

Digital Audio Player with Onboard AI Equalization

ECE 445 Design Document | Fall 2026 | Project 57

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

Design baseline dated October 2 2026

1 Introduction

A fixed equalizer applies the same gains to every recording, although recordings differ in tonal balance. A bass preset that helps a thin track can make an already bass-heavy track excessive. Manual EQ requires the listener to choose frequency bands and readjust them as the music changes. We will build a portable player that adapts a five-band equalizer to the music and a listening goal selected through physical controls.

The player reads music from microSD and runs a small neural controller on an STM32H743. It operates without a phone, subscription, or network connection. The user chooses more bass, clearer vocals, softer treble, or bypass. Our improvement over a fixed preset is evaluated as agreement with a defined tonal target; it is not a claim of universal listener preference. Both stereo channels share the same gains.

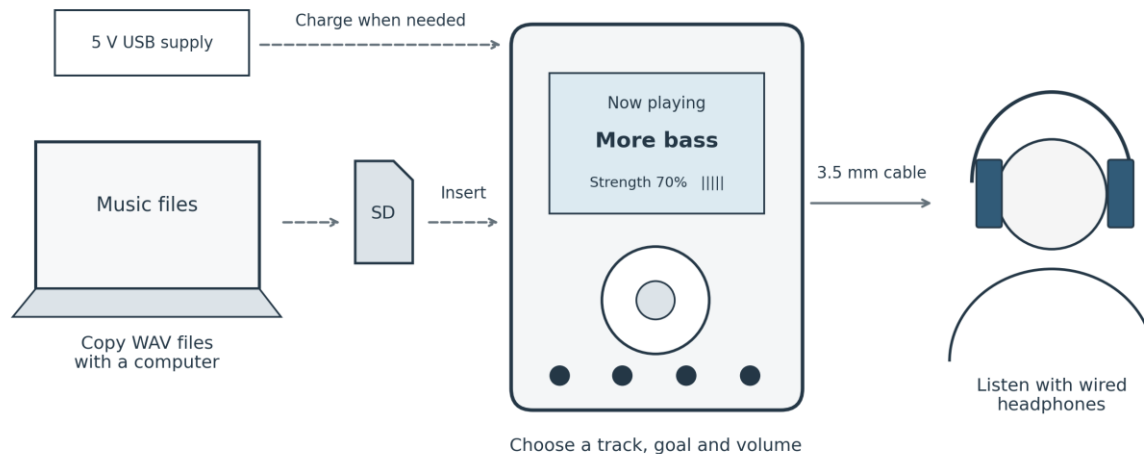


Figure 1. Use of the standalone player. The computer prepares the card; it is not required during playback.

High level requirements

- H1. 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. 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. 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 $p < 0.01$.

2 Design

System architecture

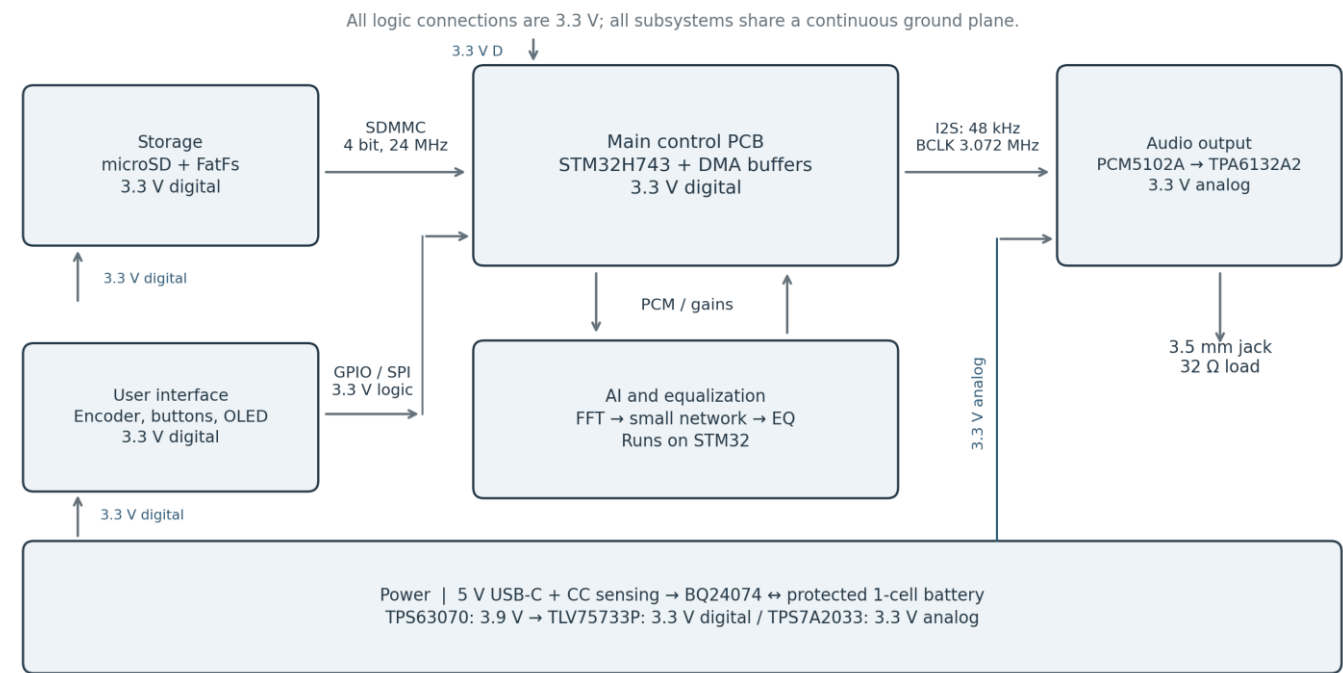


Figure 2. Six subsystem boundaries and their main electrical interfaces. The AI subsystem executes on the main processor.

Main control schedules audio and coordinates the other blocks. Storage supplies PCM; AI produces bounded gain commands; audio output converts the processed stream to a headphone signal. The user interface changes playback and listening settings. Power supplies separate digital and analog rails. DMA decouples audio transmission from file reads, inference, and display work.

Interface	Design basis
Storage to MCU	3.3 V SDMMC, four data lines, command and clock; initial clock ≤ 400 kHz, playback target 24 MHz. FAT32 only.
MCU to DAC	I2S stereo, 48 kHz LRCLK, two 32-bit slots per frame; BCLK = 3.072 MHz. 3.3 V logic.
MCU to OLED and controls	4-wire SPI at 4 MHz; 3.3 V GPIO for buttons, encoder, reset and chip select.
AI to DSP	Five float gain commands in dB, one update per 4,800 samples; shared across channels.
Power to loads	Digital 3.3 V ± 0.10 V at 500 mA; analog 3.3 V ± 0.05 V at 100 mA.
Debug and evaluation	SWD, 3.3 V UART at 115,200 baud, GPIO timing markers, sample-indexed event logs.

The required scope is WAV playback at one sample rate, three listening goals, and onboard inference. FLAC, 44.1 kHz playback, wireless access, text-entry control, on-device training, and a human listening study are optional extensions after the required verification is complete.

