PRACTICAL TOOLS

ESPERDYNE: A Dual-Band Heterodyne Monitor and Ultrasound Recorder for Bioacoustic Field Surveys

Ravi Umadi [1]

[1] Lehrstuhl für Zoologie, TUM School of Life Sciences, Technische Universität München, Liesel-Beckmann-Str. 4, 85354, Freising-Weihenstephan, Germany

Abstract.

1. Background. Ultrasonic monitoring is essential for ecological studies of bats and other animals, yet high-performance field devices remain prohibitively expensive and inaccessible–particularly in biodiversity-rich regions with limited research infrastructure. Existing low-cost options often lack real-time listening and recording features. There remains a critical need for versatile, affordable, and field-ready tools that support acoustic behavioural research, educational and conservation outreach.

2. New Tool. I introduce *Esperdyne*, an open-source, dual-channel ultrasound monitoring and recording system based on the ESP32-S3 microcontroller. With a component cost under €75, Esperdyne combines real-time heterodyne monitoring, stereo recording from a retroactive ring buffer, and an intuitive rotary-based user interface with OLED display. It supports full-duplex 192 kHz audio, dual-band tuning for simultaneous FM/CF monitoring, and real-time playback via headphones or a speaker. All audio processing–including adjustable carrier frequency mixing, gain control, and file-saving logic–is implemented without reliance on fixed-rate audio libraries.

3. Applications. Esperdyne has been tested in field conditions and shown to reliably detect high-SNR calls and harmonics from free-flying bats. A companion tool *Bat Reviewer* supports rapid inspection, playback, and export of selected recordings. Together, these tools enable portable, solo-operated acoustic surveys with minimal training. Beyond ecological research, Esperdyne is suitable for education, outreach, and preliminary field assessments in remote or resource-constrained settings. Its modular design encourages hardware customisation and firmware extension by interdisciplinary teams.

4. Availability and Implementation. Full hardware schematics, firmware, and software tools are publicly available. The system can be built using hobbyist-accessible components and standard Arduino tooling. By sharing this system openly, I aim to lower technical barriers and foster broader participation in ultrasound-based biodiversity monitoring and conservation. Esperdyne demonstrates how microcontroller-based platforms can bridge gaps between affordability, usability, and scientific capability–supporting global efforts in soundscape ecology.

Keywords: bat echolocation, ultrasound monitoring, ESP32-S3, open hardware, heterodyne audio, field recording, conservation technology

Corresponding author: Ravi Umadi– Email: ravi.umadi@tum.de

1 INTRODUCTION

Acoustic monitoring of bat calls is a widely used method for species surveys, identification, and even the discovery of previously unknown taxa [1–4]. It is also a key educational tool in soundscape ecology, providing an accessible gateway into understanding animal communication and nocturnal biodiversity [5–10]. The most effective tools for active field surveys—particularly when tracking bats, on the move through challenging terrain—are lightweight, reliable, and intuitive to operate [11].

Despite recent advances, the range of available heterodyne detectors and field recorders remains limited, especially in the lower cost segment. Many professional-grade devices cost several hundred to several thousand euros. For example, one of the most widely used heterodyne-enabled field recorders, the Pettersson D1000X [12], retails for approximately €3600, weighs 600 grams (excluding batteries), and requires CompactFlash cards for storage. Other commercial bat detectors with recording capabilities are typically priced around €1400, while non-recording devices often start at €250. In Germany, the SSF BAT3-Detektor is a commonly used monitoring unit with limited recording capabilities and a cost of approximately €350. Handheld devices from manufacturers like Dodotronics provide mobile-compatible plug-and-play solutions, but often require additional proprietary software and cost over €250.

While these devices offer robust features and professional reliability, their high cost remains a significant barrier—particularly for educators, students, and conservation projects in the Global South. From personal experience, I recall the considerable challenge of acquiring a bat detector early in my research career in India: the device cost over a month's salary and incurred additional import duties. I was struck by the irony that regions rich in bat diversity often lack access to the tools necessary to study them. In contrast, the tools originate from areas with relatively low bat diversity.

Open-source projects such as AudioMoth [13] and RPi Bat Projekt [14] have addressed part of this gap by providing low-power, autonomous passive recording platforms. These systems are well-suited for long-term monitoring and programmable deployment but lack the capability for real-time listening and user-directed recording, which can be critical for behavioural studies and educational outreach.

I recently developed a multichannel field recorder for bioacoustic behavioural experiments as part of my ongoing work in *embedded ultrasonics*. Combined with custom microphone arrays built using a modular toolkit [15–17], this system offers both real-time listening and retroactive capture of ultrasonic events, enabling novel experimental paradigms. Building on this foundation,

I present the design and implementation of *Esperdyne*[1]—a compact, low-cost, open-source dual-band heterodyne monitor and stereo ultrasound recorder built around the ESP32-S3 microcontroller.

Esperdyne allows simultaneous tuning of two frequency bands, enabling live parallel monitoring of frequency-modulated (FM) and constant-frequency (CF) bats—a feature rarely found even in high-end detectors. The system features an intuitive user interface with rotary-encoder-based control, OLED display, real-time headphone and speaker output, and a tap-to-save ring buffer allowing up to 20 seconds of retroactive recording. It is built entirely from off-the-shelf, hobbyist-accessible components, with fully open-source firmware and hardware.

The prototype was assembled on a perforated PCB and field-tested during the summer of 2025 along the river Isar in southern Germany. Figures in this paper provide representative recordings, and reference files from these surveys are available for comparison and testing by users building their own units.

The Esperdyne provided a high-fidelity listening experience with minimal self-noise—an issue common in other bat detectors that artificially boost gain to reveal weak signals. The perceived audio quality depends strongly on the microphone and DAC components used, yet the system, as tested, produced results comparable to commercial units costing over 10 times as much. With a build cost under €75, Esperdyne offers a powerful, affordable alternative for researchers, educators, and students worldwide.

Complementing the hardware, I also developed *Bat Reviewer*, a lightweight MATLAB-based graphical interface for browsing, inspecting, and curating ultrasonic recordings – available as cross-platform standalone executables [18]. The tool enables users to navigate entire recording sessions, visualise waveforms and spectrograms, and listen to heterodyned audio in real time with adjustable carrier frequency and playback volume. Export functions streamline the selection of high-quality calls and facilitate organised data curation. Combining visual inspection with auditory evaluation in a single interface, *Bat Reviewer* shortens the review cycle and supports more systematic field data processing.

In summary, Esperdyne and BatReviewer demonstrate how open-source, low-cost solutions can lower barriers to acoustic monitoring. By coupling accessible hardware with flexible software, these tools provide a practical foundation for research and educational applications, particularly in regions where commercial systems are prohibitively expensive. Beyond immediate utility, they also highlight the potential of community-driven development in advancing the field of bioacous-

[1] The name combines *ESP* (the ESP32 microcontroller) and *dyne* (from heterodyne), with *er* referencing *Vesper*—Latin for "evening" and the root of the bat family name *Vespertilionidae*.

tics and ensuring broader participation in biodiversity monitoring.

