduction achieved by chirp 1+EPI-Y and chirp 4+EPI-Y indicates that the added EPI-Y information was the main contributor to the EPI improvement. In this sequence, the acoustic energy was concentrated around a dominant spectral region, so explicitly including EPI-Y information gave the TF better amplitude and phase information at the frequencies most important for cancellation. As a result, the difference between the two combined TFs was smaller than the difference between chirp-only and chirp+EPI-Y calibration. Chirp 4+EPI-Y therefore should not be treated as superior simply because chirp 4 performed best in the initial FFE session. For EPI, chirp 1 and chirp 4 acted as comparable broadband baselines, while the EPI-Y component provided the sequence-specific information needed to reduce the dominant peak.

EPI improvement remained channel-dependent.
The improvement from the combined TFs was more apparent in CH2 than in CH1. In CH2, the combined TFs produced lower residual amplitudes near the EPI-dominant band than chirp-only calibration, especially in the 33-dynamics EPI measurement. This channel dependence matches the broader CH1–CH2 imbalance observed during calibration and FFE evaluation. It indicates that adding more relevant TF information can improve EPI cancellation, but the final reduction still depends on the channel-specific playback path, headphone coupling, and acoustic transfer to the measurement position.

Peak-dominated residuals limited broadband calibration.
The EPI results also show why broadband calibration alone was insufficient for this sequence type. Chirp 4 provided the strongest broadband chirp baseline, but its chirp-only TF did not reduce the dominant EPI residual as effectively as the combined TFs. The improvement observed after adding EPI-Y information indicates that broader chirp coverage alone was not enough and that the calibration also needed sequence-specific information near the dominant EPI spectral component. Unlike the FFE measurements, where the acoustic energy was spread more broadly across the analysis band, EPI energy was concentrated around narrow dominant peaks. This made the reduction more sensitive to small amplitude or phase errors at those specific frequencies.

The limitation was not confined to the dominant EPI peak. Compared with the representative FFE sequences, EPI also showed higher residual energy across a broader portion of the analysis band. This suggests that the strong peak-dominated response may have influenced the overall broadband reduction, because small amplitude or phase errors near the dominant component can dominate the reduction metric while residual energy remains elsewhere in the band.

In addition to the broader residual energy, residual components remained near harmonic-related regions, including the component around 1.65 kHz, approximately twice the dominant EPI spectral component. This is consistent with the EPI-Y calibration response, where some harmonic-related acoustic components were observed even though they were weakly represented in the differentiated gradient input. Adding more sequence-specific calibration content may therefore be useful, although increasingly

targeted calibration could also overemphasize components that are not stable or consistently relevant during live playback.

The remaining residuals show that the EPI limitation was not simply a matter of missing spectral information in the calibration. Even after adding EPI-Y information around the dominant peak, cancellation remained sensitive to amplitude mismatch, phase error, harmonic components, and temporal changes in the scanner acoustic response. This makes EPI more demanding than the representative short FFE sequences, where the acoustic energy was distributed more broadly, and the redesigned broadband calibration gradients were generally sufficient.

Repeated EPI playback challenged the fixed LTI model.
The repeated noise-only and anti-noise-only measurements further suggest that the EPI limitation involved time-varying amplitude mismatch. The RMS-based estimated reduction decreased across consecutive runs 1–5 and increased again after the pause before run 6, eventually returning close to its initial value by run 8. This pattern suggests that the mismatch between the scanner noise and the generated anti-noise was not fully constant over time. However, because this reduction was estimated from separate recordings, it does not directly represent live cancellation of the EPI sequence.

The peak-based analysis showed that the EPI limitation was not explained only by the initial TF estimate. The anti-noise-only spectra were relatively repeatable, while the noise-only spectra changed more across repetitions. In particular, the dominant noise-only peak tended to decrease during the first consecutive runs, even though the broadband and narrowband RMS values remained relatively stable. This means that the scanner response changed locally around the dominant EPI component, without a comparable change in the overall noise level. Under these conditions, a fixed TF and fixed anti-noise scaling can become less accurate over time, even when the global RMS level and timing appear stable.

These observations point to a limitation of the fixed LTI model for EPI. The combined TF improved spectral characterization near the dominant EPI peak, but cancellation still degraded when the sequence response changed over repeated playback. The remaining error therefore cannot be explained by missing spectral information in the initial calibration alone. A fixed TF assumes that the acoustic response remains constant, while the repeated EPI measurements suggest that the dominant peak can change over time. This makes the fixed LTI assumption a major limitation for high-duty-cycle, peak-dominated sequences such as EPI.

Repeated EPI playback may alter the gradient system's acoustic response through temperature-related changes. Because EPI has a high-duty-cycle readout, a time-varying mechanical response could shift the amplitude, phase, or resonance behavior of the dominant acoustic component. The partial recovery after pauses is consistent with this interpretation, although the mechanism was not directly measured. Overall, the data point to time-varying scanner acoustic behavior, with gradient heating as a plausible but unconfirmed contributor. This makes sequence-specific, repeated, or adaptive correction an important open problem for EPI and other peak-dominated sequences.

The recovery observed after pauses of at least three minutes suggests that allowing sufficient rest between runs may help restore cancellation closer to its initial level. However, adding rest periods is difficult in real PNC measurements because scan time is limited. For future fMRI applications, the trade-off between acquisition efficiency and cancellation stability will need careful consideration. Defining the optimal measurement strategy by balancing the number of runs, rest intervals, and the specific outcomes of interest remains an important next step toward applying PNC in EPI-based functional imaging.

In summary, the EPI evaluation showed that combined chirp+EPI-Y TFs were more suitable for peak-dominated EPI measurements than chirp-only calibration. Chirp 1+EPI-Y and chirp 4+EPI-Y produced similar reductions, indicating that the targeted EPI-Y information was the main source of improvement once the dominant EPI band was included. This refines the FFE conclusion. Chirp 1 and chirp 4 are both strong broadband calibration candidates, but EPI requires additional sequence-specific TF information because its acoustic energy is concentrated around narrow dominant peaks. Residual amplitude mismatch, harmonic-related components, channel dependence, and repetition-dependent degradation still remained. Thus, sequence-specific TF information improved EPI PNC performance, but longer EPI acquisitions may require repeated or adaptive correction to maintain stable cancellation.