Physical design and assembly

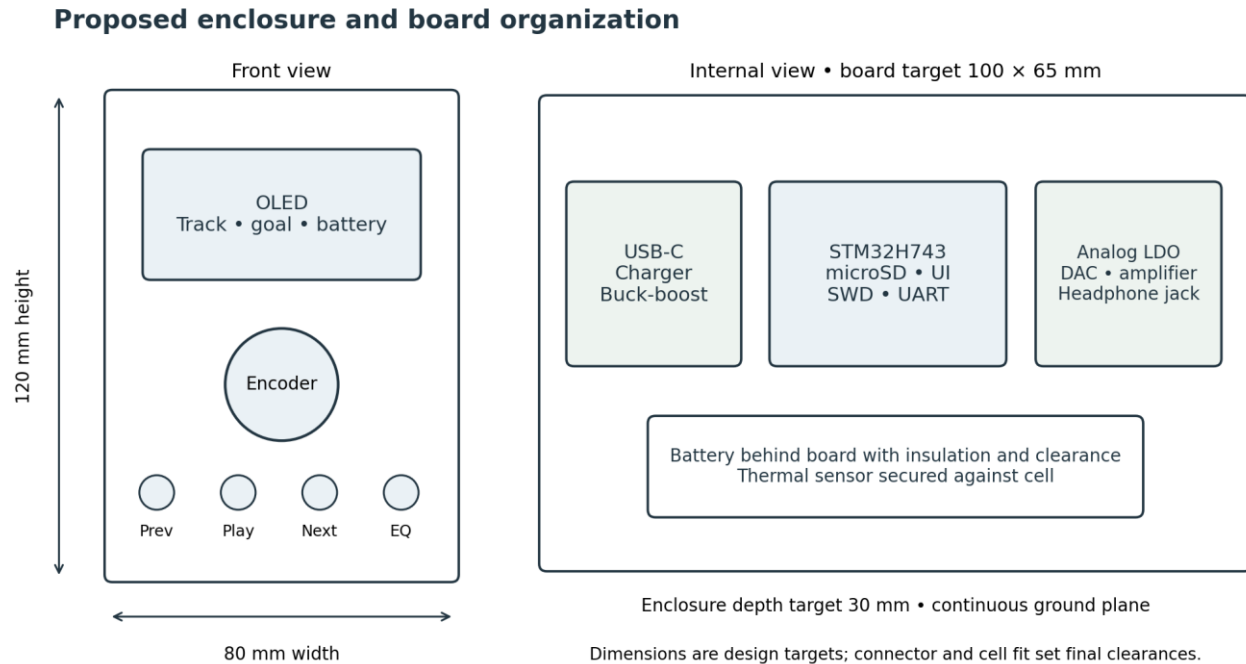


Figure 3. Proposed mechanical envelope and functional placement. This dimensioned concept defines packaging intent and is not a routed PCB drawing.

The starting enclosure target is 120 × 80 × 30 mm, with a 100 × 65 mm four-layer PCB, a front display, an encoder, and four buttons. USB-C, microSD, and the headphone jack remain accessible from the edges. A screw-fastened back allows battery service. These dimensions reserve space for the selected cell, wiring bends, standoffs, and an insulating barrier; the footprint and enclosure fit review determines the release dimensions.

Place the charger, buck-boost inductor, and high-current switching loops near USB-C and away from the DAC and amplifier. Keep the analog chain near the headphone jack. A continuous ground plane supports short return paths; separate analog and digital supply routing does not mean splitting the ground plane. Do not route SD, display, or clock traces through the analog area. Add local decoupling at every supply pin and ground vias at the exposed thermal pads. [3–9]

Board connections and bring up

Use ICs and their support components on the custom PCB. The display and controls may connect through short headers. Development boards support early tests only. Provide test points for battery, charger output, 3.9 V, both 3.3 V rails, LRCLK, BCLK, DAC output, amplifier output, and ground. Keep SWD and UART accessible after enclosure assembly.

Bring up one rail at a time using a current-limited supply. Confirm polarity, rail resistance, voltage, and clock activity before enabling audio. Then verify SWD, storage, DAC tones, controls, and inference in that order. Assemble the cell only after the charger and protection review. The fabrication package must contain a pin-level schematic, footprint checks, ERC/DRC results, and a BOM reconciled to the PCB; the connection diagrams here are the engineering basis for that package.

Main control and real time firmware

The STM32H743ZIT6 provides a Cortex-M7, internal flash and SRAM, SAI, SDMMC, timers, and debug access. It supports the local DSP and inference needed for H1 and H3. A crystal-based clock is preferred to an untrimmed internal oscillator for the ± 100 ppm audio-clock target. Use a 25 MHz HSE crystal, short oscillator traces, and load capacitors calculated from the chosen crystal and board parasitics. Follow the manufacturer supply, VCAP, reset, and boot connections. [3]

Function	Proposed MCU allocation	Connection
SDMMC1	PC8–PC12 and PD2	D0–D3, CK and CMD; AF12. Card detect uses a spare GPIO.
SAI1 block A	PE4, PE5, PE6	FS, SCK, SD; AF6. No external DAC master-clock line.
OLED SPI2	PB13, PB15; PB12, PD3, PD4	SCK, MOSI; GPIO CS, DC, RESET.
Controls	PA0, PA1, PA2; PC0–PC3	Encoder A/B and push; four buttons.
Battery ADC	PA3	Divided battery voltage; select supported ADC input.
USB CC status	PB8, PB9	I2C1 SCL/SDA, AF4, with 3.3 V pull-ups.
Audio enable	PE7, PE8	DAC XSMT and amplifier EN, held inactive during reset.
Debug	PA13, PA14; PD8, PD9	SWDIO/SWCLK; USART3 TX/RX, AF7.
Power management	PG0–PG4	Charger mode, charge permit, status and shutdown control.

This is a proposed allocation for the LQFP144 device. Validate all alternate functions, reset states, boot/debug reservations, and remaining power-status inputs together in the STM32 configuration before schematic release. [3]

Scheduling and memory

Audio processing runs every 480 stereo frames (10 ms) at the highest task priority below essential interrupts. SD reads refill a 128 KiB input ring in a lower-priority task. Inference runs every 100 ms and is preemptible by audio; display and UART work have still lower priority. Interrupts acknowledge DMA and signal work without running inference or blocking on files. The timing measurement includes feature extraction and coefficient updates in whichever task executes them.

Allocate at least 40 ms of output queue, or 15,360 bytes at two 32-bit slots per frame. Keep DMA buffers aligned to cache-line boundaries in DMA-accessible SRAM, not DTCM. Use non-cacheable MPU regions or explicit clean/invalidate operations with clear buffer ownership. Logging uses a bounded queue; logging overflow is counted and never blocks audio.

Firmware data flow and verification

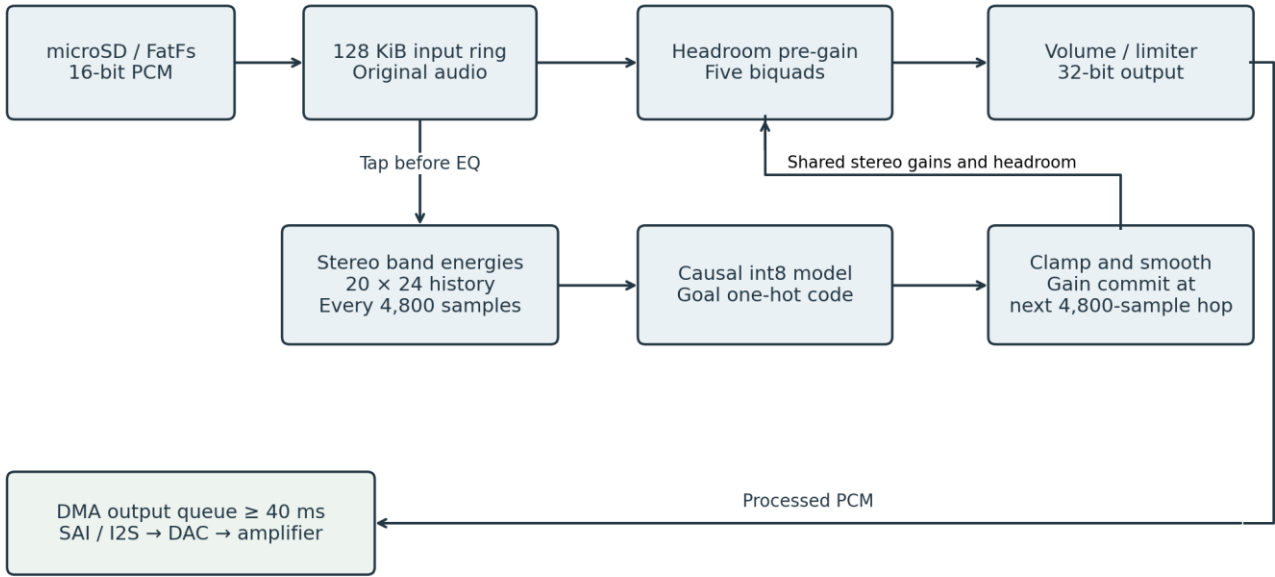


Figure 4. The feature branch observes original audio. Inference changes future EQ gains at fixed sample boundaries while the output queue continues playing.

