

# Towards a Low Cost, Microcontroller-Based Class-D Audio Amplifier

Erik Genthe

Deutsches Elektronen-Synchrotron DESY  
Hamburg, Germany  
e-mail: mail@erikgenthe.de

Timm Bostelmann and Sergei Sawitzki

FH Wedel (University of Applied Sciences)  
Wedel, Germany  
e-mail: {bos|saw}@fh-wedel.de

**Abstract**—In this implementation study, a simple, low cost digital (class-D) audio amplifier is presented. Class-D audio amplifiers are much more efficient than prior (analogue) types of amplifiers. However, such amplifiers still have significant power losses due to quiescent currents independent of their output power. The design goal was to reduce the complexity of the hardware design as much as possible while keeping high audio quality and power efficiency. This is done by covering as much complexity as possible with a standard microcontroller using its internal hardware Pulse-Width Modulation (PWM) for signal generation. A key part of this work lies in the measurement and evaluation of the implemented amplifier prototype. It is shown that the Total Harmonic Distortion Plus Noise (THD+N) is less than 1 % for frequencies up to 500 Hz. Further analysis shows that quantization is the main source of this noise. At 10 kHz, the measured THD+N is 7.0 %, which is audible. However, this high noise level is not caused by quantization. The prototype has a quiescent current of 159 mA. The prototype demonstrates the potential of this simple approach in terms of energy efficiency. In addition, significant improvements could be achieved by incorporating noise shaping algorithms and multi-level switching techniques.

**Keywords**—Measurement; Class-D; Microcontroller.

## I. INTRODUCTION

There are several classes of power amplifiers. Only class-D amplifiers come into consideration for battery powered applications because of their outstanding efficiency, which is often over 90 % at high output levels. Analogue audio amplifiers usually have quite high quiescent currents, which limit the efficiency at low output levels. In contrast, the TAS5630B integrated class-D amplifier Integrated Circuit (IC) has a very low quiescent current [1]. A high quality board embracing a TAS5630B with very good specifications is offered by WONDOM [2]. The detailed specifications are:

- Power supply voltage 25 V to 48 V
- Idle Power: 4.8 W
- Efficiency at high power: 91 % to 96 %
- Switching frequency: 400 kHz
- THD+N = 1 % @ 246 W, 4  $\Omega$

The downsides are:

- Idle power of 4.8W is still dominating the power consumption for “normal hearing conditions”.
- The TAS5630B is “mostly” monolithic [1]. The switching-transistors are internal, so they cannot be replaced to improve the characteristics of the device or for repairing purposes.
- Extra cost for the class-D IC.

The idea of this paper is to create and evaluate a simple and low cost class-D amplifier. Complexity and cost are to be held

low by using a microcontroller that receives the audio signal, handles all the amplifier logic and controls a simple circuit of Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) drivers, which includes a low pass filter. The difficulty is to do this fast and precise enough. Since the microcontroller is usually required anyway (e.g., for wireless signal transmission), this way the system complexity and cost can be reduced in comparison to using a separate class-D controller IC. In the end, two circuit boards shall be used to implement a full-bridge amplifier.

The implementation goals are:

- No noticeable distortions. (THD < 1 % for up to 80 % output power)
- Quiescent current resulting in an idle power significantly lower than 4.8 W
- Efficiency at high output power > 90 %
- Low temperature of components to achieve a long lifespan
- Output voltages of up to 48 V
- No Analog to Digital Converter (ADC)
- Frequency range of 20 Hz to 20000 Hz

The remainder of the paper is organized as follows: In Section II, the key challenges of this implementation study are identified and analyzed. In Section III, the implementation and the ways to overcome these challenges are discussed. Since measurement and validation are key parts of this work, the measurement setup is also described here. In Section IV, the results are presented and analyzed in detail. Finally, in Section V, a summary of this work is provided and future improvements are outlined.