2 METHODS

The Esperdyne system is built around the ESP32-S3 (Espressif Systems (Shanghai) Co., Ltd.) [19] microcontroller, selected for its dual-core architecture, onboard Pseudo-Static Random Access Memory (PSRAM), and two independent I2S ports—crucial for full-duplex ultrasonic audio processing. It integrates a high-fidelity WM8782 24-bit stereo ADC (Cirrus Logic, Inc., Austin, Texas, USA) [20] for input and a PCM5102A DAC (Texas Instruments, Dallas, Texas, USA) [21] for audio output.

Audio from the SPU0410LR5H-QB analogue omnidirectional MEMS ultrasonic microphones (Knowles Electronics, LLC. Itasca, IL, USA) is digitised via the WM8782 ADC and captured through the I2S0 input interface. The digitised samples are processed for heterodyne conversion (based on user-set carrier frequencies) and output via the PCM5102A DAC through I2S1. Audio is simultaneously written into a ring buffer stored in PSRAM, allowing buffering up to 20 seconds on the onboard 8MB PSRAM with a single channel at a given sampling rate and bit depth.

All components were assembled on 6x8 cm perforated prototyping PCBs using hand-soldered point-to-point wiring.

2.1 User Interface and OLED Display

The Esperdyne device features a minimal yet intuitive user interface built around a rotary encoder, three tactile buttons, and a 128×32 pixel OLED display. This interface allows real-time adjustment of heterodyne parameters and clear feedback on device status.

Rotary Encoder and Parameter Selection. The rotary encoder serves as a dual-mode control for adjusting the heterodyne **carrier frequency** (F) and **playback gain** (V). A short press on the encoder toggles between these two editable parameters. The currently selected parameter is indicated by displaying its corresponding label (`F` or `V`) on the OLED. Rotating the encoder clockwise increases the value, while counter-clockwise rotation decreases it. Adjustments are applied in real-time to the active audio channel.

Tap-to-Save. A separate *tap-to-save* button triggers the recording function, writing the last 5 seconds of buffered audio to the SD card as a WAV file. While writing the file, a message "Saving File" is displayed.

Channel 2 Frequency Adjustment. When the device is in frequency tuning mode, a dedicated

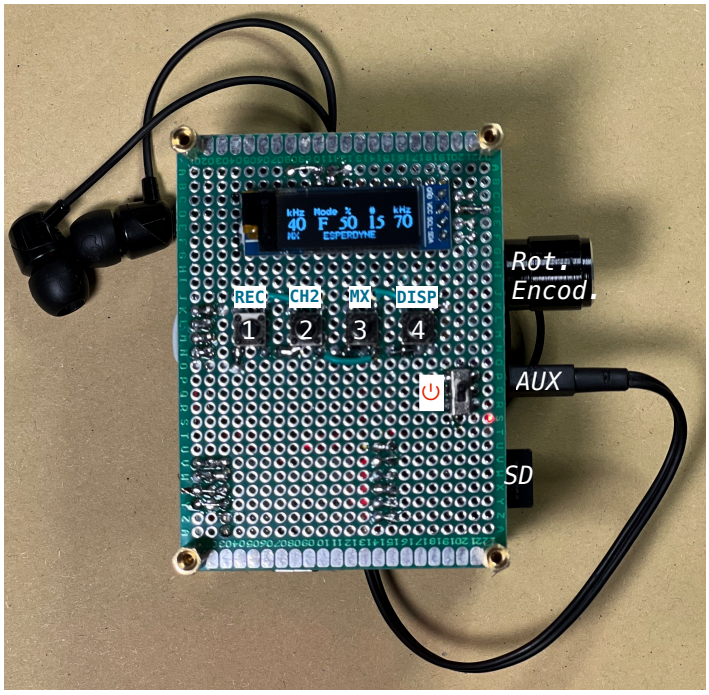

Figure 1: Esperdyne Field Recorder. The device is based on the ESP32-S3 MCU with dual I2S input, rotary encoder, OLED display, and microSD card. It includes a tap-to-save recording button (REC), channel 2 tuning hold button (CH2), and an OLED screen toggle (DISP). The front face includes ports and controls: SD card slot, rotary encoder, AUX Output, and power button. The interface allows real-time stereo heterodyne output with independent frequency tuning for each channel, while a ring buffer stores the last 5 seconds of audio, triggered for saving upon user input. The two carrier frequencies are tuned to 40 kHz and 70 kHz, and the display shows the current number of recorded files to be 15. The volume is set at 50 %.

button enables editing of the second heterodyne channel (CH2). By holding this button while rotating the encoder, users can independently adjust the carrier frequency for CH2, allowing both channels to be tuned separately for dual-band monitoring.

Listening Mode A push button toggles the dual-channel heterodyned output between two modes. In *MX mode*, signals from both microphones are mixed and played back as a combined output on both channels. True stereo monitoring is enabled in *ST mode*, routing each microphone input to its corresponding output channel.

Display Toggle. An additional *screen toggle* button turns the OLED display on or off, allowing users to reduce power consumption during extended operation or eliminate screen glare in low-light environments.

OLED Layout and Visual Feedback. The OLED interface updates dynamically in response to user input, providing clear visual feedback. It uses the `Adafruit_GFX` library with `FreeSerif` and `FreeSerifItalic` fonts for enhanced legibility. The screen layout is structured as follows:

- **Top row:** Labels and units for all displayed parameters
- **Middle row:** Channel 1 frequency (left), current adjustment mode (F: Frequency / V: Volume),

file index, and Channel 2 frequency (right)

- **Bottom row:** Output mode (MX/ST) on the left and device logo in the centre.

This layout ensures immediate visibility of all critical parameters, enabling quick and intuitive adjustments during field deployment.