All DSP state is indexed by sample count. A feature frame completed at hop k produces a gain command scheduled for hop $k + 1$. Reset filter, limiter, history, and model state identically at each evaluation track. Offline mode follows the same deadlines logically and uses the same DSP code, coefficients, quantized weights, and output conversion. This makes live-versus-offline comparison meaningful.

ID and role	Requirement	Verification and pass criterion
MCU1 H1	Stereo I2S at 48 kHz within ± 100 ppm.	Measure LRCLK over a ≥ 10 s gate using a frequency counter with ≤ 20 ppm uncertainty. Repeat during SD, display, and inference activity. Pass: 47,995.2–48,004.8 Hz, with measurement uncertainty included.
MCU2 H1 H3	Each 10 ms audio block < 5 ms; each inference < 50 ms.	Cycle-counter and GPIO timestamps over the full 60-minute run: 360,000 blocks and 36,000 scheduled inferences. Report maxima. Pass: every observed block and inference below its limit; zero missed scheduled updates.
MCU3 H1	Output buffer capacity ≥ 40 ms; coherent DMA data.	Inspect linker map and buffer sizes. Compare known-pattern transfers with cache enabled and collect underrun counters during playback. Pass: capacity $\geq 15,360$ bytes for the chosen format, no stale/corrupt blocks, and zero underruns.

Storage and user interface

Storage design

FAT32 and uncompressed stereo 16-bit 48 kHz PCM WAV keep the playback path small and predictable. The reader parses RIFF chunks rather than assuming a fixed 44-byte header; it checks the format before queueing audio. A 128 KiB read-ahead ring holds 0.683 s at 192,000 bytes/s, exceeding the 500 ms requirement. Read 8–32 KiB sequential chunks when space is available. The measured 0.5 MB/s throughput target is 2.60 times the playback rate. [3]

The microSD socket connects directly to 3.3 V SDMMC. Use CMD/D0–D3 pull-ups and local decoupling; begin with 47 kΩ pull-ups, 100 nF plus 10 μF at the socket, and optional source-series resistor footprints on the clock. The clock has no pull-up. The initial library target is 100 tracks with names displayed by scrolling. Invalid files, read failure, or removal stop that stream, mute output, and return to the menu. Offline evaluation uses separate output files and may run slower than playback.

User interface design

Four buttons control play/pause, previous, next, and bypass. The encoder scrolls lists or changes the selected value; a push selects the highlighted item or exits an adjustment. The now-playing screen offers track, goal, volume, strength, and battery status. The encoder accesses a short menu for goal and strength, so all functions remain available without a computer. The 128 × 64 SSD1306 module uses 4-wire SPI at 4 MHz. [13]

Use 10 kΩ pull-ups to 3.3 V on the encoder and buttons. Sample at 1 kHz, reject invalid quadrature transitions, and debounce switches for 5 ms. Schedule screen refresh at 20 Hz; a 1,024-byte framebuffer transfer needs about 2.05 ms at 4 MHz before command overhead. This leaves margin within 100 ms, but the verification measures the visible result, including debounce, scheduling, transfer, and panel refresh.

Boot and reset set volume to –40 dB relative to maximum. A warning precedes settings above –10 dB, while the user retains control. These electrical settings do not establish a safe acoustic level for every headphone. Normal bypass uses a ≥50 ms crossfade between aligned dry and wet paths; for demonstration, use a frozen level-matching calibration so loudness alone does not dominate the comparison. [21]

State	Entry and behavior
BOOT	Mute amplifier; initialize rails, card, settings and buffers; start at low volume.
MENU or PAUSED	Show track/settings; retain standalone control; output silence.
PLAYING	Refill buffers, process PCM and accept controls while DMA transmits.
ERROR	Mute, stop failed stream, show cause, and permit another selection.
SHUTDOWN	Ramp audio down, mute hardware, close files, and disable the main regulator.

Storage and interface verification

ID and role	Requirement	Verification and pass criterion
STO1 H1	Effective sequential read rate ≥ 0.5 MB/s.	Read a 64 MiB test file through the production FatFs/SDMMC path, five times from freshly opened files. Use elapsed wall time and MB = 10^6 bytes. Pass: every run ≥ 0.5 MB/s. Log the longest individual read stall.
STO2 H1	Read-ahead capacity ≥ 500 ms and zero storage underruns.	Inspect capacity and run the H1 playlist on the selected card with normal UI/inference activity. Pass: $\geq 96,000$ input bytes, and zero storage-caused underruns during 60 minutes. Capacity alone does not prove underrun freedom.
STO3 H1	Supported files play; malformed/unsupported files fail safely.	Use valid WAVs with extra RIFF chunks plus mono, 44.1 kHz, 24-bit, truncated, and malformed cases. Remove the card once in a controlled playback test. Pass: supported files play; other cases mute and return to the menu without reset.
STO4 H1	Browse and select at least 100 tracks.	Load 100 supported tracks with short and long filenames. Select first, middle and last entries and perform previous/next at boundaries. Pass: correct file selected and menu remains responsive.
UI1 H1	Visible response < 100 ms from physical input.	Record the electrical button edge and a photodiode over a toggled OLED test patch with an oscilloscope. Test 20 events per button and encoder action while playing. Pass: every edge-to-visible transition < 100 ms. A 240 fps video is a fallback only if its uncertainty still proves the bound.
UI2 H1	No missed/double encoder counts through 3 detents/s.	Rotate 60 detents in each direction at 1 and 3 detents/s; compare physical detents with the decoded count. Pass: all counts and directions match.
UI3 H1	All controls work standalone; low startup volume.	Run the complete boot, select, play, pause, previous/next, goal, strength and bypass sequence. Repeat ten power cycles. Pass: each action works and every boot begins at -40 dB.

Recorded evidence

Each verification record identifies board revision, firmware hash, card model, supply condition, instrument, calibration, raw log location, and pass/fail result. Use the same firmware build for integration and evaluation wherever possible. A change affecting a verified interface triggers the relevant test again, with both results retained.

AI and equalization design

The controller supports H3 by varying the correction with the current recording. Every 4,800 samples, compute a 4,096-sample Hann-windowed spectrum from each channel and average the spectral powers. This stereo-energy feature avoids cancellation from out-of-phase channels. Integrate 24 log-spaced bands from 20 Hz to 16 kHz and apply log scaling and training-set normalization. Twenty frames provide approximately two seconds of past context. [10]

Band	EQ filter	Measurement region
Low	100 Hz low shelf; slope $S = 1$	20–150 Hz
Low-mid	300 Hz peak; $Q = 1$	150–500 Hz
Mid	1 kHz peak; $Q = 1$	500 Hz–2 kHz
Presence	3.2 kHz peak; $Q = 1$	2–5 kHz
Treble	8 kHz high shelf; slope $S = 1$	5–16 kHz

Goal	Offsets from reference R	Gain sign constraints
More bass	Low +4 dB; low-mid +1.5 dB	Low ≥ 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 ≤ 0 dB

Use the RBJ biquad forms. Each gain is limited to ± 6 dB, then clamped to the goal constraints and scaled by strength from 0 to 1. The same gains drive both channels. Reference R is the bandwise median of the training segments after removing each segment’s mean band level, then recentered to zero mean. Add the goal offsets and recenter again to obtain target T. [10]

Model and training

The implementation target is a causal, depthwise-separable temporal convolutional network with approximately 25,000 parameters. Its inputs are the 20×24 feature history and a three-element goal code; its outputs are five bounded gains. Quantization-aware training exports int8 weights and activations. Use STM32Cube.AI or ST Edge AI Core, with TensorFlow Lite for Microcontrollers as a fallback. The target budgets are 150 KiB model flash and 64 KiB activation workspace; the timing requirement remains < 50 ms per inference.