## II. CHALLENGES

This section focuses on the challenges which have to be addressed in order to achieve the defined performance goals, which are the switching frequency, steepness of switching slopes, and keeping electromagnetic radiation low.

### A. High Switching Frequency

The switching frequency is directly depending on the Pulse Width Modulation (PWM) frequency of the microcontroller. It is possible to adjust the PWM frequency to sufficiently high values with modern microcontrollers. However, this comes with a trade-off between the PWM frequency and the PWM resolution based on the PWM base clock frequency, as will be discussed further below.

Table I summarizes some key characteristics of several microcontrollers. Even though there are more powerful options,

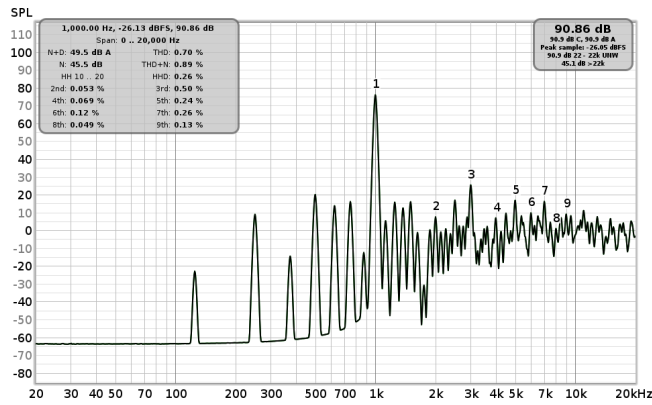


Figure 1. Exemplary Fast Fourier Transformation (FFT) of a synthetic 1 kHz signal with a sampling rate of 78125 Hz and a bit depth of 10 bit at a volume of -20dB, using the software REW [8].

the most convenient choice would be a microcontroller with Bluetooth on-chip, like ESP32 or STM32WB55. Compared to the STM32WB55, the ESP32 has a higher PWM base clock frequency of 80 MHz. Following this, it had to be verified, if the ESP32s PMW base clock frequency is sufficient to realize the specified objective of less than 1 % THD+N. This would make the ESP32 the preferred option.

The PWM base clock frequency  $f_{clk}$  is given by the microcontroller. In case of an ESP32, it is 80 MHz. The PWM configuration requires one parameter, the PWM resolution  $r_{pwm}$ . Given this parameter and the PWM base clock frequency  $f_{clk}$ , the PWM frequency  $f_{pwm}$  can be calculated as shown in (1). The ESP32 supports only PWM resolutions that are powers of two. As the frequency response should range up to 20 kHz, a PWM frequency of at least 40 kHz is needed.

$$f_{pwm} = \frac{f_{clk}}{r_{pwm}} = \frac{80 \text{ MHz}}{1024} = 78125 \text{ Hz} \quad (1)$$

Using the ESP32, the ideal resolution is therefore 1024 and the ideal PWM frequency is 78125 Hz.

To find out if it is possible to reach the 1 % goal, using this configuration, a digital 1 kHz audio signal with a depth of  $\log_2(1024) = 10$  bit and a sampling rate of 78125 Hz was generated. Using the software Room Equalization Wizard (REW) [8], the harmonics and noise at different volume levels were calculated. The resulting THD+N values are 0.074 % @0dB, 0.89 % @-20dB (Figure 1) and 9.8 % @-40dB.

The analysis of the synthetic signal shows that the best achievable audio quality, heavily depends on the volume level. This issue can be addressed by altering the power line voltage of the output stage rather than decreasing the volume digitally. Using this technique, the PWM base clock frequency of an ESP32 is sufficiently high to achieve a THD+N of less than 1 %. Given the fact that a full bridge amplifier is implemented, it should be possible to double the resolution (resulting in 11 bit), by switching both half bridges independently, which should

reduce the THD+N even further. For more information refer to Section III-D. Consequently, it is unnecessary to discuss the options without Bluetooth capabilities to make use of a higher PWM base clock frequency, which makes the ESP32 the preferred choice.