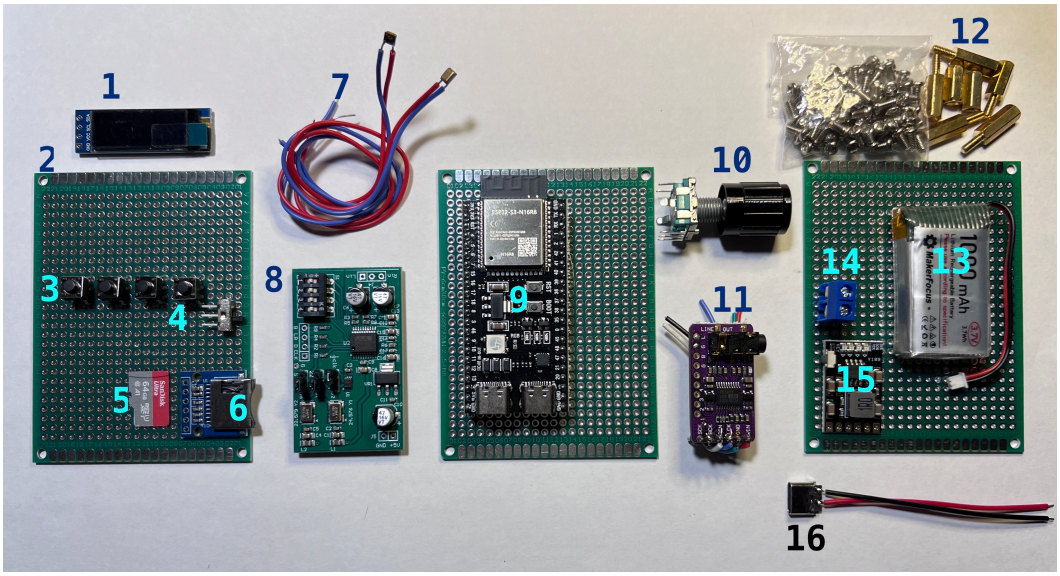

Figure 2: Components of the Esperdyne device: **(1)** 128x32 OLED Display, **(2)** Perforated PCB **(3)** Tactile push buttons (for control), **(4)** Power On/Off Switch **(5)** microSD card, **(6)** SD Card Module, **(7)** Knowles MEMS Microphones with leads (I2S input), **(8)** WM8782 I2S ADC, **(9)** ESP32-S3 main controller board with USB-C and PSRAM, **(10)** Rotary encoder with push-button, **(11)** PCM5102A stereo DAC breakout board, **(12)** Set of brass hex nuts and screws for mounting, **(13)** LiPo battery (1000 mAh), **(14)** Terminal block for battery input, **(15)** LiPo battery charging module (with protection), **(16)** USB-C Port for battery charging connection. See Appendix A for material cost and sources.

2.2 Firmware Overview

The Esperdyne firmware is implemented using the Arduino framework. The operating system manages high-speed real-time audio input/output, heterodyne demodulation, and user interaction, all optimised for responsive field use. Below, I summarise the core architecture and signal flow.

2.2.1 Signal Flow and I2S Configuration

Dual I2S peripherals are initialised for independent high-fidelity audio input and output. I2S0 is configured as the receiver, acquiring audio from external ADCs at 192 kHz with two channels. I2S1 acts as the transmitter, delivering processed audio to the DAC. The implementation logic may be found in supplementary material B.1.

2.2.2 Ring Buffer Management

A 5 s stereo ring buffer is dynamically allocated in external PSRAM to store the most recent audio samples. This enables a *tap-to-save* feature, writing the last 5 seconds of audio to a WAV file on microSD upon user command. Buffer writes occur in real time in the main audio loop. The implementation logic can be found in the supplementary material B.2.

2.2.3 User Interaction and Editable Parameters

The rotary encoder and push buttons provide intuitive control over frequency, volume, display state, recording action, and monitoring mode. Depending on input state, the encoder adjusts either the *frequency* or the *volume*, while the *channel 2 edit* button allows independent tuning. The implementation logic is provided in supplementary material B.3

2.2.4 Saving and Playback

When the record button is pressed, the firmware stops I2S input, copies the ring buffer to a WAV file using a custom header writer, then resumes streaming. File names auto-increment to prevent overwrites (`REC000.WAV` to `REC999.WAV`).

The firmware prioritises low-latency response and stability, with state-based logic ensuring no recordings are interrupted mid-stream and all inputs are debounce-guarded. Delays are added strategically (e.g., post PSRAM allocation) to avoid race conditions during startup.

A splash screen and OLED UI provide visual feedback, showing frequencies, gain, mode, file index, and playback mode in a compact 128x32 layout (see Figure 1).

2.3 I^{2}S Configuration

Both I2S0 and I2S1 are set to operate at a 192 kHz sampling rate. The master clock (MCLK) for the ADC and DAC is derived from the APLL and set to 256× the sample rate (49.152 MHz) to satisfy audio codec requirements. The I2S input is configured in standard stereo mode with 16-bit samples. Direct Memory Access (DMA) transfers data from the ADC to the PSRAM buffer without CPU intervention, ensuring minimal dropouts [22]. The output I2S stream is synchronised to the same sample rate, allowing full-duplex operation.

The ESP32-S3 microcontroller is configured to operate in full-duplex I2S mode using separate ports for input and output. The ADC (WM8782) interfaces via `I2S0` (input), and the DAC

(PCM5102A) interfaces via `I2S1` (output). The implementation overview is provided in supplementary material B.4.

2.4 Heterodyne Signal Processing and Mix Mode Playback

The Esperdyne firmware implements real-time dual-channel heterodyning on a 192 kHz I2S audio stream. Incoming ultrasonic signals from the stereo ADC (WM8782) are acquired using `I2S0` and processed in blocks of 1024 samples. Each channel is heterodyned independently by multiplying the raw audio with a synthetic carrier wave, whose frequency is user-adjustable from 10 kHz to 85 kHz in 5 kHz steps.

Each stereo sample pair's left and right channels are first normalised to a floating-point range and scaled. These are then multiplied by cosine functions with accumulating phase (`phase1` and `phase2`) calculated as:

$$\theta_n = \theta_{n-1} + 2\pi f_c / f_s$$

where f_c is the carrier frequency (per channel), and f_s is the 192 kHz sample rate. After phase wrapping and gain adjustment, the output is optionally mixed down to mono if mix mode is enabled. Both channels are averaged and sent to both stereo output channels in mix mode, producing a unified audio stream. If mix mode is off, the heterodyned left and right signals are routed to the corresponding DAC output channels.

To ensure timing coherence, the output is streamed over `I2S1` with identical sample rate and bit depth settings, allowing full-duplex operation.

Mix Mode Handling. A dedicated toggle button switches between stereo (ST) and mix (MX) modes. The state is reflected on the OLED display and applied in real-time during the audio processing loop. The corresponding logic is conditionally executed per sample block and does not introduce latency beyond buffer processing time.

2.5 ADC Configuration

The I^2S ADC (WM8782) [20] employed in this system is a 24-bit multi-bit sigma–delta converter with integrated digital anti-aliasing filters. In addition to the minimal analogue front-end conditioning, the oversampling modulator and decimation stage provide effective suppression of out-of-band components, with the internal digital filter maintaining a flat passband response up to approximately 0.45 f_s and offering more than 65 dB attenuation above 0.55 f_s [20]. This ensures that signals of interest, including ultrasonic content below the Nyquist frequency, are reliably captured

without significant aliasing artefacts. The hardware setup, including I^2S line configuration, sampling resolution, and master clock synchronisation, follows the design described for the Batsy4-Pro system [15], unless otherwise specified.

Notably, the ADC sampling rate represents the principal limiting factor of the Esperdyne system. Although the ESP32-S3 operates at a core frequency of 240 MHz and includes a dedicated audio Phase-locked loop (PLL) capable of generating master clocks up to approximately 128 MHz [23], the WM8782 constrains acquisition to lower rates. In contrast, the PCM5102A DAC can accept input rates up to 384 kHz, which remain well within the achievable APLL range. Thus, while the processing and clocking capabilities of the ESP32-S3 are sufficient for higher-fidelity acquisition and playback, the present configuration is ultimately limited by the maximum conversion rate of the ADC. Future iterations employing higher-bandwidth converters could therefore exploit the full potential of the system's clocking and processing resources.

