

Digital Audio Player with Onboard AI Equalization

ECE 445 Project Proposal | Fall 2026

Siddharth Salapaka (svs9) • Andrew Cheng (ac158) • Advait Anand (aanand10)

1 Introduction

1.1 Problem

A fixed equalizer applies the same gains to every song, even though recordings have different tonal balances. A bass preset that helps one track can make another sound too heavy. Manual EQ gives the listener more control, but it requires choosing frequency bands and readjusting them as the music changes. We want a portable player that adjusts its EQ to both the music and a simple listening goal.

Physical controls and local storage let users listen without a phone, subscription or network connection. This avoids recurring service costs and supports use where internet access is limited. A separate player also adds battery and electronic waste, so we will use removable storage and a serviceable battery connection. Musical preferences differ across genres and cultures; users will choose a goal and adjust its strength. WHO identifies excessive sound exposure from personal audio devices as a hearing risk [1], so startup volume and output limiting are also part of the design.

1.2 Solution

We will build a battery-powered player with a microSD card, rotary encoder, buttons, OLED, stereo DAC and headphone amplifier on a custom PCB. Users select a track and choose more bass, clearer vocals, softer treble or bypass. A small neural network running on an STM32H743 uses recent spectral features to select gains for five EQ bands. Both channels use the same gains, and firmware smooths changes and limits digital peaks.

Training will happen on a computer; playback and inference will run entirely on the player. We will compare the finished device with an optimized fixed preset using a frozen tonal-balance metric. The comparison measures how closely the output follows the selected target.

1.3 Visual aid



Figure 1. The user loads music onto a card, chooses a goal on the player and listens through wired headphones. Dashed arrows show setup and charging connections.

1.4 High level requirements

- H1 — Reliable standalone playback: The player shall play stereo, 16-bit, 48 kHz PCM WAV files for at least 3 hours on battery with adaptive EQ enabled and complete a 60-minute continuous playback test with zero buffer underruns.
- H2 — Audio quality: With EQ bypassed, each output channel shall have a frequency response within ± 1 dB from 20 Hz to 20 kHz and THD+N below 0.1% at 1 kHz and 0.2 Vrms into a 32 Ω load.
- H3 — Adaptive EQ: For each of the three goals, the player shall reduce mean tonal-balance error by at least 20% versus a training-optimized fixed preset on 100 frozen, held-out test tracks, with lower error on at least 65 tracks and a one-sided Wilcoxon signed-rank test of $p < 0.01$, using the procedure in Section 2.4.