### B. Steep Switching Slopes

To achieve high efficiencies, the switching slopes have to be as steep as possible. It is essential to choose a MOSFET with a low gate charge in order to achieve sharp switching slopes, as well as a low on-resistance for overall efficiency. The STP100N6F7 transistor was selected due to its on-resistance of  $R_{DS(on)} = 5.6 \text{ m}\Omega$  and its gate charge of  $Q_G = 30 \text{ nC}$  [9]. The approximate losses, caused by switching times were calculated, based on values provided by the datasheet (2). For the calculation the switching frequency is  $f_{PWM} = 78125 \text{ Hz}$ , the supply voltage is  $V_{GS} = 10 \text{ V}$ , the MOSFET gate charge is  $Q_G = 30 \text{ nC}$ , four MOSFETs are used, and there are two switching slopes per period [9].

$$P_{loss} = f_{PWM} \cdot Q_G \cdot V_{GS} \cdot N_{FETs} \cdot 2 \\ = 78125 \text{ Hz} \cdot 30 \text{ nAs} \cdot 10 \text{ V} \cdot 4 \cdot 2 = 0.187 \text{ W} \quad (2)$$

Steep switching times with low resistance MOSFETs require high gate currents. If the gate driving unit cannot deliver sufficient current, the switching time will be longer and the efficiency lower. The required gate current to achieve the given switching times is derived from (3). The calculated maximum gate current of 2.1 Ampere (3) is too high for a microcontroller output. Besides that, most microcontrollers have 3.3 V or 5 V voltage levels, whereas most MOSFETs need higher voltages (i.e., 10 V) to achieve very low resistance. Therefore, a MOSFET driver is needed that can output more than 2 A at 10 V. With such a driver, it should be possible to keep switching losses low.

$$I_{G\_max} = \frac{10 \text{ V}}{4.7 \Omega} = 2.1 \text{ A} \quad (3)$$

As a result, the requirements for the MOSFET driver are:

- Gate drive current of 2 A or more
- $V_{OUT}$  of 10 V or more
- Floating channel that makes it possible to drive the high-side MOSFET without an additional level-shift module.
- Two delay-matched channels on one chip.
- 3.3 V input

The IR2110 IC satisfies these requirements [10].

### C. Keeping Eletro-Magnetic Radiation Low

To achieve high efficiencies, steep switching slopes are required. However, steep switching slopes provide a wide spectrum of frequencies. Therefore, the high current paths need to enclose only a tiny area to keep the radiated energy low. The design of the Printed Circuit Board (PCB) takes this aspect into account. For more information on the PCB, see Section III-A. The guidelines used to build a PCB with low radiation are provided in [11].

TABLE I. MICROCONTROLLER (MCU) COMPARISON.

MCU	Clock	PWM Clock	Power consumption	Ref:
ESP32 (Bluetooth)	240 MHz Dual	80 Mhz	0.1 W, 0.33 W (BL on)	[3]
TEENSY 4.1	600 MHz Dual	150 Mhz	0.5 W	[4]
STM32H723	550 MHz Single	275 Mhz	0.266 W	[5]
STM32H7A3	280 MHz Single	280 Mhz	0.17 W	[6]
STM32WB55 (Bluetooth)	64 MHz Dual	64 Mhz	0.175 W	[7]

### III. IMPLEMENTATION

This section documents details of the project implementation and shows how the challenges were overcome.

#### A. The power switching circuit

The amplifier consists of two parts: The microcontroller and the power switching circuit, which is controlled by the microcontroller. The power switching circuit was developed with the goal of achieving high efficiency, low radiation, and the possibility of being controlled by 3.3V microcontroller outputs.

The schematic of the developed PCB is shown in Figure 2. It comprises a MOSFET half bridge, a MOSFET driver (IR2110) and a low pass filter. Using two of these PCBs, a full bridge amplifier can be built.