A practical upgrade path for users requiring higher sampling rates is the direct replacement of the WM8782 with a higher-bandwidth I^2S converter such as the AK5572 (e.g., the KaamosTech AK5572-V2 module) [24]. This converter supports substantially higher sampling frequencies, upto 768kHz, while remaining compatible with the differential MEMS microphones used in the present design. In principle, the hardware modification is limited to substituting the ADC module and making minor adjustments to the I^2S wiring. However, operating at higher acquisition rates substantially increases the data throughput, and the firmware must therefore be adapted accordingly – for example by adjusting ring-buffer sizes and managing PSRAM usage to prevent premature buffer saturation. Preliminary tests conducted on the same ESP32-S3 platform demonstrate stable acquisition at sampling rates up to 250 kHz, indicating that the microcontroller can accommodate higher-bandwidth converters with appropriate firmware optimisation. A fully integrated prototype incorporating this ADC has not yet been field-tested due to seasonal unavailability of bats, but ongoing updates to the open-source repository will include support for higher-fidelity configurations as they are validated.

2.6 Power Supply and Switching

The device is powered by a 3.7 V 1000 mAh Li-ion battery connected via a charging module. The module features overcharge protection and provides USB-C recharging support with a standard 5 V DC. The system is turned on and off via a *Single Pole, Double Throw* (SPDT) toggle switch wired to the battery output. Runtime tests showed an operational lifetime of approximately 7 hours on a full charge, depending on output volume and SD card activity.

2.7 Data Processing and Curation

To streamline inspection and organisation of field recordings, I developed a custom MATLAB-based graphical interface, *Bat Reviewer* (v1.0), shown in Figure 3. Executable standalone packages for the software are provided for MacOS and Windows operating systems that can be freely downloaded and installed at no cost, with no usage restrictions [18]. The tool provides an integrated environment for browsing, listening to, and exporting ultrasonic recordings collected during field sessions. Files stored in a directory are automatically detected and listed, allowing users to navigate sequentially or by direct selection. This eliminates the need to load each recording manually and supports rapid scanning through large datasets.

The interface's central panel displays waveform and spectrogram representations of the selected file. The waveform offers a quick overview of the amplitude envelope, while the spectrogram provides a time–frequency view with configurable parameters such as FFT size, window type, and overlap. These options allow the user to adjust the spectral resolution interactively.

For auditory inspection, Bat Reviewer V1.0 incorporates a heterodyne listening function. Users may select any channel in multi-channel recordings and apply a tunable carrier frequency between 15 kHz – (0.5$\times F_s$) kHz. The heterodyned signal is automatically resampled to 44.1 kHz for playback on standard audio hardware, with an adjustable gain control to optimise listening levels. This functionality enables real-time auditory assessment of call structure, complementing the visual analysis provided by the spectrogram.

Recordings of interest can be copied directly to a user-specified destination folder via the export panel. A flexible collision-handling policy allows overwriting, automatic renaming, or prompting before replacement, ensuring that curated datasets are built consistently without accidental data loss.

Together, these features allow efficient review and curation of raw recordings, facilitating the identification of high-quality calls and reducing the time required for downstream analysis. Combining visualisation, listening, and file management in a single interface, *Bat Reviewer* provides a practical solution for sorting and preparing large volumes of acoustic data collected in the field.

3 RESULTS AND DISCUSSION

The *Esperdyne* prototype was assembled manually on a perforated PCB (Figure 1), with wires hand-soldered between corresponding pins of the components (Figure 2). The fully assembled unit weighed approximately 150 g. To protect exposed solder joints, the underside of the device and

contact points were covered with insulating tape.

Field testing was conducted along the river Isar, where bat activity was monitored and recorded over 30 minutes. Around 20 recordings were captured as bats foraged in the riparian forest during this time. The files were manually reviewed and imported into a custom application called `Bat Reviewer`, which allows browsing waveform data from a folder and selectively copying desired files to a target location. This setup enables efficient selection and curation of recordings for further analysis.

A key component of the workflow was the development of the *Bat Reviewer* interface (Figure 3), designed to accelerate the processing and curation of field recordings. While the Esperdyne device provides high-quality ultrasonic data, subsequent analysis requires rapid identification of high-SNR calls among numerous files generated in a session. The GUI addresses this by combining waveform and spectrogram displays with real-time heterodyne playback, adjustable carrier frequency, and volume control. These features allow users to cross-check spectral features visually while listening to heterodyned audio, a process familiar to field ecologists. Files of interest can be copied directly into curated folders using built-in export tools, which handle naming conflicts automatically. This integration of visualisation, auditory assessment, and file management reduces the time required for manual screening and ensures that only the most relevant recordings are retained for further quantitative analysis.

From my field session, a randomly selected recording was used to extract a sequence of five calls, presented in Figure 4 as both waveforms and spectrograms. As typical of in-field bat recordings, the signals contain both primary calls and multiple echoes—likely arising from ground and vegetation reflections. Notably, the third harmonic was prominently visible as the bat passed the recording site, suggesting a relatively close approach. Despite this, the signal did not clip, owing to the large dynamic range of the selected MEMS microphone.

The wide dynamic range of the microphone is a critical selection criterion for the intended application, particularly given the variability in call amplitude among bat species—some being "whispering" bats. In contrast, others vocalise at levels up to 130 dB SPL [25–31]. I recommend including a microphone preamplifier with adjustable gain for more versatile designs to accommodate a broader range of vocalisation intensities.

A feature to toggle the OLED display on and off was deemed beneficial during testing, as the screen's brightness can be disruptive in low-light conditions. While this feature may be a personal preference, it can be readily deactivated or customised in the firmware by extending the button press duration required to toggle it, thereby preventing accidental deactivation of the visual dis-