2 Design

2.1 Block diagram

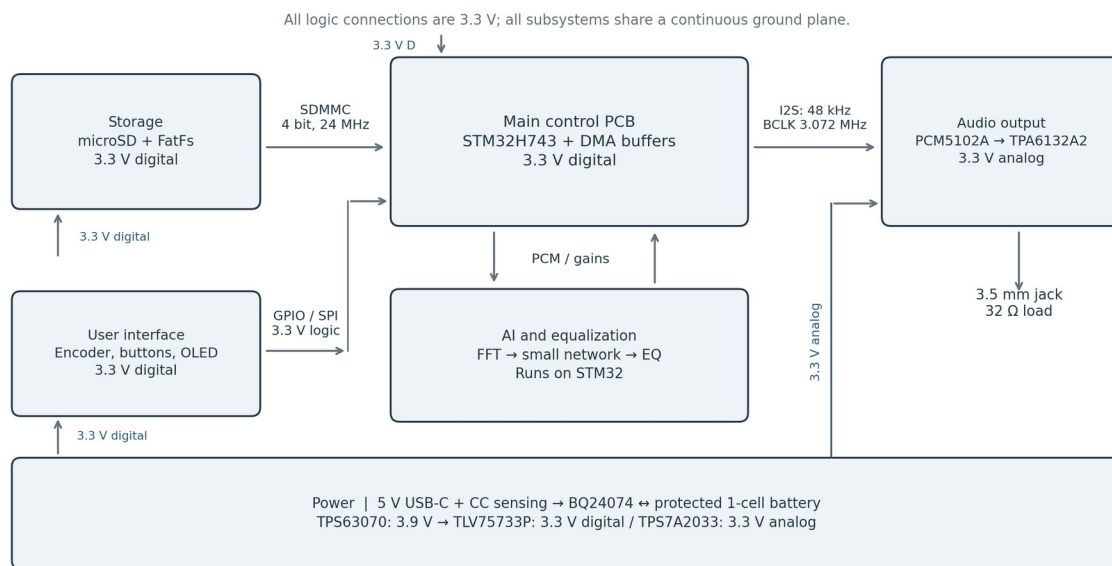


Figure 2. Six subsystems and their interfaces. The AI subsystem is firmware on the main processor. Power also supplies the storage and user interface; battery and USB status reach the MCU through ADC and GPIO inputs.

2.2 Subsystem overview

Main control PCB: The STM32H743 schedules playback, DSP and controls. DMA moves audio to the DAC while background tasks read files and update the screen. The board includes the crystal, decoupling, reset circuit, SWD and UART headers.

Storage: A microSD card and FatFs provide WAV samples to the playback buffer over SDMMC. Unsupported files produce a message instead of entering the audio pipeline.

AI and equalization: Spectral features from the original audio feed a small model. Its bounded gains control a stereo-linked five-band EQ before digital volume and peak limiting.

Audio output: The PCM5102A converts the I2S stream to analog audio. An attenuator and TPA6132A2 drive wired headphones through a 3.5 mm jack.

User interface: GPIO inputs read the wheel and buttons; SPI updates the OLED with the track, goal, volume, strength and battery status.

Power: A protected cell and USB charger feed the buck-boost regulator. Separate digital and analog LDOs supply the rest of the board, with voltage monitoring and controlled shutdown.

2.3 Subsystem requirements

The bullets below define required operating behavior and interfaces. Implementation choices and verification provisions are described in the surrounding text.

2.3.1 Main control PCB

We will use the STM32H743ZIT6 as the main processor. A four-layer board with a continuous ground plane will separate switching power from the analog audio section. These requirements support H1 and H3.

- The SAI peripheral shall output stereo I2S at 48 kHz using 3.3 V logic to the audio subsystem.
- The firmware shall process each 10 ms audio block in less than 5 ms and complete each gain update within 50 ms without delaying audio DMA.
- DMA shall read valid, current audio samples from DMA-accessible SRAM. Cache maintenance or non-cacheable mapping shall keep the output buffer coherent with the processor.

Verification provisions: UART logging will record underruns, resets, inference time and commanded gains.

2.3.2 Storage

To support H1, playback will use FAT32 microSD cards and uncompressed WAV only. This avoids codec and sample-rate conversion work during integration.

- The 3.3 V SDMMC interface shall sustain at least 0.5 MB/s, with enough buffering for at least 500 ms of stereo audio.
- The storage subsystem shall parse stereo 16-bit, 48 kHz PCM WAV headers and supply samples to the playback buffer. Invalid files and read failures shall stop that stream, mute output and return control to the track menu.

Implementation scope: The player will maintain an on-device track list. Card removal and unsupported files will return the user to the menu with an error message.

2.3.3 User interface

An encoder with a push switch, four buttons and a 128 × 64 SSD1306 OLED provide standalone operation for H1. The buttons handle play/pause, previous, next and bypass; the wheel selects tracks, volume, goal and strength.

- Inputs shall use debounced 3.3 V GPIO, and the display shall update over SPI with visible feedback within 100 ms of a control action.
- All three goals, strength control and bypass shall be accessible without a computer, and boot volume shall start low.

2.3.4 Power

The BQ24074 charger, TPS63070 buck-boost, TLV75733P digital LDO and TPS7A2033 analog LDO form the power path [5–8]. We propose a protected single-cell battery, subject to its datasheet and course battery review. This supports H1 and keeps switching noise away from H2.