4.4. Comparison with Previous PNC Performance

A direct dB-by-dB comparison with the previous PNC study is limited by differences in hardware, playback configuration, measurement session, channel arrangement, and analysis bandwidth. The previous work evaluated reduction over 0.3–4 kHz and 0.5–2 kHz, whereas the present headphone-based setup focused on 0.5–4 kHz and 0.5–2 kHz because playback below 500 Hz was less reliable. For this reason, the comparison is more useful as a qualitative, within-system assessment of whether the rebuilt Python workflow reproduced the expected PNC behavior and whether the redesigned calibration gradients improved performance relative to sl014.

The rebuilt pipeline reproduced baseline PNC behavior and enabled improved reduction.
The results support both the rebuilt implementation and the calibration redesign. The original sl014 calibration produced measurable FFE reduction across sessions, indicating that the Python workflow preserved the basic PNC functionality. Chirp 1 and chirp 4 then improved on this baseline, with more reliable reduction than sl014 in both broadband and narrowband analyses, even though the equalizer quality differed between sessions. This shows that the redesigned broadband calibrations improved regular-sequence PNC performance relative to the original calibration approach under the current hardware and playback conditions.

Frequency- and sequence-dependent limits were consistent with prior work.
The frequency-dependent behavior was broadly consistent with previous PNC observations. Earlier PNC measurements showed the strongest reduction in the SPL-intensive mid-frequency range, where many dominant MRI acoustic components were located [39]. In the present FFE measurements, narrowband reduction was usually higher than broadband reduction, and the strongest visible suppression occurred mainly around dominant components below approximately 1.5 kHz. This pattern suggests that the redesigned calibrations worked best in frequency regions where the sequence noise was well represented by the calibration and could be reproduced accurately by the playback chain. Residual components near 0.5–0.6 kHz also show that the lower boundary of the present headphone equalization and playback range limited the achievable reduction.

Broadband calibration did not fully address the peak-dominated EPI measurements. Adding EPI-Y information reduced the residual near the dominant EPI component, and chirp 1+EPI-Y and chirp 4+EPI-Y produced similar reductions. For this EPI case, the targeted EPI-Y information mattered more than the choice between the two broadband chirp baselines. Harmonic-related residuals and repetition-dependent degradation still remained, so a more sequence-specific initial TF did not fully resolve the EPI limitation.

This is also consistent with the main model limitation identified in the previous PNC study. A fixed LTI model is limited when the scanner acoustic response changes over time [39]. In the present results, this was clearest for EPI. The combined TF improved the representation of the dominant spectral region, but cancellation still degraded during repeated playback. Overall, the present work extends the previous study by showing that redesigned broadband calibration improved FFE reduction relative to sl014, EPI-targeted information improved cancellation near the dominant EPI component, and stable cancellation of peak-dominated or time-varying responses will likely require more than a fixed LTI TF.

4.5. Limitations and Future Work

This work has several technical limitations, each of which suggests directions for future improvement. These limitations concern the scope of validation, the phantom and playback setup, the assumptions of the TF model, and the translation of the findings to future MRI applications.

Validation covered a finite set of cases.
Although the rebuilt Python implementation was sufficiently accurate for the tested PNC experiments, validation was limited to a finite set of test cases. Minor discrepancies in filtering, interpolation, TF derivation, and boundary handling remained between the original and new implementations. Future work should therefore expand the range of DSP and workflow validation cases, including additional calibration waveforms, regular sequence types, EPI-like peak-dominated inputs, low-SNR conditions, and ambiguous alignment cases.

Phantom measurements did not capture in vivo conditions.
Experimental validation was conducted only using phantoms, because approval for human subject experiments had not yet been obtained. While this enabled controlled comparison of calibration strategies, it does not fully replicate in vivo acoustic conditions. Headphone fit, acoustic loading, microphone placement, ear canal geometry, subject positioning, and bone conduction may all alter the acoustic transfer path and affect the achieved cancellation. Future work should therefore prioritize in vivo evaluation to assess whether the technical improvements translate into improved subject experience and reduced scanner noise interference during auditory fMRI.

Session and setup variability limited calibration comparisons.
A major limitation of the phantom experiments was their sensitivity to setup and session conditions. Several comparisons were conducted across different measurement sessions, and the results were affected by phantom positioning, headphone fit, microphone placement, and equalizer quality. Future experiments should separate calibration-gradient effects from session-level playback variability more systematically. This could include repeated setup-removal and replacement measurements, repeated headphone refitting tests, multiple independent equalizer derivations within the same session, and counterbalanced acquisition orders. The equalizer filter reduction quality metric should also be evaluated as a session-level quality-control measure before comparing calibration gradients.

Channel imbalance requires hardware-path testing.
Channel-specific behavior remains another key limitation. The weaker reduction observed in CH1 compared with CH2 limits the interpretation of channel-averaged reduction metrics, because average values can mask poor performance in one channel. Future work should isolate hardware and acoustic-path differences before attributing channel imbalance to calibration or algorithmic limitations. This could include headphone driver testing, channel-matching measurements, alternative playback hardware, or speaker-array-based delivery for more robust anti-noise reproduction.

Low-frequency playback limited cancellation.
The present headphone-based playback setup also limited cancellation near the lower edge of the analysis band. Residual components around 0.5–0.6 kHz remained difficult to cancel, suggesting that the playback and equalization chain may not reproduce the required anti-noise amplitude and phase accurately at low frequencies. Future work should improve low-frequency equalization, characterize headphone response under phantom and in vivo loading, and consider alternative playback hardware if low-frequency residuals remain dominant.

EPI may require adaptive or time-varying correction.
For more complex sequence types, especially EPI, future work should address the limitations of the fixed LTI model. One direction is to design more sequence-specific calibration waveforms that include relevant harmonic or peak-dominated components, while avoiding overemphasis of spectral components that are not stable across playback conditions. Another direction is to retain the LTI model but add adaptive gain and phase correction using live residual monitoring during or between repetitions. More flexible acoustic models, including nonlinear, time-varying, thermal, or data-driven models, should also be investigated to determine whether they improve prediction accuracy and cancellation stability during longer EPI acquisitions.

Future comparisons need matched analysis bands.
Future comparisons with previous PNC results should use matched analysis bands where possible, or report reductions across both the current and previous frequency ranges. This is important because the present headphone-based setup was less effective below 500 Hz, while previous work included the 300–500 Hz range. Within-setup comparisons, such as the improvement of redesigned chirp gradients over sl014, therefore remain more informative than raw absolute reduction values across different hardware configurations.

5

Conclusion

This thesis redeveloped and evaluated a Python-based PNC framework for MRI acoustic noise reduction. The original LabVIEW/MATLAB-based pipeline was successfully rebuilt in Python, including the main DSP, calibration, regular sequence processing, audio recording, AFG control, and workflow interface. Validation showed that the rebuilt implementation preserved the main DSP and workflow behavior of the original system, with full-workflow errors remaining below 5% for the evaluated blocks. The Python implementation also reduced computation time and improved usability, supporting its role as an improved replacement for the original system.