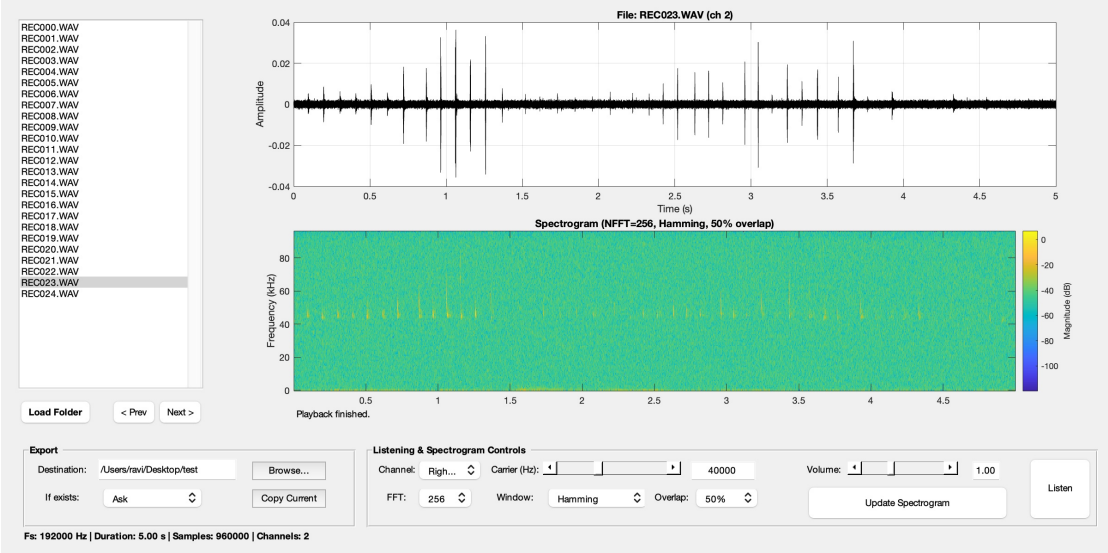

Figure 3: Bat Reviewer V1.0. The interface provides an integrated environment for browsing and visualising ultrasonic recordings. On the left, the file list allows quick navigation through multiple recordings collected in a session, with playback and export options for selected files. The main panel displays waveform (top) and spectrogram (bottom), with user-configurable FFT size, window type, and overlap. Additional controls enable heterodyne listening with adjustable carrier frequency and playback volume. This combination of visual inspection and listening facilitates rapid sorting and selection of recordings of interest from collected datasets.

play. All other functions remain operational while the display is inactive.

The MX and ST hearing modes were designed with multi-species foraging scenarios in mind, such as those observed in tropical field sites. In particular, the second-channel heterodyne mode enables clearer species separation, especially for CF bats. By default, the MX mode mirrors the signal across both channels, but can be toggled to proper stereo monitoring via a simple push button. Since some field researchers use carrier frequency tuning to estimate species identity, being able to set slightly offset heterodyne frequencies on the two channels enables estimation of upper and lower call frequency bounds—especially useful for FM bats. To the best of my knowledge, such a feature is currently not present in existing commercial detectors.

By default, the system records 5-second stereo audio clips, a choice informed by the observation that high-SNR close-range bat passes—most useful for signal analysis—typically last only a few seconds. While more distant calls are still detectable, they often lack sufficient signal quality. This constraint ensures higher-quality data capture and avoids unnecessary accumulation of low-utility recordings. The included file selection tools support rapid sorting of the acquired session data.

The onboard file counter offers a quick visual indicator of recording progress and session productivity, allowing users to gauge their data yield in real time.

Device development included detailed testing of SD card write speeds. Using the `SdFat` library and a block size of 4096 samples, a sustained write speed of $\approx$ 1.5 MB/s was achieved. SPI-based

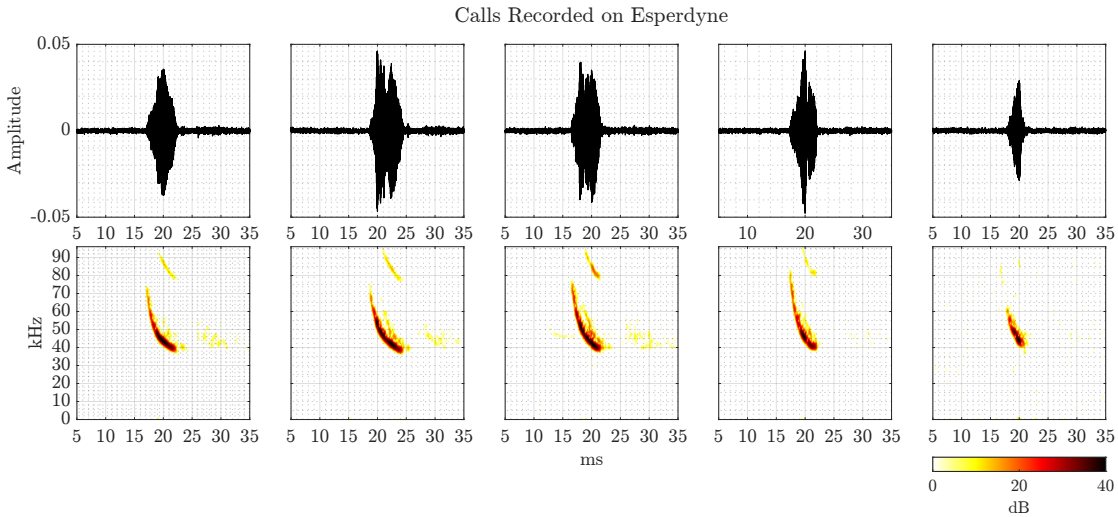

Figure 4: Field-recorded bat call sequence captured using the Esperdyne device. The data were collected in late August 2025 along the river Isar, in an area characterised by thick riparian vegetation. The waveform and spectrogram panels illustrate a sequence of echolocation calls emitted as the bat flew across the microphone. Notably, environmental echoes from the ground and vegetation are visible as smeared features following the primary call pulses, and the third harmonic of the call is distinctly captured.

SD access in the ESP Arduino core is typically not optimised for high-speed operations [32, 33], as SD cards are usually used for low-bandwidth tasks like logging text or sensor data. In this application, a 3.8 MB audio file is written in about 3 seconds. A high-speed, ultra-grade microSD card is strongly recommended to avoid data loss or file corruption.

The minimalist UI, purpose-built feature set, and specific affordances for real-time monitoring and retroactive recording make this system a unique and accessible tool. Developed on a widely available microcontroller platform, the device extends the capabilities of ESP-based systems beyond conventional audio libraries, which typically limit users to fixed sample rates optimised for consumer audio. In overcoming these limitations—as I previously did for the Teensy 4.1-based *Batsy4-Pro* [15]—this work contributes to the growing ecosystem of open-source, field-ready tools for global bioacoustic research.

Recent years have seen substantial growth in low-cost and open-source acoustic recorders. Platforms such as Solo demonstrated that inexpensive single-board computers could support flexible bioacoustic monitoring [34], while AudioMoth has become a widely adopted full-spectrum logger, offering high sampling rates, low power consumption, and weatherproof autonomous deployment [13, 35]. Other community-driven efforts, including TeensyBat [36], integrate heterodyne, frequency-division, and full-spectrum modes on more capable microcontrollers, though typically with greater assembly complexity. Smartphones paired with ultrasonic front-ends provide portable alternatives, but remain constrained by limited native bandwidth and proprietary app ecosystems [37].

Within this landscape, Esperdyne fills a distinct niche. It is a self-contained, hand-held detec-

tor that combines dual-channel tunable heterodyne monitoring with ring-buffered full-spectrum recording at 192 kHz, built entirely on open and easily reproducible ESP32-S3 hardware. Compared with devices like AudioMoth, Esperdyne lacks long-term autonomous operation and higher sampling rates, but provides real-time stereo heterodyne feedback – an uncommon feature in low-cost systems and particularly valuable for teaching and rapid field assessments.