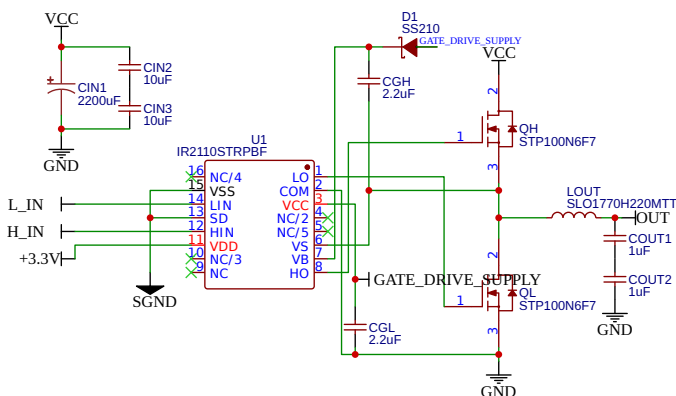
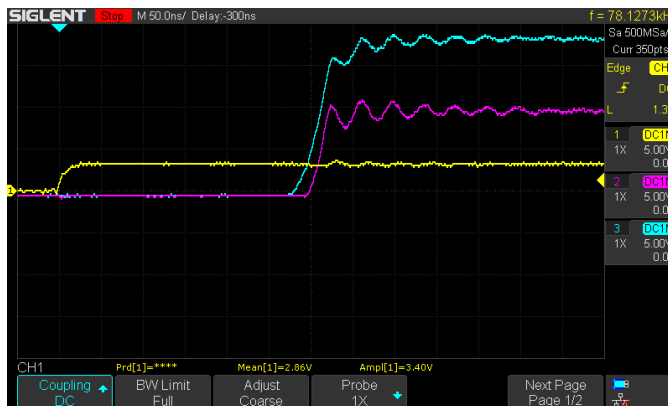


Figure 2. Schematic of the power switching circuit.

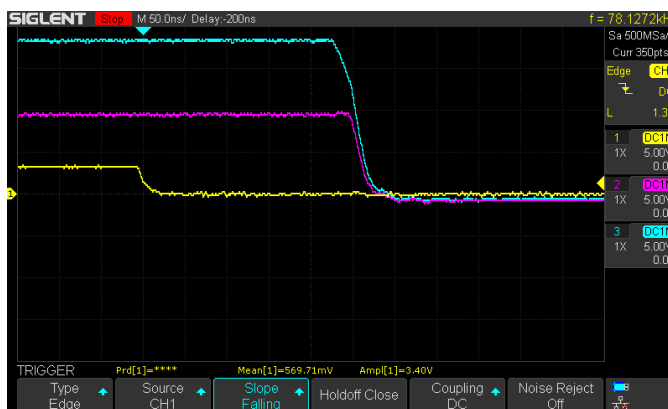
#### B. Timing

The correct timing of the switching slopes is a crucial part of a class-D amplifier. Any delay between the slopes of the high- and low-side MOSFETs causes distortions. If the slopes overlap, the supply current is shorted to ground, which results in destroyed MOSFETs.

Oscilloscope measurements (Figure 3) show that the turn on delay is 300 ns and the turn off delay is 250 ns. The turn-on/off delay is the time between a switching slope at an output of the microcontroller and the slope of the modulated power signal of the final amplifier. Fortunately, with chosen MOSFETs (STP100N6F7) the turn-on-delay is only 50ns slower than the turn-off-delay. The switching resolution of the ESP32 is  $1/80 \text{ MHz} = 12.5 \text{ ns}$ , so it is possible to improve the timing without further hardware.



(a) Rising edge.



(b) Falling Edge.