Calibration redesign improved phantom-based PNC performance, but no single calibration approach was superior across all metrics. Compared with sl014, the chirp-based gradients provided stronger excitation and generally improved TF stability and live reduction. Chirp 1 and chirp 4 were retained as the preferred broadband calibration candidates because they provided the most reliable overall performance. These findings show that calibration quality should be evaluated using both TF-domain metrics and live acoustic reduction, while also considering channel robustness and playback-chain limitations.

The redesigned calibration approaches also improved regular sequence phantom reduction, especially for short FFE sequences. Chirp 1 and chirp 4 provided more reliable FFE reduction than sl014 in both broadband and narrowband analyses across the three measurement sessions. Their relative performance depended on session-specific equalizer quality, with chirp 4 performing best in the initial session and chirp 1 performing best in the repeated sessions. These results indicate that chirp 1 and chirp 4 are the preferred broadband calibration options for representative FFE sequences, while also showing that playback-chain quality and low-frequency equalization can affect downstream performance. The separate chirp 4 versus chirp 4+EPI-Y comparison showed that adding EPI-Y information did not substantially compromise FFE broadband reduction, supporting the later use of combined TFs for EPI-focused evaluation.

PNC performance differed clearly between short representative FFE and EPI. Short FFE reduction was generally stronger and more repeatable, while EPI showed lower reduction and clearer temporal degradation. The combined chirp+EPI-Y TFs improved EPI reduction compared with chirp-only calibration and reduced residual amplitudes near the dominant EPI component around 825 Hz, but residual peaks, harmonic-related components, and repetition-dependent changes remained. These findings indicate that peak-dominated and high-duty-cycle sequences require more than broadband calibration alone, especially when the dominant acoustic response changes over repeated playback.

Overall, the rebuilt Python system was validated as a robust and efficient framework for phantom-based MRI acoustic noise reduction. The redesigned chirp calibrations improved regular-sequence reduction relative to sl014, with chirp 1 and chirp 4 providing the most reliable broadband performance under the present setup. Combined chirp+EPI-Y TFs improved EPI cancellation relative to chirp-only calibration, but repeated EPI playback remained limited by residual mismatch and time-varying acoustic behavior. Future work should focus on channel matching, low-frequency playback, time-varying acoustic modelling, adaptive correction, and in vivo evaluation.

Bibliography

[1] Oliver Behler and Stefan Uppenkamp. “Activation in human auditory cortex in relation to the loudness and unpleasantness of low-frequency and infrasound stimuli”. In: *PLoS ONE* 15.2 (Feb. 2020). ISSN: 19326203. DOI: 10.1371/journal.pone.0229088.

[2] Oliver Behler and Stefan Uppenkamp. “The representation of level and loudness in the central auditory system for unilateral stimulation”. In: *NeuroImage* 139 (Oct. 2016), pp. 176–188. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2016.06.025.

[3] Julius S.. Bendat and Allan G.. Piersol. *Random data : analysis and measurement procedures*. Wiley, 2010, p. 604. ISBN: 9780470248775.

[4] Graham A. Blackman and Deborah A. Hall. “Reducing the effects of background noise during auditory functional magnetic resonance imaging of speech processing: Qualitative and quantitative comparisons between two image acquisition schemes and noise cancellation”. In: *Journal of Speech, Language, and Hearing Research* 54.2 (Apr. 2011), pp. 693–704. ISSN: 10924388. DOI: 10.1044/1092-4388(2010/10-0143).

[5] Robert E. Brummet, J. Michael Talbot, and Peter Charuhas. “Potential Hearing Loss Resulting from MR Imaging”. In: 169 (1988), pp. 539–540.

[6] Carrie M Carr et al. “Evaluation of hearing loss in young adults after exposure to 3.0T MRI with standard hearing protection”. In: *The Journal of the Acoustical Society of America* 151.3 (Mar. 2022), pp. 1913–1921. ISSN: 0001-4966. DOI: 10.1121/10.0009824.

[7] Rebecca S. Dewey et al. “Comparison of continuous sampling with active noise cancelation and sparse sampling for cortical and subcortical auditory functional MRI”. In: *Magnetic Resonance in Medicine* 86.5 (Nov. 2021), pp. 2577–2588. ISSN: 15222594. DOI: 10.1002/mrm.28902.

[8] Francesco Di Salle et al. “fMRI of the auditory system: Understanding the neural basis of auditory gestalt”. In: *Magnetic Resonance Imaging* 21.10 (2003), pp. 1213–1224. ISSN: 0730725X. DOI: 10.1016/j.mri.2003.08.023.

[9] Ruth EC Evans et al. *Perceived patient burden and acceptability of whole body MRI for staging lung and colorectal cancer; comparison with standard staging investigations*. Tech. rep. 2018. URL: https://academic.oup.com/bjr/article/91/1086/20170731/7491366.

[10] Andrew J. Fagan et al. “7T MR Safety”. In: *Journal of Magnetic Resonance Imaging* 53.2 (Feb. 2021), pp. 333–346. ISSN: 15222586. DOI: 10.1002/jmri.27319.

[11] Søren A. Fuglsang et al. “BOLD fMRI responses to amplitude-modulated sounds across age in adult listeners”. In: *Imaging Neuroscience* 2 (July 2024), pp. 1–14. ISSN: 28376056. DOI: 10.1162/imag{_}a{_}00238.

[12] Søren A. Fuglsang et al. “Mapping cortico-subcortical sensitivity to 4 Hz amplitude modulation depth in human auditory system with functional MRI”. In: *NeuroImage* 246 (Feb. 2022). ISSN: 10959572. DOI: 10.1016/j.neuroimage.2021.118745.

[13] Deborah A. Hall et al. “Acoustic, psychophysical, and neuroimaging measurements of the effectiveness of active cancellation during auditory functional magnetic resonance imaging”. In: *The Journal of the Acoustical Society of America* 125.1 (Jan. 2009), pp. 347–359. ISSN: 0001-4966. DOI: 10.1121/1.3021437.

[14] Deborah A. Hall et al. “Functional magnetic resonance imaging measurements of sound-level encoding in the absence of background scanner noise”. In: *The Journal of the Acoustical Society of America* 109.4 (Apr. 2001), pp. 1559–1570. ISSN: 0001-4966. DOI: 10.1121/1.1345697.