- The buck-boost shall provide the intermediate rail needed by the LDOs. The digital rail shall provide 3.3 V ± 0.10 V at 500 mA continuously, and the analog rail shall provide 3.3 V ± 0.05 V at 100 mA continuously.
- USB-C charging shall use a 5 V input and draw no more than 500 mA. Charge current and temperature limits shall follow the selected cell and charger datasheets, and the design shall not assume higher source current without detecting it [5, 10].
- The battery monitor shall provide a low-battery warning and controlled shutdown before the hardware protection cutoff. Independent undervoltage protection and NTC-based charge inhibition shall follow the selected cell limits.

2.3.5 AI and equalization

This subsystem supports H3 within the playback deadlines in H1. CMSIS-DSP will compute spectral band energies from recent unprocessed stereo audio. The model will use a short sequence of these spectral features covering about two seconds of past audio, followed by a one-hot code for the selected goal.

A small temporal neural network will output five EQ gains and run on the Cortex-M7 using an embedded inference format. Gains are bounded to ± 6 dB and constrained by the selected goal. Because the five filters overlap, adjusting one gain affects several measurement bands. The model approximates the jointly fitted mapping from an offline optimizer; Section 2.4 also compares it with a direct rule-based controller.

The EQ will use RBJ biquad formulas [9]. Shelf filters use slope $S = 1$; peaking filters use $Q = 1$. Frequencies stay fixed, and both stereo channels share the gain settings.

Band	Filter	Measurement region
Low	100 Hz low shelf	20–150 Hz
Low-mid	300 Hz peak	150–500 Hz
Mid	1 kHz peak	500 Hz–2 kHz
Presence	3.2 kHz peak	2–5 kHz
Treble	8 kHz high shelf	5–16 kHz

For each non-silent training segment, we subtract the mean of its five band levels in dB. The bandwise median is then recentered to zero mean to form reference R. Each target adds the offsets below and is recentered again.

Goal	Offsets from R	Sign constraint
More bass	Low +4 dB; low-mid +1.5 dB	Low gain ≥ 0 dB
Clearer vocals	Low-mid -2 dB; mid +1 dB; presence +3 dB	Low-mid ≤ 0 dB; presence ≥ 0 dB
Softer treble	Presence -1.5 dB; treble -4 dB	Treble gain ≤ 0 dB

Training and validation data will come from the broader FMA dataset [11] using artist-disjoint train, validation and test splits across several genres. The final graded test set will contain 100 frozen, held-out tracks. An offline optimizer may use a centered audio window, including future samples, to generate fitted gain labels; the deployed model remains causal and uses only past audio. Track IDs, licenses, splits, targets and the metric will be frozen before final fitting.

- The model shall update gains at 10 Hz and finish inference within 50 ms. A missed update shall hold the previous gains, and invalid outputs shall ramp to bypass.
- Gain changes shall be smoothed to avoid audible steps. Headroom management and a look-ahead limiter shall keep digital output peaks at or below -1 dBFS, and silence shall hold the current gains.

2.3.6 Audio output

The PCM5102A converts the I2S stream to analog audio [3]. The TPA6132A2 headphone amplifier drives the 3.5 mm output [4]. A resistive attenuator will be selected so the DAC cannot overdrive the amplifier, and digital volume will be applied before I2S output.

- Each channel shall meet H2 into a $32\ \Omega$ load and provide at least 0.50 V_{rms} before clipping.
- DAC mute and amplifier enable shall be sequenced during startup, shutdown and track changes to prevent audible pops. Analog coupling and output filtering shall follow the device reference circuits.

2.4 Verification of adaptive EQ

The final comparison will use 100 frozen, artist-disjoint test tracks at full EQ strength. For each goal we will compare the neural controller against one constant five-band preset fitted on the training split using multiple optimizer starting points. Both conditions use the same gain bounds, sign constraints, filters and limiter.