Train from approximately 2,000 licensed FMA tracks using artist-disjoint training, validation, and test splits of approximately 70/15/15. Stratify by genre and record track IDs, license terms, preprocessing and hashes. Generate labels with a bounded offline optimizer over centered three-second windows, using the exact filter responses plus gain-size and temporal-smoothness penalties. Train with gain MSE and an auxiliary tonal loss; tune all weights and model dimensions only on validation. The label generator can use future audio, while the deployed model cannot. [11]

DSP protection and AI verification

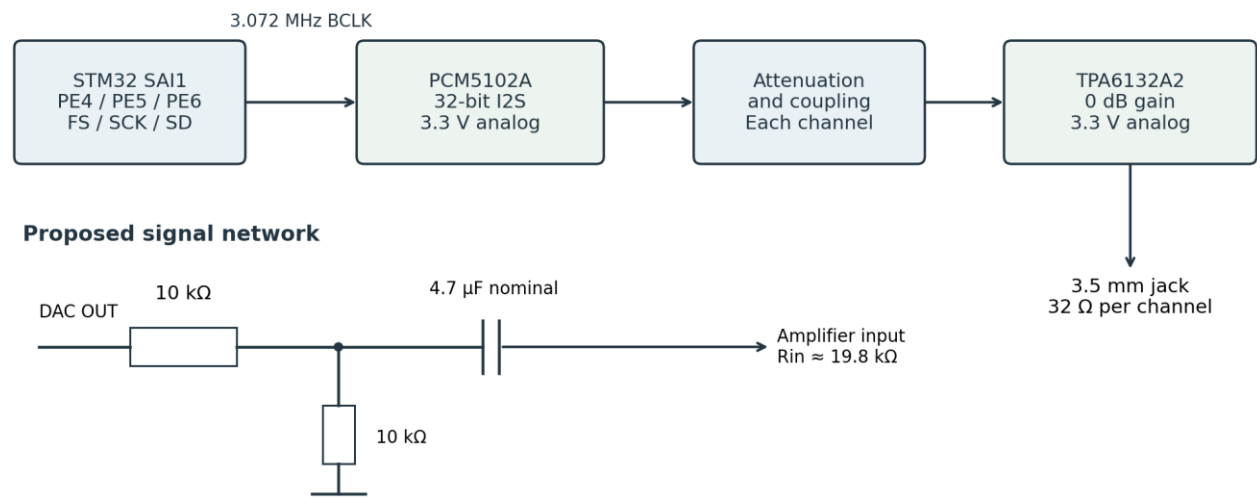
Smooth commanded gains with a 300 ms one-pole time constant and recompute coefficients every 48 samples (1 ms). The smoothing factor is $1 - \exp(-0.001/0.300) \approx 0.00333$; even a 12 dB target jump initially moves about 0.040 dB, below the 0.05 dB step bound. Recompute stable filter coefficients from bounded gains instead of blindly interpolating coefficient vectors. Use float DSP internally and convert once to signed 32-bit I2S samples.

Reserve headroom using the maximum combined filter response on a 64-frequency logarithmic grid, plus 0.5 dB. A 2 ms look-ahead limiter with immediate attack and a 50 ms release enforces a -1 dBFS sample-peak ceiling. The grid is a headroom estimate, not proof against every transient; the limiter supplies the final sample bound. This is not an analog true-peak or hearing-safety guarantee.

Below -60 dBFS input RMS, hold gains. A missed inference holds the previous gains and increments a counter. Non-finite or out-of-range pre-clamp outputs cause a 500 ms ramp to zero EQ gains. A stale-completion watchdog detects an inference transaction that has not produced a fresh sequence number for three hops; it also ramps to bypass. Repeated gain values alone are valid on steady content and do not trigger a fault. These fallbacks preserve audio, but a timing miss still fails the normal-operation timing test.

ID and role	Requirement	Verification and pass criterion
AI1 H3	10 Hz updates; inference < 50 ms.	Record sample index, completion sequence and elapsed cycles during the H1 run. Pass: all scheduled updates complete within 50 ms and are applied at the intended next-hop boundary.
AI2 Design target	Model ≤ 150 KiB flash and ≤ 64 KiB activation RAM.	Record generated model report and linker map. Pass target: both budgets met with total firmware and buffers fitting their actual SRAM banks; otherwise reduce the model before integration.
AI3 H3	All gains obey ± 6 dB and goal constraints.	Exercise every goal and strength on validation data and boundary inputs. Pass: every post-clamp command lies within its permitted interval; both channels receive identical values.
AI4 H1 H3	Gain step ≤ 0.05 dB/ms; digital peaks ≤ -1 dBFS.	Log gains and peaks over validation playback and 100 goal/bypass transitions. Pass: both numerical bounds hold with zero underruns; normal bypass crossfade ≥ 50 ms.
AI5 H1	Faults hold or ramp gains without stopping audio.	Inject NaN, infinity, range failure, stale sequence and one late completion. Pass: invalid/stale outputs reach zero EQ gain within 500 ms plus one audio block; late completion holds the prior gains.
AI6 Design target	Quantized gain error ≤ 0.25 dB RMS versus float model.	Run the same validation features through deployed and float models. Report RMS across all five gains and goals; meet target before frozen test evaluation.

Audio output design and verification



Complementary input AC-grounding, charge-pump capacitors and DAC output RC follow the device reference circuits.

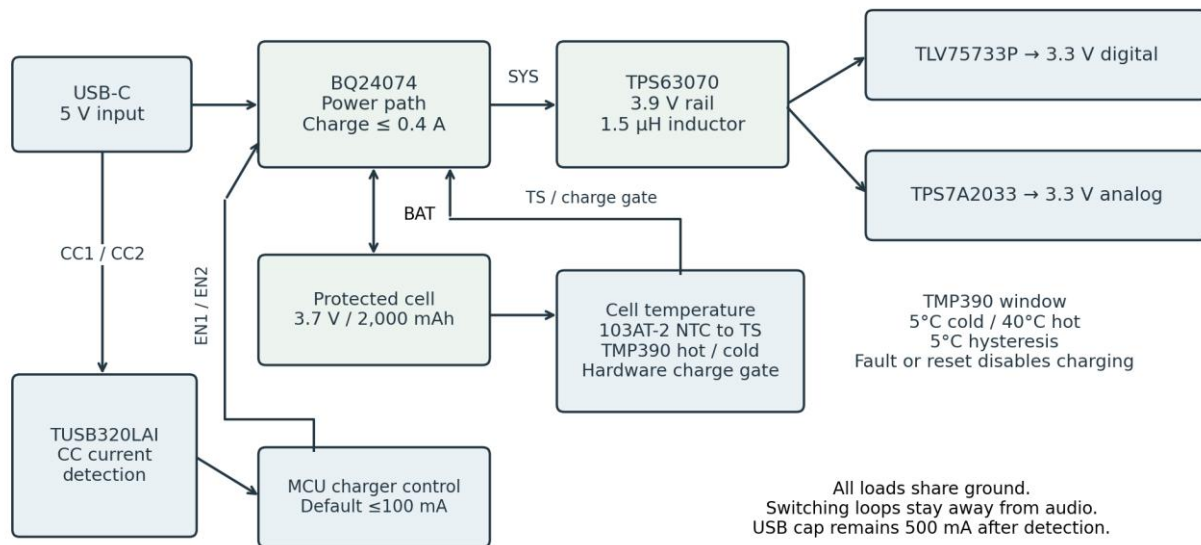
Figure 5. Audio connections and proposed per-channel attenuation. Support networks follow the TI application circuits. [4, 5]