[15] Takashi Hamaguchi et al. “Acoustic noise transfer function in clinical MRI. A multicenter analysis”. In: *Academic Radiology* 18.1 (Jan. 2011), pp. 101–106. ISSN: 10766332. DOI: 10.1016/j.acra.2010.09.009.

[16] Heledd C Hart, Alan R Palmer, and Deborah A Hall. *Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex*. Tech. rep. May 2002, pp. 177–190. URL: www.elsevier.com/locate/heares.

[17] Robert A. Hedeen and William A. Edelstein. "Characterization and prediction of gradient acoustic noise in MR imagers". In: *Magnetic Resonance in Medicine* 37.1 (1997), pp. 7–10. ISSN: 07403194. DOI: 10.1002/mrm.1910370103.

[18] Sarah M. Jacobs et al. "Image quality and subject experience of quiet T1-weighted 7-T brain imaging using a silent gradient coil". In: *European radiology experimental* 6.1 (Aug. 2022), p. 36. ISSN: 25099280. DOI: 10.1186/s41747-022-00293-x.

[19] Emanuel Kanal, Frank G. Shellock, and Lalith Talagala. *Safety Considerations in MR Imaging*. Tech. rep. 1990, pp. 593–606.

[20] Hannah R. Kurdila et al. "Survey of Acoustic Output in Neonatal Brain Protocols". In: *Journal of Magnetic Resonance Imaging* 54.4 (Oct. 2021), pp. 1119–1125. ISSN: 15222586. DOI: 10.1002/jmri.27656.

[21] L Jancke. *Intensity coding of auditory stimuli: an fMRI study*. Tech. rep. 1998. DOI: 10.1016/s0028-3932(98)00019-0.

[22] Dave R.M. Langers et al. "fMRI activation in relation to sound intensity and loudness". In: *NeuroImage* 35.2 (Apr. 2007), pp. 709–718. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2006.12.013.

[23] Nokhaeng Lee, Youngjin Park, and Gun Woo Lee. "Frequency-domain active noise control for magnetic resonance imaging acoustic noise". In: *Applied Acoustics* 118 (Mar. 2017), pp. 30–38. ISSN: 1872910X. DOI: 10.1016/j.apacoust.2016.11.003.

[24] Mingfeng Li et al. "In situ active control of noise in a 4 T MRI scanner". In: *Journal of Magnetic Resonance Imaging* 34.3 (Sept. 2011), pp. 662–669. ISSN: 10531807. DOI: 10.1002/jmri.22694.

[25] Michael J. B. McGrory et al. "Fast and silent MRI using nonlinear gradient fields at the ultrasonic gradient switching frequency of 20 kHz with a Point Spread Function framework reconstruction". In: *Magnetic Resonance in Medicine* (Aug. 2024). ISSN: 0740-3194. DOI: 10.1002/mrm.30230.

[26] Mark McJury and Frank G. Shellock. *Auditory noise associated with MR procedures: A review*. July 2000. DOI: 10.1002/1522-2586(200007)12:1<37::AID-JMRI5>3.0.CO;2-I.

[27] Mark J. McJury. *Acoustic Noise and Magnetic Resonance Imaging: A Narrative/Descriptive Review*. Feb. 2022. DOI: 10.1002/jmri.27525.

[28] Shashikant R. More et al. "Acoustic noise characteristics of a 4 Telsa MRI scanner". In: *Journal of Magnetic Resonance Imaging* 23.3 (Mar. 2006), pp. 388–397. ISSN: 10531807. DOI: 10.1002/jmri.20526.

[29] Jonathan E. Peelle. "Methodological challenges and solutions in auditory functional magnetic resonance imaging". In: *Frontiers in Neuroscience* 8 JUL (2014). ISSN: 1662453X. DOI: 10.3389/fnins.2014.00253.

[30] Mark E Quirk et al. *Anxiety in Patients Undergoing MR Imaging*. Tech. rep. 1989.

[31] Michael E Ravicz and Jennifer R Melcher. *Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: Examination of noise conduction through the ear canal, head, and body a)*. Tech. rep. 2001, pp. 216–231.

[32] Carlos V. Rizzo Sierra et al. "Acoustic fMRI noise: Linear time-invariant system model". In: *IEEE Transactions on Biomedical Engineering* 55.9 (Sept. 2008), pp. 2115–2123. ISSN: 00189294. DOI: 10.1109/TBME.2008.923112.

[33] Markus Röhl and Stefan Uppenkamp. "Neural coding of sound intensity and loudness in the human auditory system". In: *JARO - Journal of the Association for Research in Otolaryngology* 13.3 (June 2012), pp. 369–379. ISSN: 15253961. DOI: 10.1007/s10162-012-0315-6.

[34] N. B. Roozen, A. H. Koevoets, and A. J. den Hamer. “Active vibration control of gradient coils to reduce acoustic noise of MRI systems”. In: *IEEE/ASME Transactions on Mechatronics* 13.3 (June 2008), pp. 325–334. ISSN: 10834435. DOI: 10.1109/TMECH.2008.924111.

[35] *Safety Guidelines for Magnetic Resonance Imaging Equipment in Clinical Use*. Tech. rep. Medicines and Healthcare products Regulatory Agency, 2021.

[36] Richard Salvi and Adam Sheppard. *Is noise in the MR imager a significant risk factor for hearing loss?* Feb. 2018. DOI: 10.1148/radiol.2017172221.

[37] S. Schmitter et al. “Silent echo-planar imaging for auditory FMRI”. In: *Magnetic Resonance Materials in Physics, Biology and Medicine* 21.5 (Sept. 2008), pp. 317–325. ISSN: 09685243. DOI: 10.1007/s10334-008-0132-4.

[38] Irina S Sigalovsky and Jennifer R Melcher. *Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers*. Tech. rep. May 2006, pp. 67–76.

[39] Paulina Šiurytė and Sebastian Weingärtner. *Lowering The Acoustic Noise Burden in MRI with Predictive Noise Canceling*. Apr. 2024. DOI: 10.1101/2024.04.28.24305337. URL: http://medrxiv.org/lookup/doi/10.1101/2024.04.28.24305337.

[40] Kenneth I. Vaden et al. “Cingulo-opercular adaptive control for younger and older adults during a challenging gap detection task”. In: *Journal of Neuroscience Research* 98.4 (Apr. 2020), pp. 680–691. ISSN: 10974547. DOI: 10.1002/jnr.24506.

[41] Ziyue Wu et al. “Evaluation of an independent linear model for acoustic noise on a conventional MRI scanner and implications for acoustic noise reduction”. In: *Magnetic Resonance in Medicine* 71.4 (2014), pp. 1613–1620. ISSN: 15222594. DOI: 10.1002/mrm.24798.