Figure 3. Measurements of slopes at different places on the PCB. Yellow: Output. Pink: MOSFET source. Cyan: MOSFET gate

#### C. Audio Measurement setup

For the audio measurement, a Behringer UCA202 audio interface was used. Its properties are provided in Table II. Most audio interfaces like the Behringer UCA202 yield low linear- and non-linear distortion for both input (ADC) and output (Digital to Analog Converter DAC) [12]. The downside is their low sample rate. If a signal is discretized with a sample rate lower than half of its highest frequency component, lower-frequency noise (aliasing effects) is generated. However, measurements have shown that the inbuilt anti-alias filter of the audio interface prevents this effectively. For the chosen modulation frequency of 78125 Hz, no aliasing effects were detectable.

If not specified, all measurements were done with a 10 V supply voltage, resulting in a maximum sinus output of around 7.07 Vrms.

TABLE II. PROPERTIES OF THE BEHRINGER UCA202 INTERFACE [13].

Property	Value
Model	Behringer UCA202
Sample rate	48 kHz
Bit depth	16 bit
Input impedance	27 kΩ
Max. input level	2 dBV
Max. input voltage	1.26 V Root Mean Square (RMS)

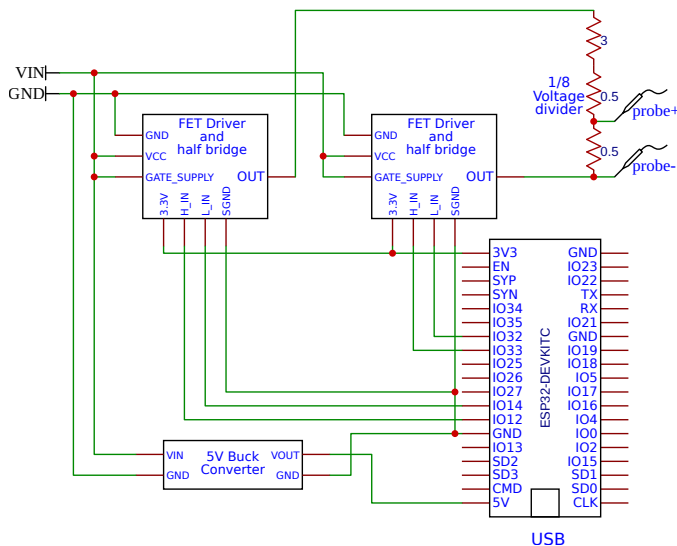


Figure 4. Schematic of test board. The detailed schematic of MOSFET driver and half bridge is shown in Figure 2.

The measurement setup is shown in Figure 4. It contains the following components:

- The ESP32 microcontroller
- Step-down voltage converter to provide the required voltage for the ESP32
- Two half bridges with integrated gate driver IC and LC-Filter as described in Section III-A
- Dummy load of 4 Ω, 20 W implemented as a voltage divider

The ESP32 provides four switching signals, high- and low-side for each half bridge amplifier. The outputs of the half-bridge amplifier PCBs are connected each to one side of the 4 Ω dummy load, which is realized using seven 1 Ω cement resistors to provide a 1/8 Voltage divider, allowing to connect the Behringer UCA202 audio interface to it, which has a maximum specified input voltage of 1 Vrms.

#### D. Fullbridge enhancement

If a full bridge amplifier is implemented, the resolution can be doubled by proper switching. For instance, if the resolution of each half bridge is 1024, in a full bridge amplifier, a value of 500.5 can be modulated by setting one half bridge duty cycle to 500 and the other to 501, effectively doubling the resolution to 2048. The result is a reduction in THD+N of almost 50 %, as shown in Table III. In Figure 5, a comparison of the frequency