The PCM5102A uses BCLK, LRCLK and DATA and generates its internal system clock from BCLK. Strap it for I2S; hold XSMT inactive until the clocks and buffers are valid. The TPA6132A2 is set to 0 dB gain. For each channel, 10 kΩ series and 10 kΩ shunt resistors feed a coupling capacitor and the amplifier input. Use 4.7 μF nominal coupling capacitors with at least 2.2 μF effective capacitance, and the manufacturer’s complementary-input grounding and charge-pump networks. [4, 5]

At the amplifier’s typical 19.8 kΩ input resistance, the divider factor is $(10\text{ k}\Omega \parallel 19.8\text{ k}\Omega) / [10\text{ k}\Omega + (10\text{ k}\Omega \parallel 19.8\text{ k}\Omega)] \approx 0.399$. The DAC’s 2.1 Vrms full-scale output therefore gives about 0.84 Vrms before limiting, or 0.75 Vrms at –1 dBFS. This supports the ≥ 0.50 Vrms target; loaded measurements must confirm clipping and component tolerance. Apply digital volume before the limiter. [4, 5]

ID and role	Requirement	Verification and pass criterion
AUD1 H2	Bypass response ± 1 dB, 20 Hz–20 kHz, each channel into 32 Ω.	Use the calibrated tone/sweep procedure in system verification. Pass: every measured response point, normalized at 1 kHz, lies within ± 1 dB including uncertainty.
AUD2 H2	THD+N $< 0.1\%$ at 1 kHz, 0.2 Vrms, 32 Ω.	Capture both loaded channels with 20 Hz–20 kHz analysis bandwidth, unweighted. Pass: each of three captures per channel $< 0.1\%$.
AUD3 H2 support	Maximum unclipped level ≥ 0.50 Vrms at 1 kHz into 32 Ω.	Increase tone level while inspecting waveform and THD+N. Pass: ≥ 0.50 Vrms with no visible clipping and THD+N $< 1\%$.
AUD4 Design target	Startup/shutdown/track-change transient < 10 mV peak with silent PCM.	Measure each loaded output on a DC-coupled scope for ten events of each type. Report peak; ordinary music amplitudes are excluded from this transient test.

Power design



Functional connection diagram. Temperature gate and CC logic supplement the original power path.

Figure 6. Proposed charging and power architecture. The intermediate rail is 3.9 V, consistent with the tolerance basis.

The BQ24074 power path shares input current between the system and the cell. The TPS63070 maintains a 3.9 V intermediate rail from battery or charger output. TLV75733P supplies the digital 3.3 V rail and TPS7A2033 supplies the analog 3.3 V rail. This separates supply noise while retaining one ground reference. Their required continuous capacities are 500 mA and 100 mA, respectively. [6–9]

Adafruit 2011 is the candidate protected 2,000 mAh cell; the associated PKCELL LP-803860 datasheet states a 1,900 mAh minimum capacity, 4.2 V charge voltage and 0–45°C charging range. The two-wire pack has no thermistor. Add a thermistor against the cell for BQ24074 TS monitoring. [12]

A bare 103AT-2 on BQ24074 nominally permits about 0–50°C, wider than the candidate cell's charging range. Add a TMP390A2 thermal window against the cell with 5°C cold and 40°C hot trips and 5°C hysteresis. SETA = 2.15 kΩ and SETB = 18.7 kΩ. Its active-low fault outputs remove charge permission through a 2N7002 gate; reset or absent control power also disables charging. Verify thermal contact and actual trip points before cell use. [6, 16]

USB-C CC detection uses a TUSB320LAI configured as a sink. Its integrated Rd termination replaces separate parallel CC pull-downs. Default to the charger's ≤ 100 mA input mode. Permit ≤ 500 mA only after detecting a Type-C source advertising at least 1.5 A; remain at the lower setting for unrecognized/default sources. The design never raises the project's 500 mA cap. At low input power, the battery can supplement playback and charging may slow or stop. [6, 15]

Power parameters and verification

Parameter	Proposed value and rationale
Intermediate regulator	387 k Ω / 100 k Ω feedback, 0.1%: $0.8 \times (1 + 3.87) = 3.896$ V. Use the TPS63070 1.5 μ H reference LC network and verify effective capacitance. [7]
Charge current	RISET = 2.49 k Ω , 1%: approximately 0.357 A typical and <0.396 A using the datasheet maximum factor and resistor minimum. [6]
Input limit	Use the fixed USB100/USB500 modes. EN2 remains low; EN1 is raised only after valid CC detection. Charger controls default low. [6]
Battery measurement	100 k Ω / 100 k Ω divider, 0.1%, with 100 nF at ADC; 4.2 V maps to 2.1 V. Calibrate against DMM and use adequate ADC acquisition time.
Low battery	Warning at 3.5 V; controlled shutdown at 3.3 V after sustained low readings for 1 s; calibrate thresholds within ± 0.05 V. Independent cell protection remains active.
Temperature control	TMP390A2 trips nominally below 5°C or above 40°C, with 5°C hysteresis. Charge remains disabled until initialization and sensor checks complete.
Analog rail accuracy	TPS7A20 $\pm 1.5\%$ worst-case accuracy equals ± 49.5 mV at 3.3 V. This nearly consumes the ± 50 mV rail allowance; verify DC accuracy and transients, or select a tighter regulator if the rail fails. [9]

ID and role	Requirement	Verification and pass criterion
PWR1 H1 H2	Digital 3.20–3.40 V at 500 mA; analog 3.25–3.35 V at 100 mA.	With a battery simulator at 3.3, 3.7 and 4.2 V, test 10%, 50% and 100% rail loads, then simultaneous maximum loads. Repeat with USB attached. DMM measures DC; scope checks load transitions. Pass: rail limits maintained, including uncertainty.
PWR2 H1	USB input never exceeds 500 mA; source authorization honored.	Test default, 1.5 A and 3 A Type-C advertisements with an inline meter. Pass: default ≤ 100 mA and recognized-source total input ≤ 500 mA during startup, charge and playback.
PWR3 Safety	Warning, shutdown and thermal inhibition act before cell limits.	Sweep simulated battery voltage; inject NTC open/short and TMP390 fault conditions without stressing a cell. Pass: thresholds occur in order, files close, audio mutes and charger disables. Validate thermal trips against a reference sensor.
PWR4 H1	Battery runtime ≥ 3 h.	Run the fixed full-charge playback protocol. Pass: elapsed playback to controlled shutdown ≥ 3 h, with zero resets or underruns.
PWR5 Safety target	Components within ratings; exterior $< 45^\circ\text{C}$ at 25°C ambient.	Log cell, regulator, charger and exterior temperatures under playback and playback-plus-charge until stable. Pass required ratings; report the exterior-temperature target separately.

Integrated verification of playback and audio

H1 playback and battery runtime

1. Freeze a ten-track, genre-varied playlist and file hashes for repeatable playback. Fully charge the reviewed cell to normal charger termination, let it rest 30 minutes, and disconnect USB. Use two $32\ \Omega$, 1%, $\geq 0.25\ \text{W}$ non-inductive loads. Record measured ambient temperature near 25°C .
2. Before the run, set bypass and play a $-12\ \text{dBFS}$ 1 kHz tone. Adjust volume to 0.2 Vrms per channel, then retain that setting. Select more bass at 100% strength, keep the OLED continuously active at its documented midrange contrast setting, and loop the playlist. Record firmware hash, battery identification and initial voltage.
3. For the first 60 minutes, log underruns, resets, inference timing, minimum buffer fill and battery voltage. Exercise each control regularly without changing the calibrated volume. Continue the same playlist to controlled shutdown; a low-battery warning does not end the timer.
4. H1 passes only when the 60-minute run has zero underruns and total playback is at least 3 hours. Report resets and inference misses as additional subsystem failures. Repeat the timing test on the other two goals; extend the battery test to any goal found to consume more power than the original condition.