High-end commercial detectors (e.g. Pettersson D1000x, Wildlife Acoustics Song Meter and Echo Meter Touch, Elekon Batlogger-Batcorder) remain the benchmark for low-noise, high-bandwidth, weatherproof monitoring. However, these systems are costly, closed-source, and generally not user-modifiable. Esperdyne is not intended to replace such instruments, but to offer an accessible, extensible alternative where affordability, transparency, and interactive listening are prioritised. It thus serves as a practical complement to both autonomous open-source loggers and proprietary scientific detectors, supporting broader participation in bat monitoring and acoustic education.

As a development prototype, the presented assembly may be a reference for more robust designs. Specifically, a printed circuit board (PCB) could be created to accommodate existing components or consolidate key ICs into a fully integrated, manufacturable system. All components are commercially available with published specifications, and the firmware is openly released for non-commercial applications. Institutional workshops and collaborative efforts across engineering and biology departments may use this foundation to produce reliable, field-ready hardware suitable for all-weather, all-terrain deployment.

4 CONCLUSION

Esperdyne demonstrates how low-cost, modular, and fully open-source embedded systems can be leveraged to advance bioacoustic research, especially in regions with limited access to commercial instruments. The presented system lowers the technical and financial barriers to ultrasound-based fieldwork by combining real-time heterodyne monitoring, retroactive recording, and a simple user interface on a widely available microcontroller platform. The inclusion of software tools to browse, audit, and organise field recordings adds a layer of usability critical for researchers and educators on the move. As a derivative of the broader *Embedded Ultrasonics* initiative, Esperdyne complements platforms such as Batsy4-Pro, but fills a critical niche for portable, solo-operation, cost-sensitive deployments. This work invites community-driven evolution by openly sharing the hardware design, firmware, and desktop tools—encouraging interdisciplinary collaboration, local fabrication, and field-ready innovation. Ultimately, devices like Esperdyne widen participation in ecological monitoring and conservation science and reflect a larger shift toward democratised,

decentralised scientific instrumentation for a more inclusive future in field biology.

SUMMARY OF REVISIONS

This revised version (V2) incorporates substantial improvements based on reviewer and editor feedback. The main updates are as follows:

- **Improved accessibility:** The companion software is now distributed as stand-alone executables for macOS and Windows, removing the need for a MATLAB installation and improving accessibility for researchers and educators in low-resource settings.
- **ADC upgrade pathway:** A new paragraph outlines practical options for replacing the WM8782 ADC with higher-bandwidth alternatives such as the AK5572. Preliminary tests demonstrating stable operation at sampling rates of 250 kHz and above are mentioned, together with notes on firmware implications for future versions.
- **Expanded comparative context:** The manuscript now positions Esperdyne within the broader landscape of low-cost and open-source bioacoustic recorders (e.g. AudioMoth, TeensyBat, smartphone-based detectors), clarifying its complementary niche in education, rapid field assessment, and accessible ultrasonic monitoring.
- **General text and structural improvements:** Sections of the methods, discussion, and figure captions have been streamlined for clarity and coherence. The code snippets have been moved to the *Supplementary Information*

These revisions strengthen the manuscript's clarity, accessibility, and practical relevance, while preserving its core contribution as an open, low-cost tool for ultrasonic bioacoustic monitoring.

CODE AND DATA AVAILIBILITY

The following GitHub repositories contain the firmware for Esperdyne and the Bat Reviewer tool:
1. `https://github.com/raviumadi/Embedded_Ultrasonics/tree/main/Esperdyne`
2. `https://github.com/raviumadi/Bat-Reviewer.git`
3. Umadi, R. (2025). Bat Reviewer: Version 1.0 - With Installation Packages [Computer software]. Zenodo. `https://doi.org/10.5281/zenodo.17870453`

FUNDING SOURCES

This study did not receive any external funding. All costs were borne personally by the author.

COMPETING INTERESTS

The author declares no competing interests.

ETHICAL STATEMENT

No human or animal participants were involved in the research in a manner requiring specific ethical approval.

ACKNOWLEDGEMENTS

The author thanks colleagues and supervisor Uwe Firzlaff at TUM for their support and critical feedback while developing the prototypes and this manuscript.

Manuscript Information

Version: V2
Last updated: December 10, 2025
Words in Text: 4088
Tables and Figures: 5

REFERENCES

1. Pham, L. K., Tran, B. V., Le, Q. T., Nguyen, T. T. & Voigt, C. C. Description of Echolocation Call Parameters for Urban Bats in Vietnam as a Step Towards a More Integrated Acoustic Monitoring of Urban Wildlife in Southeast Asia. *Diversity* **13**, 18. doi:10.3390/d13010018. https://www.mdpi.com/1424-2818/13/1/18 (2025) (1 Jan. 2021).
2. *Australian Bat Acoustic Data Collection* data.gov.au. https://www.data.gov.au/data/dataset/australian-bat-acoustic-data-collection (2025).
3. Redgwell, R. D., Szewczak, J. M., Jones, G. & Parsons, S. Classification of Echolocation Calls from 14 Species of Bat by Support Vector Machines and Ensembles of Neural Networks. *Algorithms* **2**, 907–924. doi:10.3390/a2030907. https://www.mdpi.com/1999-4893/2/3/907 (2025) (Sept. 2009).
4. Tabak, M. A., Murray, K. L., Reed, A. M., Lombardi, J. A. & Bay, K. J. Automated Classification of Bat Echolocation Call Recordings with Artificial Intelligence. *Ecological Informatics* **68**, 101526. doi:10.1016/j.ecoinf.2021.101526. https://www.sciencedirect.com/science/article/pii/S1574954121003174 (2025) (May 1, 2022).
5. https://www.dmws.nl. *Acoustic Ecology of European Bats* Veldshop. https://www.veldshop.nl/de/acoustic-ecology-der-europaeischen-fledermaus.html?source=facebook (2025).
6. Pijanowski, B. C. *et al.* Soundscape Ecology: The Science of Sound in the Landscape. *BioScience* **61**, 203–216. doi:10.1525/bio.2011.61.3.6. https://doi.org/10.1525/bio.2011.61.3.6 (2025) (Mar. 1, 2011).
7. Tuneu-Corral, C. *et al.* Ecological Indices in Long-Term Acoustic Bat Surveys for Assessing and Monitoring Bats' Responses to Climatic and Land-Cover Changes. *Ecological Indicators* **110**, 105849. doi:10.1016/j.ecolind.2019.105849. https://www.sciencedirect.com/science/article/pii/S1470160X1930843X (2025) (Mar. 1, 2020).
8. Bergmann, A. *et al.* The Soundscape of Swarming: Proof of Concept for a Noninvasive Acoustic Species Identification of Swarming Myotis Bats. *Ecology and Evolution* **12**, e9439. doi:10.1002/ece3.9439. https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.9439 (2025) (2022).
9. Schoeman, R. P., Erbe, C., Pavan, G., Righini, R. & Thomas, J. A. in *Exploring Animal Behavior Through Sound: Volume 1: Methods* (eds Erbe, C. & Thomas, J. A.) 217–267 (Springer International Publishing, Cham, 2022). doi:10.1007/978-3-030-97540-1_7. https://doi.org/10.1007/978-3-030-97540-1_7 (2025).