For each non-overlapping three-second output segment, let L be its five band-energy levels in dB. Subtract their mean to get relative tonal shape. Segment error is the root-mean-square distance between that shape and the zero-mean target T . Energy is averaged across channels, and silent segments are excluded using a fixed threshold chosen before final testing.

$$e = \text{RMS across five bands of } [(L - \text{mean}(L)) - T]$$

We will discard the startup portion, average segment errors within each track, and give every track equal weight. H3 requires at least a 20% reduction in mean error, lower error on at least 65 of the 100 tracks for each goal, and a one-sided Wilcoxon signed-rank test with $p < 0.01$. The same filters, gain bounds, sign constraints and limiter are used for both controllers.

The graded test set will be processed on the player using the same DSP and model path as normal playback. A smaller fixed subset will also be captured from the headphone output into 32 Ω loads using a stereo audio interface to confirm that the analog output follows the digital evaluation. We will measure the interface noise floor and bypass response and apply the same correction to every condition.

Host simulation will also compare the model with a causal rule-based controller that maps band error directly to bounded gains. Validation tests will confirm that the device implementation matches the host model, that the applied gains respond when the spectral balance changes, and that live playback follows the same controller behavior as the graded evaluation path.

Before final testing, we will verify goal and bypass transitions, H2 audio quality, playback timing and battery runtime under fixed documented conditions.

2.5 Tolerance analysis

The main hardware risk is the power budget: an undersized cell or insufficient LDO headroom could cause early shutdown or audio interruptions. The following calculations are design estimates, not measured results.

Battery runtime

We allocate an average load of 350 mA at 3.3 V, including the processor, SD card, display, audio and margin. A 2,000 mAh cell has 7.4 Wh nominal energy. Assuming only 80% is usable before shutdown and 75% reaches the 3.3 V rails gives:

$$P_{\text{load}} = 3.3 \times 0.350 = 1.155 \text{ W}$$

$$t_{\text{run}} = (3.7 \times 2.0 \times 0.80 \times 0.75) / 1.155 = 3.84 \text{ hours}$$

With a 20% increase in average load, the estimate falls to 3.20 hours. For the three-hour requirement, the maximum average 3.3 V load is 1.48 W, or about 448 mA, under these assumptions. We will measure load, conversion efficiency and usable capacity on the first board.

Regulator headroom and heat

The 3.9 V intermediate rail provides headroom for the digital LDO. At -2%, its input is 3.822 V. Allowing +1% on the 3.3 V output leaves 0.489 V of headroom. The TLV757P specifies up to 0.425 V dropout at 1 A through 85 °C for the selected SON package [7], leaving about 64 mV even at that higher current. Our continuous digital requirement is 500 mA.

At 500 mA, the digital LDO dissipates approximately $(3.9 - 3.3) \times 0.5 = 0.30 \text{ W}$. At a 350 mA total average load, both LDOs together dissipate roughly 0.21 W. We will use the thermal-pad package with ground copper and measure surface temperatures under worst-case playback and charging. The charger may reduce current thermally; fast charging is not a requirement.

3 Ethics safety and societal impact

3.1 Honest evaluation and responsible use

ACM Code of Ethics Sections 1.2 and 1.3 address avoiding harm and honest claims; Section 2.5 addresses thorough system evaluation [12]. We will keep the test set separate and report measured results against each requirement. The tonal metric measures agreement with a chosen curve, not listener preference or improved hearing. The clearer-vocals mode adjusts the whole mix rather than separating the singer.

Sections 1.5 and 1.6 address intellectual work and privacy [12]. We will record the license and attribution for each FMA track, use only material whose terms permit the planned processing, and avoid redistributing source audio. Music stays on the player, and normal operation will not collect listening history or upload data. Any later listening study would be optional, voluntary and discussed with course staff before recruitment.

3.2 Electrical battery and hearing safety

Early tests will use a current-limited bench supply. Before battery testing, we will complete the required course training and obtain the battery-circuit review required by the ECE 445 battery policy. That policy also requires additional fire safety training, lithium safety bags for charge/discharge tests, and protection against operation outside the allowed cell voltage range [13, 14]. We will not work alone in the lab.

The final device will use a protected, documented cell, an independent undervoltage cutoff, charge-temperature sensing, insulated wiring and a keyed battery connector. We will check connector polarity before assembly, keep the cell away from sharp edges and hot components, and stop using any damaged or swollen cell. Short-circuit tests will use a current-limited supply or battery simulator rather than deliberately shorting the cell. NTC faults will be tested using substitute resistances.