A

PNC Canonical Function Mapping

This appendix details the canonical function blocks and artifacts mapped from the legacy LabVIEW implementation to the current Python architecture.

A.1. System Workflows and Canonical Blocks

Table A.1 outlines all system orchestrations, including general utilities, system calibration, and regular sequence runs in PNC.

Table A.1: Canonical blocks detailing the purpose and transition from legacy to Python implementation.

ID	Canonical block	Purpose	Legacy implementation	Python implementation	Main artifacts / outputs
General System Blocks					
C00	Session initialization and workspace setup	Create experiment folder, initialize paths, prepare run context	`PNCcal.vi` / `PNC.vi` init logic	`PNC.py`, `src/main.py`, startup window / workdir selection	working directory, run paths (A00)
C01	Settings and shared state	Hold user settings, calibration parameters, runtime state	LabVIEW settings/calibration clusters, in-memory variables	`src/settings.py`, `src/state.py`	config values, run state
C02	Logging and run traceability	Write process logs and result logs	`process_log_message.vi`, csv logging logic	`src/utils/logger.py`	.csv or .log logs (A01)
C03	Audio recording service	Start, collect, and close microphone recording sessions	`labview_recording_initiate.vi`, `labview_recording_collect.vi`, `labview_recording_close.vi`	`src/utils/pnc_audio.py`, `src/utils/record_audio.py`, `src/utils/record_audio_sd.py`	.wav recordings, buffer data (A02)

Continued on next page

Table A.1: Canonical blocks detailing the purpose and transition (continued).

ID	Canonical block	Purpose	Legacy implementation	Python implementation	Main artifacts / outputs
C04	AFG communication	Upload waveform(s), configure triggering, start/stop/delete AFG content	Tektronix AFG31000 driver VIs, `write_seq_to_afg.vi`, `write_arr_xyz_to_afg.vi`, `write_short_tfwx_to_afg.vi`, `afg_delete_files.vi`	`src/utils/afg_driver.py`, `src/utils/afg_sequences.py`	uploaded waveforms, sequence files, output scale
C05	Plotting	Plot diagnostics and derived results for user inspection	no standalone plotting block, integrated in MATLAB DSP scripts	`src/utils/plotter.py` and plotting inside DSP/runtime	plots, visual QA (A09)
C06	General DSP utilities	General calculation supporting advanced DSP	`sigdelay2.m`, `window_clip_edges.m`, `correv.m`, `sigshift.m`, MATLAB bandpass and resample	`src/dsp/general.py`	Processed data or array
C07	User interface	User friendly interface for the PNC program	`PNCcal.vi` / `PNC.vi` front panel windows	`PNC.py` UI window called from `src/ui`	UI display
Calibration Workflow					
K00	Calibration controller	Top-level orchestration of calibration workflow	`PNCcal.vi`	calibration workflow in `src/main.py`, `src/runtime/cal_*` orchestration	full calibration run
K01	Background noise reference	Measure microphone/environment baseline noise	`background_noise_reference.vi`, `estimate_background_noise.m` + recording VIs	`src/runtime/cal_step0_background_ref.py`, `src/dsp/background.py` + recording utils	background RMS (A02, A09, A11)
K02	I/O scale derivation	Determine amplitude scaling for EQ derivation	no I/O scale implemented	`src/runtime/cal_step0_io_scale.py`, `src/dsp/io_ramp.py` + AFG/recording utils	I/O scale (A02, A09, A10)
K02b	Microphone SPL calibration	Determine target playback for fMRI scaling	no microphone SPL implemented	`src/runtime/cal_step0_io_scale.py`, `src/dsp/io_ramp.py` + AFG/recording utils	I/O scale, SPL threshold result (A02, A09, A10)

Continued on next page

Table A.1: Canonical blocks detailing the purpose and transition (continued).

ID	Canonical block	Purpose	Legacy implementation	Python implementation	Main artifacts / outputs
K03	Frequency sweep test and drift estimation	Record sweeps, assess headphone response, estimate clock drift	`frequency_sweeps_initiate.vi`, `frequency_sweeps_collect.vi`, `frequency_sweeps_analyze.vi`, `update_clock_drift_value.vi`, `generate_sweep_series.m`, `process_sweeps.m` + AFG/recording VIs	`src/runtime/cal_step0_frequency_sweeps.py`, `src/runtime/cal_step0_update_clock_drift.py`, `src/dsp/sweeps.py` + AFG/recording utils	sweep recordings, drift estimate (A02, A04, A06, A09)
K04	Equalizer derivation	Derive pre-equalization filter for playback chain	`equalizer.vi`, `prepare_equalizer_input.m`, `derive_eq_filter.m`, `apply_eq_filter.m`, `plot_eq_results.m` + AFG/recording VIs	`src/runtime/cal_step0_equalizer_derivation.py`, `src/dsp/equalizer.py` + AFG/recording utils	EQ filter, tuning data, EQ plots (A02, A05, A06, A09)
K05	Calibration step 1 acquisition	Record scanner calibration noise / gradient response	Step 1 recording part of `PNCcal.vi` using recording VIs	`src/runtime/cal_step1_transfer_functions.py` + recording utils	Step 1 .wav (A02)
K06	Calibration transfer function derivation	Process step 1 recording, correct drift, derive acoustic transfer functions	`derive_transfer_functions.vi`, `derive_transfer_functions.m`	`src/runtime/cal_step1_transfer_functions.py`, `src/dsp/calibration_dsp.py`	transfer function (hxyz), averaged clips, updated drift (A03, A06, A09)
K06b	Calibration combine transfer function	Combine transfer function	-	`src/runtime/cal_step4_combine_transfer_functions.py`, `src/dsp/calibration_dsp.py`	transfer function (hxyz)

Continued on next page

Table A.1: Canonical blocks detailing the purpose and transition (continued).