H2 headphone output measurement

5. Use a lab audio analyzer or a stereo USB audio interface capable of 48 kHz, 24-bit capture, with REW or a versioned Python analysis script. Verify line-input range and connection polarity. Keep the $32\ \Omega$ load across the output while the high-impedance interface measures it. Do not connect a grounded probe across any node that is not ground-referenced.
6. Measure loopback response, noise and distortion at the actual capture levels. The setup must resolve the $\pm 1\ \text{dB}$ response requirement and have residual THD+N $\leq 0.03\%$ for the 0.1% test. If it does not, use a better instrument; do not subtract THD+N scalars to manufacture a passing result.
7. With adaptive EQ bypassed, measure a sweep plus explicit 20 Hz, 1 kHz and 20 kHz points at a fixed volume. Normalize to 1 kHz and retain any measured acquisition-path correction identically for both channels. Capture three 1 kHz trials at 0.2 Vrms per channel into $32\ \Omega$, using 20 Hz–20 kHz, unweighted THD+N.
8. H2 passes only if both channels meet both response and distortion limits with measurement uncertainty accounted for. Report maximum unclipped output and the silent-PCM transient checks separately. Record A-weighted dynamic range as characterization, with 85 dB as a design goal rather than a new high-level pass condition.

Adaptive EQ evaluation

Frozen evaluation definition

Before final fitting, commit the artist-disjoint split lists, preprocessing, metric implementation, goal targets, gain limits, startup exclusion, silence mask, baseline fitting procedure, and subset-selection seed. Select 100 of the approximately 300 test tracks by deterministic genre-stratified sampling; select a fixed ten-track subset from those 100 for live equivalence and analog confirmation. No test-set result selects model settings. [11]

Discard the first 3 seconds of every track. Score complete, non-overlapping three-second segments thereafter; omit a final incomplete segment. A segment is silent when its original stereo input RMS is below -60 dBFS, relative to full-scale amplitude 1. Use that same input-derived mask for every processing condition. Confirm every selected track has at least one eligible segment before the split is frozen.

For each output segment, compute a Hann-windowed spectrum for each stereo channel, average their powers, and sum bins within the five half-open measurement bands; include 16 kHz in the last band. Normalize by window power, apply a fixed 10^{-12} power floor, and convert to band dB values L . Remove their mean and compare with the zero-mean target T . This definition and its exact FFT/bin handling are identical for all conditions.

$$e = \sqrt{\frac{1}{5} \sum_{b=1}^5 [L_b - \text{mean}(L) - T_b]^2}$$

Average segment errors within each track, then average the 100 track errors with equal track weight. For each goal, reduction is $100 \times (\text{mean fixed error} - \text{mean adaptive error}) / \text{mean fixed error}$. The fixed preset is the best solution found by a training-only multistart search; the search is not a proof of a global optimum. Re-score candidates through the same smoothing, pre-gain and limiter path used for the neural controller.

Pass conditions and reproducibility

Process all 100 tracks on the device for each goal and both controllers at full EQ strength. H3 passes per goal only if mean error falls by $\geq 20\%$, at least 65 of 100 tracks improve strictly, and the one-sided Wilcoxon test gives $p < 0.01$. Use paired differences fixed minus adaptive, `zero_method = wilcox` and `alternative = greater`; pin the SciPy version and p-value method before testing. Report ties, usable pair count and exact procedure. Report all three goals separately; all must pass.

Run the full approximately 300-track host test once the model is frozen. Compare with the same optimized fixed preset, a causal bounded rule controller, and the non-causal label optimizer as a reference. The optimizer is not a formal upper bound for the final metric because its objective and runtime assumptions differ. Report every comparison, including cases where the simpler controller performs better.

For the fixed ten tracks and all goals, require matching per-block CRC-32 sequences between live playback and offline device output. Also byte-compare captured PCM where practical. For analog confirmation, capture the same tracks into 32 Ω at 48 kHz/24-bit, align by correlation and account for clock drift. Measure the acquisition/bypass response once, apply identical correction, and use the same input silence mask. Analog percentage error reduction must agree with digital reduction within 5 percentage points per goal.

Tolerance analysis and feasibility

Battery energy and regulator headroom

The critical hardware risk is insufficient energy or supply headroom. Use the submitted 350 mA average equivalent load at 3.3 V, 80% usable battery energy and 75% end-to-end conversion efficiency. These are design assumptions to be replaced with first-board measurements; the 75% factor includes the LDO losses and must not be counted a second time.

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

$$t_{run} = (3.7 \text{ V} \times 2.0 \text{ Ah} \times 0.80 \times 0.75) / 1.155 \text{ W} = 3.84 \text{ h}$$

Scenario	Estimated runtime	Implication
2,000 mAh; 350 mA load	3.84 h	0.84 h estimated margin
2,000 mAh; 420 mA load	3.20 h	20% higher load still above 3 h
1,900 mAh; 420 mA load	3.04 h	Minimum stated capacity leaves little margin
1,900 mAh; 420 mA; efficiency 70%	2.84 h	Fails; reduce measured load or increase capacity

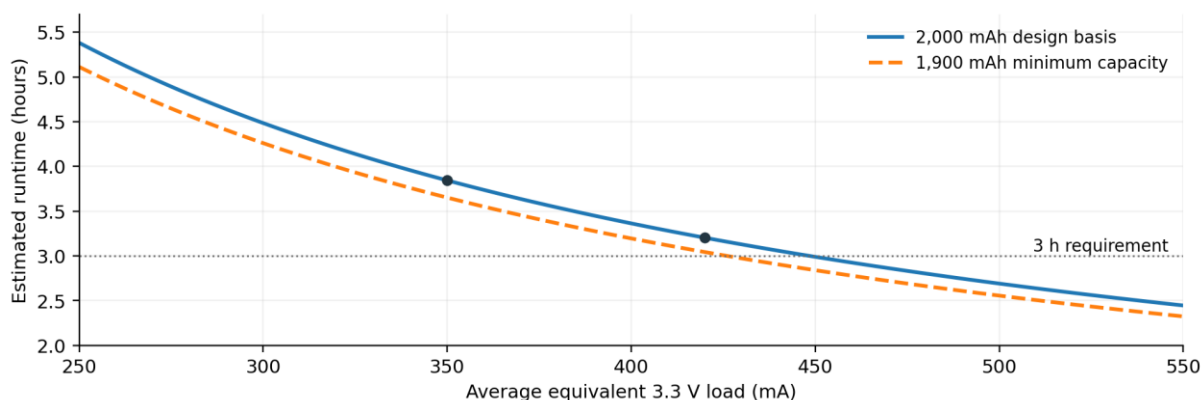


Figure 7. Calculated sensitivity, assuming 80% usable energy and 75% conversion efficiency. These curves are not measured runtime.

The 2,000 mAh assumption allows about 448 mA average equivalent 3.3 V load for 3 hours; the 1,900 mAh minimum allows about 426 mA. The chosen battery therefore needs a measured power check early in bring-up. A 20% load margin is not sufficient if usable energy or conversion efficiency also declines substantially.