10. Arzberger, S. *et al.* The Potential of Soundscapes as an Ecosystem Monitoring Tool for Urban Biodiversity. *Journal of Urban Ecology* **11,** juaf002. doi:10.1093/jue/juaf002. https://academic.oup.com/jue/article/doi/10.1093/jue/juaf002/8088407 (2025) (Jan. 8, 2025).

11. Zamora-Gutierrez, V., MacSwiney G., M. C., Martínez Balvanera, S. & Robredo Esquivelzeta, E. in *50 Years of Bat Research: Foundations and New Frontiers* (eds Lim, B. K. *et al.*) 43–59 (Springer International Publishing, Cham, 2021). doi:10.1007/978-3-030-54727-1_3. https://doi.org/10.1007/978-3-030-54727-1_3 (2025).

12. *Ultrasound Detectors – Pettersson Elektronik AB – Bat Detectors and Sound Analysis Software* https://batsound.com/product-category/detectors/ (2025).

13. Hill, A. P., Prince, P., Snaddon, J. L., Doncaster, C. P. & Rogers, A. AudioMoth: A Low-Cost Acoustic Device for Monitoring Biodiversity and the Environment. *HardwareX* **6,** e00073. doi:10.1016/j.ohx.2019.e00073. https://www.sciencedirect.com/science/article/pii/S2468067219300306 (2025) (Oct. 1, 2019).

14. NABU, e. *Raspberry Pi Bat Projekt* Raspberry Pi Bat Projekt. http://bat-pi.eu/ (2025).

15. Umadi, R. *Embedded Ultrasonics: A Microcontroller-Based Multichannel Ultrasound Recorder for Behavioural Field Studies* https://www.biorxiv.org/content/10.1101/2025.08.11.669530v1 (2025). Pre-published.

16. Umadi, R. Widefield Acoustics Heuristic: Advancing Microphone Array Design for Accurate Spatial Tracking of Echolocating Bats. *BMC Ecology and Evolution* **25,** 92. doi:10.1186/s12862-025-02441-4. https://doi.org/10.1186/s12862-025-02441-4 (2025) (Sept. 8, 2025).

17. Umadi, R. *Array WAH: Microphone Array Design Tool* version V1.1.0. Zenodo, Aug. 2025. doi:10.5281/zenodo.15691372. https://zenodo.org/records/16922627 (2025).

18. Umadi, R. *Bat Reviewer: Version 1.0 - With Installation Packages* Zenodo, Dec. 9, 2025. doi:10.5281/zenodo.17870453. https://doi.org/10.5281/zenodo.17870452 (2025).

19. Espressif. *ESP32-S3 Series Datasheet V2.0* Espressif Systems (Shanghai) Co., Ltd, Apr. 2025. https://www.espressif.com/sites/default/files/documentation/esp32-s3_datasheet_en.pdf (2025).

20. Cirrus Logic, I. N. C. WM8782 - 24-Bit 192 kHz Stereo ADC Datasheet. *Cirrus Logic.* https://statics.cirrus.com/pubs/proDatasheet/WM8782_v4.8.pdf (2025) (2020).

21. Texas-Instruments. *PCM510xA 2.1 VRMS, 112/106/100 dB Audio Stereo DAC with PLL and 32-Bit, 384 kHz PCM Interface* (Texas Instruments, USA, 2015). https://www.ti.com/lit/ds/symlink/pcm5100a-q1.pdf?ts=1754321963376 (2025).

22. Espressif. *Inter-IC Sound (I2S) - ESP32 — ESP-IDF Programming Guide v5.5 Documentation* espressif.com. https://docs.espressif.com/projects/esp-idf/en/stable/esp32/api-reference/peripherals/i2s.html (2025).

23. Espressif. *Esp32-S3 Technical Reference Manual v1.7* June 2025. https://www.espressif.com/sites/default/files/documentation/esp32-s3_technical_reference_manual_en.pdf#iomuxgpio (2025).

24. Nihtila, T. *ADC AK5572 v2 - Very High Performance AD-converter in a Compact Form Factor. https://kaamostech.com* Kaamos Tech Ltd. https://kaamostech.com/product/adc-ak5572-v2/ (2025).

25. Waters, D. A. & Jones, G. Echolocation Call Structure and Intensity in Five Species of Insectivorous Bats. *The Journal of Experimental Biology* **198,** 475–489. doi:10.1242/jeb.198.2.475. PMID: 7699316 (Pt 2 Feb. 1995).

26. Surlykke, A. & Kalko, E. K. V. Echolocating Bats Cry Out Loud to Detect Their Prey. *PLoS ONE* **3,** e2036. doi:10.1371/journal.pone.0002036. PMID: 18446226 (2008).

27. Brinkløv, S., Kalko, E. K. V. & Surlykke, A. Intense Echolocation Calls from Two 'whispering' Bats, Artibeus Jamaicensis and Macrophyllum Macrophyllum (Phyllostomidae). *Journal of Experimental Biology* **212,** 11–20. doi:10.1242/jeb.023226. https://doi.org/10.1242/jeb.023226 (2025) (Jan. 1, 2009).

28. Schuchmann, M. & Siemers, B. M. Variability in Echolocation Call Intensity in a Community of Horseshoe Bats: A Role for Resource Partitioning or Communication? *PLOS ONE* **5,** e12842. doi:10.1371/journal.pone.0012842. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0012842 (2025) (Sept. 17, 2010).

29. Brinkløv, S., Jakobsen, L., Ratcliffe, J. M., Kalko, E. K. V. & Surlykke, A. Echolocation Call Intensity and Directionality in Flying Short-Tailed Fruit Bats, Carollia Perspicillata (Phyllostomidae). *The Journal of the Acoustical Society of America* **129,** 427. doi:10.1121/1.3519396. https://asa.scitation.org/doi/abs/10.1121/1.3519396 (2022) (Feb. 2, 2011).

30. Surlykke, A., Jakobsen, L., Kalko, E. K. V. & Page, R. A. Echolocation Intensity and Directionality of Perching and Flying Fringe-Lipped Bats, Trachops Cirrhosus (Phyllostomidae). *Frontiers in Physiology* **4,** 143. doi:10.3389/fphys.2013.00143. PMID: 23825459 (2013).

31. Yoh, N., Syme, P., Rocha, R., Meyer, C. F. J. & López-Baucells, A. Echolocation of Central Amazonian 'Whispering' Phyllostomid Bats: Call Design and Interspecific Variation. *Mammal Research* **65,** 583–597. doi:10.1007/s13364-020-00503-0. https://link.springer.com/10.1007/s13364-020-00503-0 (2022) (July 2020).