The digital limiter prevents overload but does not guarantee a safe sound-pressure level with arbitrary headphones. Startup volume will be low, and the interface will warn before high-volume operation. Electrical tests will use dummy loads, and listening checks will begin at low volume and use short sessions. WHO guidance on volume limits and exposure awareness informs these choices [1].

3.3 Standards and regulations

IEC 62368-1 covers safety of audio/video equipment, including energy sources and safeguards [15]. It will guide our review of accessible conductors, heating, insulation and the enclosure. For the purchased cell, we will seek supplier evidence for IEC 62133-2 with its 2021 amendment, covering portable lithium-cell safety [16]. Formal product certification is outside this prototype's scope.

The high-speed circuitry is a potential unintentional radiator. A future marketed product would need evaluation under the applicable FCC Part 15 rules, including the radiated-emission limits in 47 CFR 15.109 [17]. We will use ground planes, short clock traces and local decoupling to reduce coupling. USB-C charging will follow the source-current advertisement limits [10].

3.4 Societal and environmental impact

Offline operation avoids recurring service costs and network dependence, but a separate player still uses materials and can become electronic waste. We will use removable storage, a serviceable battery connection and a screw-fastened enclosure. Illinois lists portable digital music players among devices covered by its electronics recycling program [18]; unwanted boards and batteries will go through appropriate campus collection routes. Since musical preferences differ across genres and cultures, we will report test-set coverage, provide a strength control and keep bypass readily available.

4 Development plan

The schedule uses the supplied Fall 2026 milestones [19]. The target is a working custom-board system by the November 2 progress demo and a feature freeze before the November 16 mock-demo week. Fall break is not part of the required build time. Board delivery dates remain estimates, so development-board firmware work will continue while fabrication is in progress.

Date or week	Planned work and checkpoint
Sep 21–25	Freeze interfaces and parts; test SD-to-DAC playback on a development board; prepare the PCB review and team contract. Use a simple enclosure; consult the machine shop by Sep 21 only if needed.
Sep 28–Oct 2	Target the Sep 29 first PCB order after audit. Finish the design document for Oct 2. Freeze data splits and the EQ metric; run the fixed-preset and simple-controller baselines.
Oct 5–9	Complete the Oct 5–7 design review. Train/export the small network and benchmark it on the MCU. Oct 8 is a contingency first-order date if the board misses Sep 29.
Oct 12–16	Show playback, controls and changing EQ on development hardware at the breadboard demo. Validate that each mode has margin over H3. Assemble the PCB when delivered; identify revisions.
Oct 19–30	Bring up power first, then MCU, storage, audio and UI. Submit any necessary revision by the Oct 29 final PCB round; do not plan on a later board order.
Nov 2–6	Demonstrate the integrated custom PCB. Measure audio response, inference deadlines and current. Complete the progress report; fix core failures before adding features.
Nov 9–13	Run battery, temperature, transition and held-out EQ tests. Finish the enclosure and freeze firmware. Preserve logs and recordings for the report.
Nov 16–20	Mock demo and presentation. Use this week for fixes, repeat only affected tests, and rehearse the final demo.
Nov 23–27	Fall break; no required lab work scheduled.
Nov 30–Dec 9	Final demo Nov 30–Dec 2; presentations Dec 3–4; complete the final report for Dec 9.

4.1 Work distribution and scope limits

The team will divide primary responsibility among hardware and power, embedded playback and controls, and model/DSP evaluation. Interfaces, PCB review and integration tests will be shared; final names for each role will be recorded in the team contract.

The required build has one sample rate, WAV playback, three listening goals and a small temporal model. Text entry, natural-language interpretation, FLAC, 44.1 kHz, wireless features, a substantially larger model and on-device training are outside the core scope. A larger dataset and listening study are optional only after the required tests pass.

If the model misses the October validation checkpoint, we will diagnose features and labels using validation data and discuss any requirement change with the TA before final testing.