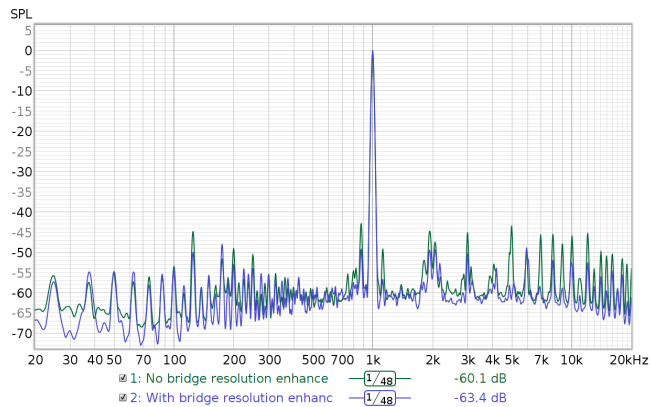


Figure 5. Comparison of the frequency spectrum with and without full bridge resolution enhancement. The two measurements of the amplifier output were done using the Behringer UCA202 audio interface and the software REW. A 1 kHz sine signal has been sent to the input.

spectrum with and without full bridge resolution enhancement is shown to underline these findings.

## IV. RESULTS

This section presents the results of the measurements that were conducted to show whether the implemented amplifier meets the defined objectives. The measurements include audio quality, power consumption, and output waveforms. The goal is to test the effectiveness of the implemented amplifier and to determine its suitability for the intended application.

### A. Audio

The low frequency response seems to be good. The frequency response is within ±0.8 dB for frequencies up to 5000 Hz, and the THD+N is lower than 1 % for frequencies up to 500 Hz.

Similar as in Section II-A, synthetic quantized signals were created and analyzed for comparison. REW calculations are provided for 11-bit, -20 dB signals at 1000 Hz and 100 Hz a THD+N of 0.46 % and 0.41 % respectively.

While the real measurement at 100 Hz is with 0.53 % close enough to argue that quantization is the main reason of distortion, this cannot be concluded for the measured value of 1.22 % at 1000 Hz.

At 10 kHz, the measured THD+N is 7.0 % which indicates that there is another source of distortion than quantization. The distortion is clearly audible at this frequency. Furthermore, if quantization is the main source of distortion, the THD+N is supposed to decrease anti proportional with the power (see "THD+N vs Power" in Figure 6). While this is the case with the 100 Hz measurements it is clearly not the case at 10 kHz.

What is the reason for the poor audio quality at high frequencies? To narrow down the problem, some of the audio data that is given to the *ampcontroller* software module was captured and the whole control loop of the *ampcontroller* software module, including the interpolation algorithm, was converted to a desktop program written in C. Then, the captured audio data has been processed with the desktop program and the

TABLE III. COMPARISON OF WITH AND WITHOUT FULL BRIDGE RESOLUTION ENHANCEMENT. THE TWO MEASUREMENTS OF THE AMPLIFIER OUTPUT WERE DONE USING THE BEHRINGER UCA202 AUDIO INTERFACE AND THE SOFTWARE REW. A 1 kHz SINE SIGNAL HAS BEEN SENT TO THE INPUT.

Before	Enhanced
Input RMS -26.09 dBFS	Input RMS -25.57 dBFS
-26.1 dBFS C, -26.1 dBFS A	-25.6 dBFS C, -25.6 dBFS A
-26.1 dBFS 22 - 22k UNW	-25.6 dBFS 22 - 22k UNW
-67.5 dBFS >22k	-71.8 dBFS >22k
Distortion at 1,000.0 Hz, -26.1 dBFS:	Distortion at 1,000.0 Hz, -25.6 dBFS:
THD: 3.03 % based on 18 harmonics [500..20000 Hz]	THD: 0.90 % based on 18 harmonics [500..20000 Hz]
HHD: 1.65 % [10..19]	HHD: 0.66 % [10..19]
N: 4.97 % [500..20000 Hz]	N: 3.14 % [500..20000 Hz]
N+D: -52.4 dBFS A	N+D: -56.6 dBFS A
<b>THD+N: 5.82 %</b> [500..20000 Hz]	<b>THD+N: 3.26 %</b> [500..20000 Hz]
2nd harmonic 0.21 %	2nd harmonic 0.16 %
3rd harmonic 0.96 %	3rd harmonic 0.49 %
4th harmonic 0.10 %	4th harmonic 0.039 %