32. Espressif. *SDMMC Host Driver - ESP32-S3 - — ESP-IDF Programming Guide v5.5 Documentation* https://docs.espressif.com/projects/esp-idf/en/stable/esp32s3/api-reference/peripherals/sdmmc_host.html (2025).

33. Schatzmann, P. *SD Read and Write Speeds on an ESP32* Phil Schatzmann. https://www.pschatzmann.ch/home/2025/03/25/test/ (2025).

34. Whytock, R. C. & Christie, J. Solo: An Open Source, Customizable and Inexpensive Audio Recorder for Bioacoustic Research. *Methods in Ecology and Evolution* **8,** 308–312. doi:10.1111/2041-210X.12678. https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12678 (2025) (2017).

35. Hill, A. P. *et al.* AudioMoth: Evaluation of a Smart Open Acoustic Device for Monitoring Biodiversity and the Environment. *Methods in Ecology and Evolution* **9,** 1199–1211. doi:10.1111/2041-210X.12955. https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12955 (2025) (2018).

36. Houwertjes, E. *TeensyBat Diy Batdetector* https://www.teensybat.com/ (2025).

37. Gili, F., Bertolino, S. & Rolando, A. Using Mobile Device Built-in Microphones to Monitor Bats: A New Opportunity for Large-Scale Participatory Science Initiatives. *Biodiversity and Conservation* **33,** 1623–1643. doi:10.1007/s10531-024-02818-9. https://doi.org/10.1007/s10531-024-02818-9 (2025) (Apr. 1, 2024).

SUPPLEMENTARY INFORMATION

A BILL OF MATERIAL

Table 1: Bill of Materials for the Esperdyne System. The listed prices are for reference only and may vary depending on your source of purchase. The indicated prices correspond to costs incurred during development between May and August 2025.

Item	Description	Count	Unit Cost (€)	Source
ESP32-S3 DevKitC-1-N16R8	Microcontroller	1	7.99	Generic Manufacturer
WM8782	I2S ADC	1	14.90	Audio-phonics.fr
Knowles SPU0410LR5H-QB	MEMS Mics	2	0.50	Digikey.de
PC5102A	I2S DAC	1	7.99	Amazon.de
SSD1306	OLED Display	1	3.25	AZDelivery.de
SD Schield	File operations	1	3.50	Amazon.de
microSD Card	File storage	1	6.95	Local purchase
Push Buttons	User Input	4	0.50	Amazon.de
Rotary Encoder	User Input	1	1.50	Amazon.de
Circuit Switch	Power Supply ON/OFF	1	0.25	Amazon.de
Perforated PCB	System Base	3	1.50	Amazon.de
Terminal Connector	Battery connections	1	0.25	Amazon.de
Battery Charging circuit	Power management	1	1.50	Amamzon.de
Hexnut Screws	PCB Stacking	Assorted	5.00	Amazon.de
LiPo Battery 3.7V	Power source	1	5.99	Amazon.de
USB C Connector	Charging	1	0.10	Amazon.de
Wiring	18/22 gauge coloured wires	Assorted	5.00	Amazon.de
Total			**72.67**	

B CODE SNIPPETS

B.1 Signal Flow Configuration

```
void setupI2SInput() {
    i2s_config_t config = { /* I2S_MODE_RX, 192kHz, 16-bit */ };
    i2s_driver_install(I2S_IN_PORT, &config, 0, NULL);
    i2s_set_pin(I2S_IN_PORT, &pins);
}
void setupI2SOutput() {
    i2s_config_t config = { /* I2S_MODE_TX, 192kHz, 16-bit */ };
    i2s_driver_install(I2S_OUT_PORT, &config, 0, NULL);
    i2s_set_pin(I2S_OUT_PORT, &pins);
}
```

Listing 1: Simplified I2S input/output configuration

B.2 Ring Buffer Logic

```
ringBuffer[writeIndex] = buffer[i];
writeIndex = (writeIndex + 1) % RING_BUFFER_SIZE;
if (ringFill < RING_BUFFER_SIZE) ringFill++;
```

Listing 2: Storing incoming samples to ring buffer

B.3 User Interaction Logic

```
if (editMode == EDIT_FREQ) {
    if (ch2AdjustMode) adjust channel 2 freq;
    else adjust channel 1 freq;
} else {
    outGain = constrain(outGain + detents * VOL_STEP);
}
```

Listing 3: Frequency vs Volume control logic

B.4 I2S Configuration

```
void setupI2SInput() {
    i2s_config_t config = {
        .mode = (i2s_mode_t)(I2S_MODE_MASTER | I2S_MODE_RX),
        .sample_rate = SAMPLE_RATE,
        .bits_per_sample = BITS_PER_SAMPLE,
        .channel_format = I2S_CHANNEL_FMT_RIGHT_LEFT,
        .communication_format = I2S_COMM_FORMAT_I2S_MSB,
        .intr_alloc_flags = ESP_INTR_FLAG_LEVEL1,
        .dma_buf_count = 10,
        .dma_buf_len = 1024,
        .use_apll = true,
        .tx_desc_auto_clear = false,
        .fixed_mclk = SAMPLE_RATE * 256
    };
    i2s_pin_config_t pins = {
        .mck_io_num = I2S_IN_MCK_IO,
```

```
            .bck_io_num = I2S_IN_BCK_IO,
            .ws_io_num = I2S_IN_WS_IO,
            .data_out_num = I2S_PIN_NO_CHANGE,
            .data_in_num = I2S_IN_DI_IO
        };
        i2s_driver_install(I2S_IN_PORT, &config, 0, NULL);
        i2s_set_pin(I2S_IN_PORT, &pins);
        i2s_zero_dma_buffer(I2S_IN_PORT);
    }
```

Listing 4: I2S Input Configuration I2S0

```
    void setupI2SOutput() {
        i2s_config_t config = {
            .mode = (i2s_mode_t)(I2S_MODE_MASTER | I2S_MODE_TX),
            .sample_rate = SAMPLE_RATE,
            .bits_per_sample = BITS_PER_SAMPLE,
            .channel_format = I2S_CHANNEL_FMT_RIGHT_LEFT,
            .communication_format = I2S_COMM_FORMAT_I2S_MSB,
            .intr_alloc_flags = ESP_INTR_FLAG_LEVEL1,
            .dma_buf_count = 10,
            .dma_buf_len = 1024,
            .use_apll = true,
            .tx_desc_auto_clear = true,
            .fixed_mclk = SAMPLE_RATE * 256
        };
        i2s_pin_config_t pins = {
            .bck_io_num = I2S_OUT_BCK_IO,
            .ws_io_num = I2S_OUT_WS_IO,
            .data_out_num = I2S_OUT_DO_IO,
            .data_in_num = I2S_PIN_NO_CHANGE
        };
        i2s_driver_install(I2S_OUT_PORT, &config, 0, NULL);
        i2s_set_pin(I2S_OUT_PORT, &pins);
        i2s_zero_dma_buffer(I2S_OUT_PORT);
    }
```

Listing 5: I2S Output Configuration I2S1