References

[1] World Health Organization, “New WHO–ITU standard aims to prevent hearing loss,” 2019. [Online]. Available: <https://www.who.int/news/item/12-02-2019-new-who-itu-standard-aims-to-prevent-hearing-loss-among-1.1-billion-young-people>

- [2] STMicroelectronics, "STM32H742xl/G and STM32H743xl/G," datasheet. [Online]. Available: <https://www.st.com/resource/en/datasheet/stm32h743zi.pdf>
- [3] Texas Instruments, "PCM510xA Audio Stereo DAC," datasheet. [Online]. Available: <https://www.ti.com/lit/ds/symlink/pcm5102a.pdf>
- [4] Texas Instruments, "TPA6132A2 Stereo Headphone Amplifier," datasheet. [Online]. Available: <https://www.ti.com/lit/ds/symlink/tpa6132a2.pdf>
- [5] Texas Instruments, "BQ2407x Standalone 1-Cell Linear Battery Charger with PowerPath," datasheet. [Online]. Available: <https://www.ti.com/lit/ds/symlink/bq24074.pdf>
- [6] Texas Instruments, "TPS63070 Buck-Boost Converter," datasheet. [Online]. Available: <https://www.ti.com/lit/ds/symlink/tps63070.pdf>
- [7] Texas Instruments, "TLV757P 1A Low-Dropout Regulator," datasheet, Rev. C. [Online]. Available: <https://www.ti.com/lit/ds/symlink/tlv757p.pdf>
- [8] Texas Instruments, "TPS7A20 300mA Low-Noise LDO," datasheet. [Online]. Available: <https://www.ti.com/lit/ds/symlink/tps7a20.pdf>
- [9] R. Bristow-Johnson and R. Toy, "Audio EQ Cookbook," W3C Working Group Note, 2021. [Online]. Available: <https://www.w3.org/TR/audio-eq-cookbook/>
- [10] USB Implementers Forum, "USB Type-C Cable and Connector Specification," Release 2.5, 2026. [Online]. Available: <https://www.usb.org/document-library/usb-type-cr-cable-and-connector-specification-release-25>
- [11] M. Defferrard, K. Benzi, P. Vandergheynst and X. Bresson, "FMA: A Dataset for Music Analysis," ISMIR, 2017; dataset repository. [Online]. Available: <https://github.com/mdeff/fma>
- [12] Association for Computing Machinery, "ACM Code of Ethics and Professional Conduct," 2018. [Online]. Available: <https://www.acm.org/code-of-ethics>
- [13] University of Illinois, ECE 445, "Safety Guidelines." [Online]. Available: <https://courses.grainger.illinois.edu/ece445/guidelines/safety.asp>
- [14] University of Illinois, ECE 445, "Safe Practice for Lead Acid and Lithium Batteries," 2016. [Online]. Available: <https://courses.grainger.illinois.edu/ece445/documents/GeneralBatterySafety.pdf>
- [15] International Electrotechnical Commission, "IEC 62368-1:2023 — Audio/video, information and communication technology equipment — Safety requirements." [Online]. Available: <https://webstore.iec.ch/en/publication/85813>
- [16] International Electrotechnical Commission, "IEC 62133-2:2017+AMD1:2021 — Safety requirements for portable sealed secondary lithium cells and batteries." [Online]. Available: <https://webstore.iec.ch/en/publication/70017>
- [17] Federal Communications Commission, "47 CFR 15.109 — Radiated emission limits." [Online]. Available: <https://www.ecfr.gov/current/title-47/chapter-I/subchapter-A/part-15/subpart-B/section-15.109>
- [18] Illinois Environmental Protection Agency, "Electronics Recycling," Consumer Electronics Recycling Act program. [Online]. Available: <https://epa.illinois.gov/topics/waste-management/materials-management/electronics-recycling1.html>
- [19] University of Illinois, ECE 445, "Calendar," Fall 2026; supplied course calendar. [Online]. Available: <https://courses.grainger.illinois.edu/ece445/calendar.asp>