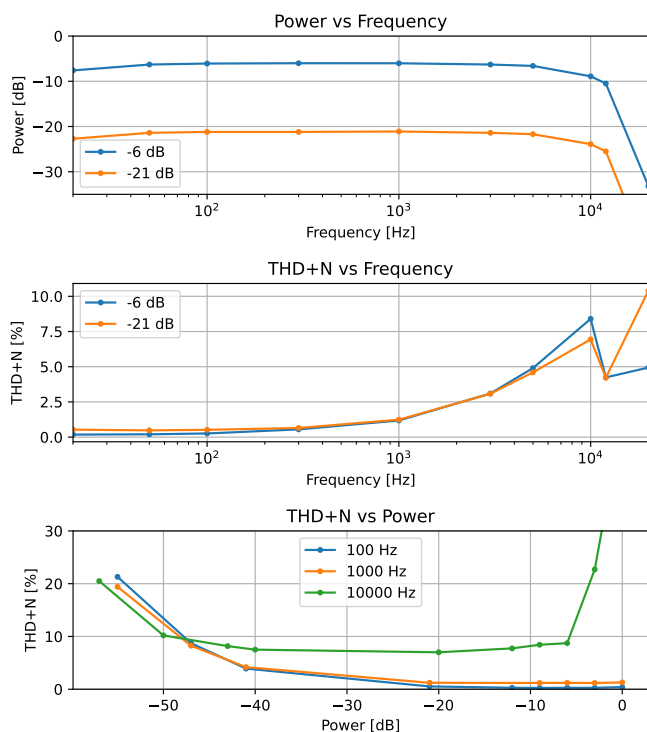


Figure 6. Visualization of audio quality measurements. Measured with Behringer UCA202 and REW as shown in Section III-C.

resulting signal was analyzed using REW. The high-frequency-problem did not occur.

Further investigation suggests that the effect can be ascribed to improper latching of the duty cycle by the microcontrollers timer-hardware. To overcome this, a different microcontroller must be used. This is a prospect for future work.

The idle power consumption is a crucial factor, particularly for mobile applications. The idle supply current was determined to be 158.5 mA with a digital multimeter and a supply voltage of 10 V. Therefore, the idle power consumption is 1.6 W at 10 V, as shown in (4).

$$P = U \cdot I = 10 \text{ V} \cdot 158.5 \text{ mA} = 1.585 \text{ W} \quad (4)$$

In case of the TAS5630B, the idle losses are 4.8 W using the full 48 V and 2.8 W using the minimal required voltage of 25 V [1], which means a reduction of at least 43 %.

To achieve very low idle losses a variable supply voltage is required, regardless of whether it is the TAS5630B or the microcontroller-based amplifier. While the TAS5630B has an absolute minimum of 2.8 W due to its minimum required supply voltage, there is no minimal required voltage for the power stage of the microcontroller-based amplifier. Therefore, it is possible to reduce the idle consumption of the microcontroller-based amplifier to the that of the microcontroller and the MOSFET driver, given that the volume is controlled by the supply current.

### B. Output waveform at different levels

The measurements shown in Figure 7 shall give an impression of how the quantization error degrades the output of the amplifier. If the signal volume is reduced to -54 dB, in the output there are only four signal levels left (see Figure 7b). This is expected, because the resolution is 2048 and the level is -54 dB (5).

$$-54 \text{ dB} \cdot 2048 \approx 0, 1995 \cdot 2048 \approx 4 \quad (5)$$

## V. CONCLUSION AND FUTURE WORK

The prototype demonstrates the potential of this simple approach in terms of energy efficiency, particularly for low-volume levels, without being limited to low-volume levels. Both noise shaping and multilevel switching could help to further increase the audio quality using the limited PWM frequency and resolution of a microcontroller.