The 387 k Ω / 100 k Ω divider gives 3.896 V nominal. Under a 2% low rail assumption, the LDO input is 3.818 V. A 1% high 3.3 V output is 3.333 V, leaving 0.485 V headroom. Against TLV757P's 0.425 V maximum dropout at 1 A through 85°C, this leaves about 60 mV. At 500 mA, LDO loss is about $(3.896 - 3.3) \times 0.5 = 0.298 \text{ W}$. Verify ripple, transient minima and temperature because this calculation alone does not guarantee board performance. [7, 8]

Timing margins and integration risks

Memory and execution budgets

Allocation	Planning budget	Reason
Input read-ahead	128 KiB	0.683 s of 16-bit stereo PCM
Output DMA queue	16 KiB	42.7 ms of two 32-bit I2S slots
Model activation workspace	64 KiB	Deployment target
FFT, stereo staging, features and limiter	96 KiB	Two-channel analysis with reusable workspace
Stacks, files, UI and diagnostics	128 KiB	Conservative software allocation
Total planned SRAM	432 KiB	Check bank accessibility as well as total capacity

A 25,000-weight int8 model stores roughly 25 kB of raw weights; the 150 KiB flash target also allows graph metadata and quantization data. The rest of firmware must fit the MCU's total flash. SRAM capacity does not by itself prove correct placement: DMA-visible rings and inference workspace may have different bank requirements. [3]

The <5 ms per 10 ms audio budget leaves at least 50% nominal CPU time for other work. A <50 ms inference every 100 ms can consume almost that entire remainder, so the two limits alone do not prove schedulability. Target ≤ 3 ms audio blocks and ≤ 30 ms inference during development, including feature work, leaving roughly 40% for storage, UI and interrupts. Only the integrated worst-case measurements establish the required margins.

Integration risks and responses

Risk	Observable trigger	Response and owner
Model lacks H3 margin	Validation improvement near/below 20%	Andrew revisits data, labels and model on validation only. Preserve the frozen test protocol and report failure honestly if H3 is not met.
CPU or memory pressure	Timing maximum approaches limit; bank overflow	Advaith prioritizes audio, reduces UI/logging load and reuses workspace; Andrew reduces the model if necessary.
Audio noise or clipping	H2 failure, rail ripple or coupling	Siddharth measures DAC and amplifier separately, checks power returns and attenuation, then revises the affected circuit.
Runtime below target	Measured energy budget predicts <3 h	Siddharth and Advaith reduce average load or qualify a larger cell, then repeat runtime and enclosure checks.
Late PCB or assembly fault	Bring-up misses planned week	Continue independent firmware/AI work on development hardware; reserve revision time and parts. The final device still requires the custom board.

Spliced-passage testing is supplementary characterization. Use clips with a ≥ 4 dB change in a goal's primary band, report gain trajectories and settling time, and confirm uninterrupted playback. Do not force a fixed gain change where the bounded optimizer's desired response is itself unchanged.

3 Cost and schedule

Candidate bill of materials

The following candidates make the design and purchasing estimate concrete. They are proposed selections, not claims of parts already purchased or footprints already released. Unit amounts are budget estimates in USD except the Adafruit battery listing at \$12.50, checked October 2 2026. Click the part or sourcing label for its supplier route. Final procurement uses the reviewed PCB BOM and current quotes.

Component and candidate	Key specification	Qty	Unit \$	Total \$
MCU STM32H743ZIT6	ST; Cortex-M7, LQFP144	1	22.00	22.00
Stereo DAC PCM5102APWR	TI; I2S DAC, TSSOP20	1	5.00	5.00
Headphone amplifier TPA6132A2RTER	TI; stereo, 0 dB configuration	1	3.00	3.00
Battery charger BQ24074RGTR	TI; 1-cell charger/power path	1	3.00	3.00
Buck-boost TPS63070RNMR	TI; adjustable, 3.9 V target	1	4.00	4.00
Digital LDO TLV75733PDRVR	TI; 3.3 V, 1 A, thermal pad	1	1.00	1.00
Analog LDO TPS7A2033PDBVR	TI; 3.3 V, 300 mA, low noise	1	1.00	1.00
USB-C CC controller TUSB320LAIRWBR	TI; sink/current detection	1	2.50	2.50
Cell thermal switch TMP390A2DRLR	TI; programmable hot/cold trips	1	2.00	2.00
Protected battery Adafruit 2011	PKCELL LP-803860 candidate; 3.7 V, 2,000 mAh	1	12.50	12.50
OLED module 0.96inch OLED (B)	Waveshare; SSD1306, 128 × 64, 4-wire SPI variant	1	15.00	15.00
Rotary encoder PEC11R-4215F-S0024	Bourns; 24 detents, push switch	1	4.50	4.50
Buttons B3F-1000	Aratas / Omron; momentary tactile	4	0.50	2.00
microSD socket 1040310811	Molex; push-pull, detect switch	1	2.00	2.00
USB-C receptacle USB4105-GF-A	GCT; USB-C receptacle, footprint review	1	1.50	1.50
Headphone jack SJ1-3523N	Same Sky; stereo 3.5 mm, through-hole	1	1.50	1.50
Cell NTC 103AT-2	Semitec; 10 kΩ at 25°C	1	1.50	1.50
HSE crystal ABM8-25.000MHZ-B2-T	Abracon; 25 MHz; calculate load caps	1	1.00	1.00

Major component subtotal: \$85.00. Manufacturer data and orderable package variants are referenced in [3–9, 12–18].

Supporting parts and total cost

Support quantities are estimates with an assembly allowance. Named regulator LC parts follow its application BOM; finalize remaining values, MPNs and reference-designator quantities during schematic capture. [7]

Item and sourcing route	Specification or status	Qty	Unit \$	Total \$
Inductor Coilcraft XFL4020-152ME	1.5 μ H; verify peak current and land pattern	1	2.00	2.00
Regulator input capacitors Murata GRM21BC71E106ME11L	10 μ F, 25 V, X7S, 0805	2	0.80	1.60
Regulator output capacitors Murata GRM21BC81C226ME44L	22 μ F, 16 V, X6S, 0805	3	0.80	2.40
Local decoupling and coupling MPNs at schematic freeze	0.1 / 1 / 2.2 / 4.7 / 10 μ F; effective C verified	40	0.20	8.00
Resistors and pull-ups MPNs at schematic freeze	1% general; 0.1% feedback and ADC ratios	40	0.10	4.00
Connectors and small protection Final footprints at schematic freeze	Battery header, SWD/UART, power switch, 2N7002, TVS/ESD, charge gate and wiring	1	15.00	15.00
microSD card 16–32 GB FAT32 card allowance	Selected card qualified by storage tests	1	10.00	10.00
Custom PCB and stencil Four-layer prototype fabrication	One order allowance; final quote uses Gerbers	1	50.00	50.00
Enclosure and fasteners Printed or machined project enclosure	Prototype material / hardware allowance	1	20.00	20.00
Test loads and cables 32 Ω non-inductive loads and adapters	Two ≥ 0.25 W 1% loads; audio and power adapters	1	15.00	15.00

Labor and project estimate

Labor uses the course planning rate of \$50/hour with a 2.5 overhead multiplier: 10 hours per member per week for 12 active weeks. These are estimated engineering costs, not student wages or hours already worked. [1, 2]

Member	Hours	Labor calculation	Estimated cost
Siddharth Salapaka	120	$\$50 \times 2.5 \times 120$	\$15,000
Andrew Cheng	120	$\$50 \times 2.5 \times 120$	\$15,000
Advaith Anand	120	$\$50 \times 2.5 \times 120$	\$15,000

Parts and fabrication: \$213.00. Spare/revision reserve: \$60.00. Shipping/tax allowance: \$25.00. Optional audio-interface access/purchase reserve: \$150.00. Total cash planning budget: \$448.00. Labor: \$45,000. Total estimated project cost: \$45,448.00.

Lab DMM, oscilloscope, bench supply, logic analyzer and rework tools are budgeted at \$0 incremental cost, subject to availability. No paid machine-shop hours are assumed. Replace allowances with quotes or receipts during purchasing.