ID	Canonical block	Purpose	Legacy implementation	Python implementation	Main artifacts / outputs
K07	Calibration anti-noise output preparation	Generate step 2/3 calibration anti-noise from TFs, polarity inversion, equalization	`prepare_calib_step2_output.vi`	`src/runtime/cal_step1_prepare_calib.py`, `src/dsp/calibration_dsp.py`	equalized X/Y/Z outputs, step 1 simulated reduction (A03, A07, A09)
K08	Calibration output selection / reuse	Reuse previously prepared calibration output instead of regenerating it	`select_calib_step2_output.vi`	`src/runtime/cal_step1_select_calib.py`	loaded equalized calibration outputs
K09	Calibration step 2/3 execution	Send calibration anti-noise to AFG and collect recording iteratively	Step 2/3 control loop in `PNCcal.vi` + AFG/recording VIs	Step 2/3 runtime flow in calibration scripts + AFG/recording utils	Step 2/3 recordings (A02)
K10	Calibration step 2/3 processing	Estimate latency and simulated/live reduction from step 2/3 recordings	`process_calibration_steps_2_3.vi`, `process_calibration_recording.m`	`src/runtime/cal_step23_process_calib.py`, `src/dsp/calibration_dsp.py`	latency, reduction metrics, corrected outputs (A06, A08, A09)
Regular PNC Workflow					
R00	Regular PNC controller	Top-level orchestration of regular sequence workflow	`PNC.vi`	regular runtime orchestration in `src/main.py`, `src/runtime/reg_*`	full regular PNC run
R00b	Optional runtime clock drift recalibration	Re-estimate drift during regular PNC and restart preparation with updated drift	no clock drift recalibration implemented	regular runtime orchestration in `src/main.py`, `src/runtime/reg_*`, `src/dsp/sweeps.py`, `src/runtime/cal_step0_update_clock_drift.py`	updated drift estimate (A04, A06)
R01	Regular sequence input loading	Load gradient waveform / sequence-specific input data	gradient input .mat loading in `PNC.vi` flow	regular workflow input loading	recording duration
R02	Regular step 1 acquisition	Record scanner noise for target regular sequence	Step 1 recording part of `PNC.vi` using recording VIs	`src/runtime/reg_step1_prepare_reg.py` + recording utils	Step 1 regular .wav (A02)

Continued on next page

Table A.1: Canonical blocks detailing the purpose and transition (continued).

ID	Canonical block	Purpose	Legacy implementation	Python implementation	Main artifacts / outputs
R03	Regular output preparation	Correct step 1 for drift, predict noise using TFs, invert and equalize playback signal	`prepare_reg_step2_output.vi`, `process_reg_output.m`, `equalize_reg_output.m`	`src/runtime/reg_step1_prepare_reg.py`, `src/dsp/reg_dsp.py`	predicted anti-noise (raw and equalized), step 1 simulated reduction (A07, A09)
R04	Regular output selection / reuse	Reuse previously prepared regular anti-noise output	`select_reg_step2_output.vi`	`src/runtime/reg_step1_select_reg.py`	loaded equalized regular output
R04b	fMRI stimulus generation and mixing	Generate fMRI stimulus, scale it, and mix it into prepared anti-noise before step 3 playback	no fMRI workflow implemented	`src/runtime/reg_step3_pnc_with_fmri.py`	updated anti-noise playback, stimulus files in `work_dir/stim` (A12)
R05	Regular step 2/3 execution	Send regular anti-noise to AFG and collect step 2/3 recordings	Step 2/3 control loop in `PNC.vi` + AFG/recording VIs	`src/runtime/reg_step23_process_reg.py` flow + AFG/recording utils	Step 2/3 regular recordings (A02)
R06	Regular step 2/3 processing	Estimate residual latency and simulated/live reduction	`process_reg_steps_2_3.vi`, `process_reg_output2.m`	`src/runtime/reg_step23_process_reg.py`, `src/dsp/reg_dsp.py`	residual latency, reduction metrics (A08, A09)

A.2. System Artifacts and Outputs

Table A.2 defines the primary inputs, artifacts, and output files utilized across all the previously defined blocks, mapping their legacy counterparts to the new Python file structures.

Table A.2: Main artifacts and outputs transitioned from legacy to Python implementation.

ID	Artifact	Purpose	Produced / updated by	Consumed / reused by	Legacy form	Python form
A00	Experiment folder	Root folder for one experiment run and its saved outputs	Session initialization (C00)	All PNC steps	working directory	`work_dir`
A01	Process / result logs	Store run trace, messages, and summary results for traceability	Logging and save steps (C02, runtime save logic)	User inspection, debugging, post-run review	`yymmdd_calib_process_log.csv` / `yymmdd_process_log.csv`	`experiment.log`

Continued on next page

Table A.2: Main artifacts and outputs transitioned from legacy to Python implementation (continued).

ID	Artifact	Purpose	Produced / updated by	Consumed / reused by	Legacy form	Python form
A02	Recordings	Store raw microphone recordings from calibration and regular-sequence steps	Acquisition / execution blocks (K01-K05, K09, R02, R05)	DSP processing, reuse, validation	`.wav` recordings	`.wav` recordings
A03	Transfer functions	Store estimated acoustic transfer functions derived from calibration step 1	Transfer function derivation (K06), combine transfer function (K06b)	Calibration output preparation (K07), regular output preparation (R03), reuse across runs	stored in `yymmdd_calibration_data_step_1.*` and also embedded in A06 output tuning file	stored in `yymmdd_calibration_data_step_1.*` and also embedded in A06 output tuning file
A04	Drift estimate	Store estimated PC clock mismatch / drift correction parameter	Sweep analysis (K03), possibly updated in step 1/2/3 calibration (K06, K10) and drift re-estimation in reg workflow (R00b)	Calibration and regular-sequence processing (K04, K06, K10, R03, R06)	stored in `yymmdd_sweep_rec.*` and also embedded in A06 output tuning file	stored in `yymmdd_sweep_rec.*` and also embedded in A06 output tuning file
A05	Equalizer filter	Store derived playback equalization filter and related parameters	Equalizer derivation (K04)	Calibration output preparation (K07), regular output preparation (R03)	stored in A06 output tuning file	stored in A06 output tuning file
A06	Output tuning file	Main persistent calibration handoff artifact containing tuning parameters reused in later runs	Calibration save / update step (K01, K03, K04, K06, K10)	Calibration (K04, K06, K10) and regular PNC workflow (R03, R04)	`yymmdd_output_tuning.mat`	`yymmdd_output_tuning.npz`
A07	Prepared equalized anti-noise output	Store generated anti-noise playback signal prepared for step 2/3 execution or reuse	Calibration / regular output preparation (K07, R03)	Output selection/reuse and step 2/3 execution (K08, R04, K09, R05)	`yymmdd_equalized_cal_inputs.mat` / `yymmdd_equalized_reg_inputs.mat`	`yymmdd_equalized_cal_inputs.npz` / `yymmdd_*_equalized_reg_inputs.npz`

Continued on next page

Table A.2: Main artifacts and outputs transitioned from legacy to Python implementation (continued).