Weekly work allocation

Primary ownership follows the signed team contract: Siddharth owns hardware and bring-up; Andrew owns training, quantization, deployment and evaluation; Advait owns storage, playback, DSP and UI. Integration is shared. Andrew maintains the decision/action list, and the team reviews it in the weekly meeting. The dates below are work targets aligned with the supplied semester plan, not claims that earlier work has been completed.

Week	Siddharth	Andrew	Advait
Sep 21–25	Define schematic blocks and verify major IC choices.	Create artist-disjoint split and target/metric specification.	Bring up SD-to-DAC WAV playback on development hardware.
Sep 28–Oct 2	Prepare PCB interfaces, power calculations and candidate BOM.	Fit fixed baselines and generate first training labels.	Specify buffers, scheduling, UI flow and timing logs.
Oct 5–9	Close design-review circuit issues; freeze footprints and order PCB.	Train first TCN; export and profile quantized candidate.	Integrate model interface and deterministic gain scheduling.
Oct 12–16	Assemble available board; check rails, clocks, SWD and storage.	Measure validation margin and quantization error.	Demonstrate UI, DSP, playback and fault recovery on development board.
Oct 19–23	Bring up DAC/amplifier and temperature/current protections.	Integrate device inference and activation-memory report.	Port audio/storage/UI to custom board; verify cache ownership.
Oct 26–30	Resolve hardware faults; place revision order if required.	Validate device/host gains and finalize evaluation scripts.	Complete offline output mode and live CRC logging.
Nov 2–6	Support planned progress demo; measure power and loaded audio.	Run validation-only end-to-end evaluation and freeze candidate.	Measure worst-case timing, UI latency and underruns on board.
Nov 9–13	Run battery, thermal and enclosure checks.	Freeze final model; run held-out H3 and analog analysis.	Run H1, equivalence and transition tests; freeze core firmware.
Nov 16–20	Repair hardware issues from planned mock demo.	Document results and repeat only affected tests with disclosure.	Fix integration defects and rehearse full demo.
Nov 23–27	Fall break; review documentation remotely if needed.	Fall break; consolidate result tables if needed.	Fall break; organize logs and report figures if needed.
Nov 30–Dec 4	Support final demonstration and hardware evidence.	Present model, comparisons, statistics and limitations.	Run demo and present timing/playback evidence.
Dec 7–9	Finalize measured BOM and hardware report; checkout.	Finalize evaluation report and reproducibility archive.	Finalize firmware report and archive tagged release.

The critical path is schematic freeze → fabrication → power/audio bring-up → integration → verification. Dataset, model and firmware development proceed in parallel before board arrival. The October revision window and spare budget absorb a first-board fault. Final test results are not reused as validation data; any post-test repair is identified and its retesting disclosed.

4 Ethics safety and societal impact

Responsible evaluation and use

ACM principles on avoiding harm, honest claims, respecting intellectual work, privacy, and thorough evaluation apply directly. We will preserve artist-disjoint test data, report all required comparisons and failures, and identify estimated values separately from measurements. The tonal metric expresses agreement with a chosen curve; it does not measure musical enjoyment, improve hearing, or isolate a singer. Dataset licenses and attribution remain attached to the track manifest, and source music is not redistributed outside its terms. [11, 19]

Offline processing avoids subscription cost and dependence on network access. Normal operation neither uploads audio nor collects listening history. Physical controls support independent use, but a small display can be difficult for users with low vision. Readable text, simple menus and tactile controls reduce that barrier. The three goals represent selectable preferences, not cultural or genre-independent definitions of good sound. Strength and bypass keep the listener in control.

Electrical battery and mechanical safety

Battery power is necessary for the portable three-hour use case. The team will complete the course laboratory and fire-safety training, obtain review by the TA and a power-focused TA before testing the charging/protection circuit, use the required lithium safety bag for charge/discharge testing, and work with at least two people present. [20]

Begin with a current-limited supply and confirm connector polarity before attaching the cell. Keep battery terminals insulated, use a keyed connector, secure wiring against strain, and separate the cell from solder joints, sharp edges, regulators and fasteners. A damaged or swollen cell is removed from service. NTC and hardware thermal-gate faults are tested with substitute resistances or test inputs. A battery simulator supplies short-circuit/fault tests; never deliberately short the lithium cell. [12, 20]

Use dummy loads for electrical tests. Boot at low volume and keep listening checks brief. The limiter caps digital samples, but headphone sensitivity and fit determine acoustic exposure; the interface warning cannot establish safe listening for arbitrary headphones. WHO safe-listening guidance motivates conservative startup volume and user awareness. [21]

Standards and environmental impact

Guidance	Application to this project
IEC 62368-1	Review accessible energy sources, insulation, heating, mechanical protection and enclosure safeguards for audio/ICT equipment. [22]
IEC 62133-2	Seek supplier safety evidence for the rechargeable cell/pack. A UN38.3 shipping report is not an IEC 62133-2 certificate. [23]
USB Type-C and I2S interfaces	Honor CC current advertisement and charger limits; validate serial-audio format, framing and timing. [4, 15]
FCC Part 15	A future commercial version needs applicable emissions assessment; short clocks, ground returns and decoupling reduce coupling. This prototype claims no certification. [24]
Waste and serviceability	Use removable storage, a serviceable cell connection and screw fasteners. Route cells and electronics to suitable recycling programs rather than mixed waste.

5 References

- [1] University of Illinois ECE 445, "Design Document." [Source](#)
- [2] University of Illinois ECE 445, "Design Review," including labor-cost guidance and evaluation rubric. [Source](#)
- [3] STMicroelectronics, STM32H742xI/G and STM32H743xI/G datasheet, DS12110. [Source](#)
- [4] Texas Instruments, PCM510xA stereo audio DAC datasheet. [Source](#)
- [5] Texas Instruments, TPA6132A2 stereo headphone amplifier datasheet. [Source](#)
- [6] Texas Instruments, BQ2407x battery charger and PowerPath datasheet. [Source](#)
- [7] Texas Instruments, TPS63070 buck-boost converter datasheet. [Source](#)
- [8] Texas Instruments, TLV757P low-dropout regulator datasheet, rev. C. [Source](#)
- [9] Texas Instruments, TPS7A20 low-noise regulator datasheet. [Source](#)
- [10] R. Bristow-Johnson and R. Toy, "Audio EQ Cookbook," W3C Working Group Note, 2021. [Source](#)
- [11] M. Defferrard et al., "FMA: A Dataset for Music Analysis," ISMIR, 2017, and dataset repository. [Source](#)
- [12] Adafruit, protected lithium-ion battery product 2011, with linked PKCELL LP-803860 datasheet. [Source](#)
- [13] WaveShare, "0.96inch OLED (B)" and SSD1306 SPI interface documentation. [Source](#)
- [14] Bourns, PEC11R incremental encoder datasheet. [Source](#)
- [15] Texas Instruments, TUSB320LAI Type-C configuration-channel controller datasheet. [Source](#)
- [16] Texas Instruments, TMP390 dual-channel temperature switch datasheet. [Source](#)
- [17] Abracon, ABM8 crystal datasheet. [Source](#)
- [18] Molex, microSD connector 1040310811; GCT USB4105 and Same Sky SJ1-3523N product data are linked in the BOM. [Source](#)
- [19] Association for Computing Machinery, "ACM Code of Ethics and Professional Conduct," 2018. [Source](#)
- [20] University of Illinois ECE 445, "Safety Guidelines" and linked "Safe Practice for Lead Acid and Lithium Batteries." [Source](#)
- [21] World Health Organization, "New WHO-ITU standard aims to prevent hearing loss," 2019. [Source](#)
- [22] IEC, IEC 62368-1:2023, audio/video and ICT equipment safety requirements. [Source](#)
- [23] IEC, IEC 62133-2:2017 with Amendment 1:2021, portable lithium-cell and battery safety. [Source](#)
- [24] U.S. Electronic Code of Federal Regulations, 47 CFR 15.109, radiated emission limits. [Source](#)