ID	Artifact	Purpose	Produced / updated by	Consumed / reused by	Legacy form	Python form
A08	Latency and reduction results	Store simulated/live reduction metrics and latency estimates from step 2/3 processing	Processing blocks (K10, R06)	Result review, comparison, and validation	`yymmdd_calibration_data_*.mat` / `yymmdd_reg_data_*.mat`	`yymmdd_calibration_data_*.npz` / `yymmdd_*_reg_data_*.npz`
A09	Plots	Plot process and results	Plotting / inspection functions (C05)	User inspection	Plot previews (not saved)	`work_dir/figures`
A10	I/O scale	I/O scale or I/O EQ scale for equalizer and fMRI experiment	I/O scale derivation (K02), microphone SPL calibration (K02b)	Equalizer derivation (K04) and fMRI stimulus mixing (R04b)	-	`yymmdd_io_scale.npz` / `yymmdd_io_scale_EQ.npz`
A11	Background RMS	Estimate background noise floor	Background noise reference (K01)	Equalizer derivation (K04)	`yymmdd_background_test.mat`	`yymmdd_background_test.npz`
A12	fMRI stimulus data	Generated fMRI stimulus waveforms and design files used for fMRI playback/mixing	fMRI stimulus generation and mixing (R04b)	Step 3 preparation and execution when fMRI mode is enabled (R04b, R05)	-	`work_dir/stim`

B

Python-Based PNC Implementation

B.1. Source Code

The complete source code for the Python-based PNC program is hosted on the Delft University of Technology GitLab server. The repository can be accessed at `https://gitlab.tudelft.nl/mars-lab/pypnc`. Access will be granted upon request, and the repository will be made publicly available upon publication in a journal.

B.2. Graphical User Interface

This section provides extended views of the PNC graphical user interface. These additional screens detail the software initialization process, hardware configuration, and system calibration workflows that complement the main experiment control interface discussed in the primary text.

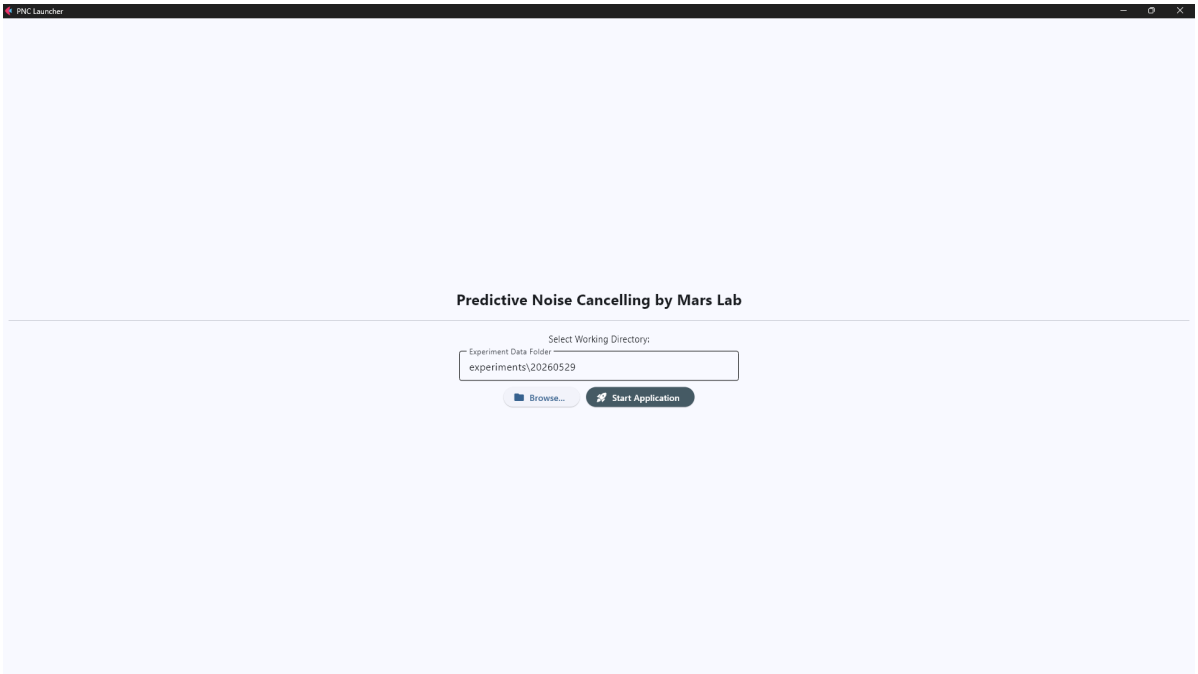


Figure B.1: The PNC Launcher interface. This initial screen allows the user to browse and select the working directory where experiment data and logs will be stored before starting the main application.

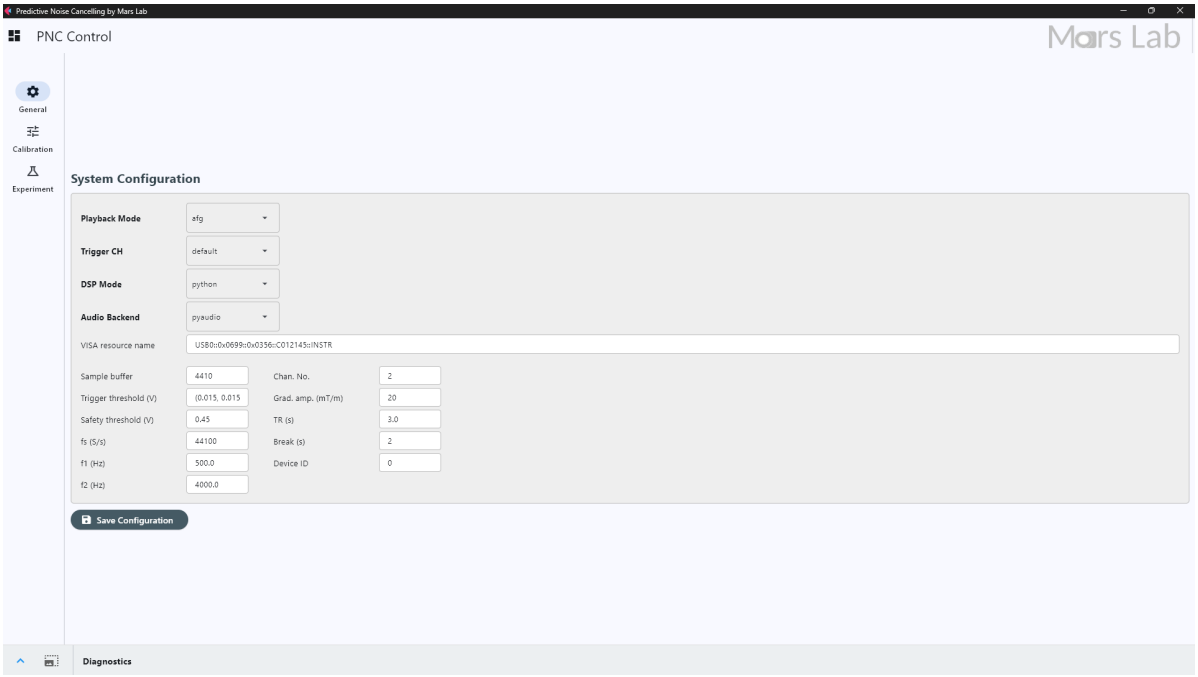


Figure B.2: The System Configuration tab under the General settings menu. This interface manages essential hardware and software parameters, including the DSP and audio backends, instrument VISA resource names, safety thresholds, and timing parameters.

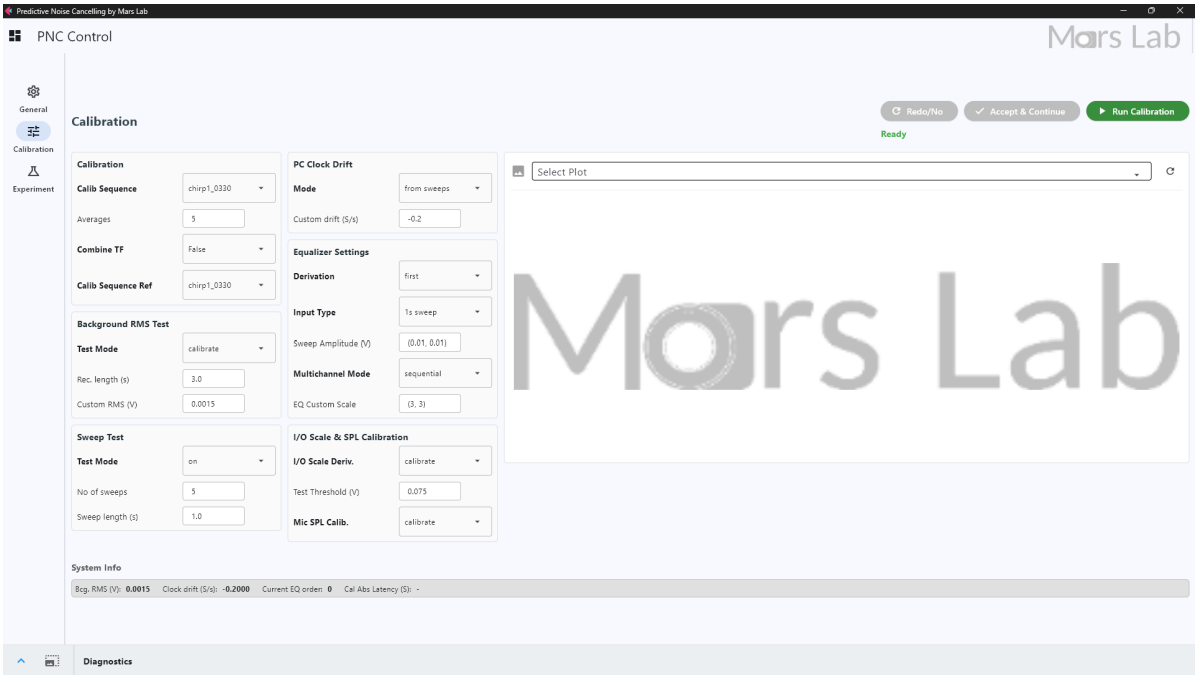


Figure B.3: The Calibration tab. This menu provides comprehensive controls for system calibration routines prior to an experiment. Key configurations include background noise floor testing, PC clock drift estimation, equalizer settings, and I/O scale calibration.

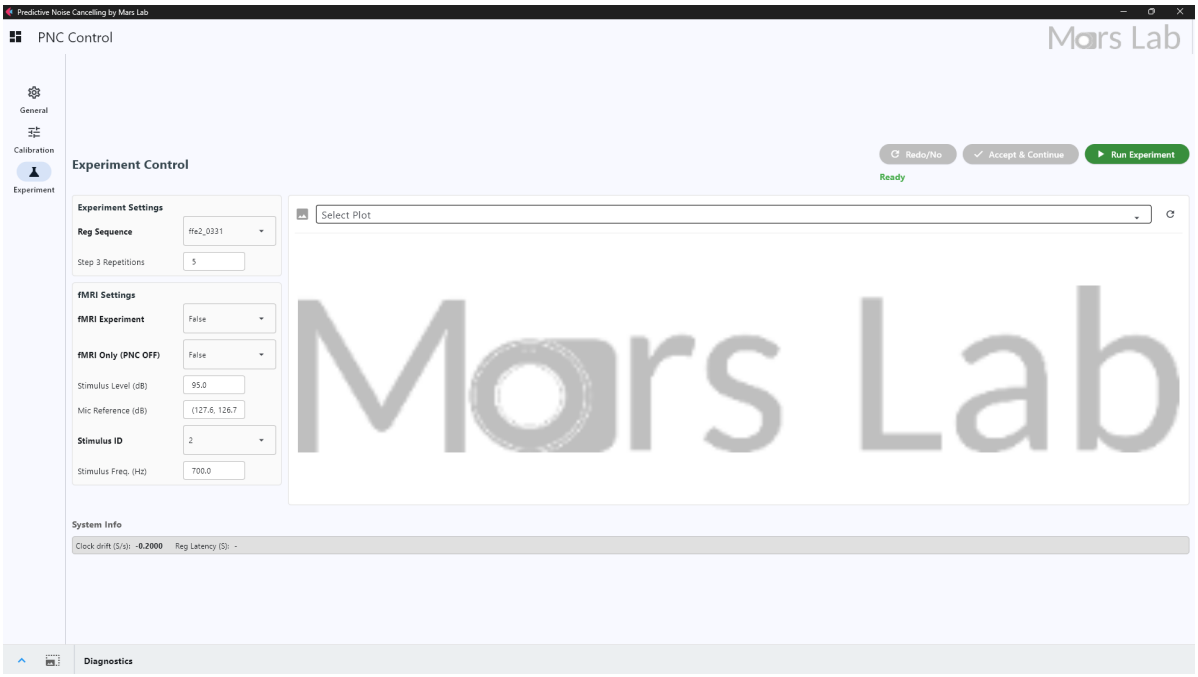


Figure B.4: The Experiment tab. This menu provides comprehensive controls for regular sequence experiments. Key configurations include configurations for sequence list and fMRI settings.