### A. Multi Level Switching

The maximum PWM base frequency of any microcontroller is limited. An option to increase audio resolution without increasing the PWM base frequency is multi level switching. While a traditional half-bridge has two levels, LOW and HIGH, it is possible to build a circuit having more levels. This way, the resolution can be increased without requiring higher PWM resolution or frequency, but at the cost of requiring more PWM channels and a more complex output stage.

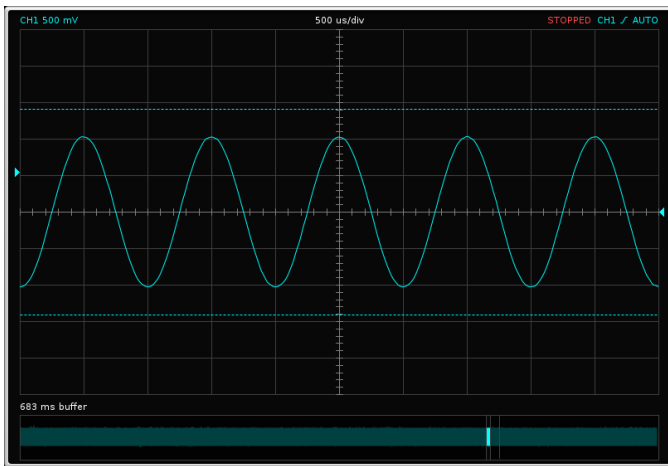
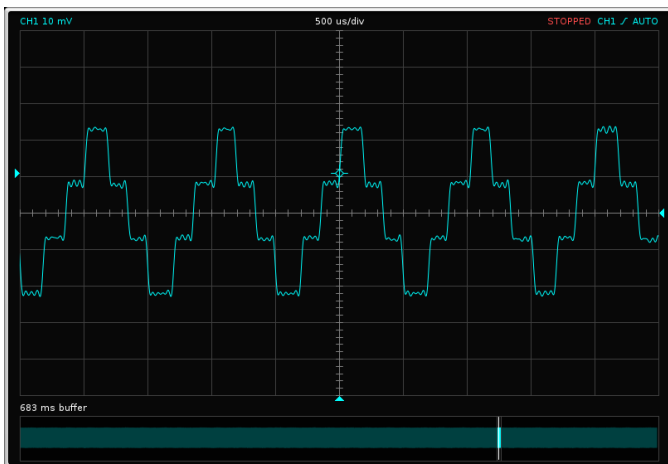
(a) -20 dB  $A = 1$  V.(b) -54 dB  $A = 20$  mV.

Figure 7. Output waveform (1 kHz) measured using Behringer UCA202 with an internal aliasing filter at different volume levels. The measurements are made without full bridge resolution enhancement.

Multi level switching has several benefits in class-D audio amplification compared to conventional two level switching, like lower idle power consumption, higher power efficiency and lower electromagnetic emission [14].

Multi level switching can be implemented by creating a power stage of two half-bridges with each four power MOSFETs and a flying capacitor.

An intermediate voltage supply is generated over the flying capacitor's terminals, which together with a smart MOSFET switching scheme result in a PWM output with three voltage levels rather than the conventional two. This doubles the switching frequency seen at the PWM output [14].

It is possible to increase the switching frequency seen at the load even further. Two half-bridges can be bridged with a relative phase shift of  $270^\circ$ , achieving a five level switching scheme across the load. This effectively quadruples the switching frequency seen at the load [14].

Furthermore, multi level switching reduces the voltage magnitude over the output filter inductor during switching,

which reduces the ripple current and relaxes the output filter requirements. Hysteresis losses in the inductor core are therefore also reduced, which improves overall power efficiency [14].

### B. Noise Shaping

Noise shaping is a technique typically used during bit depth reduction of a digital signal. Its purpose is to increase the apparent signal-to-noise ratio of the resulting signal. This is done by changing the spectral shape of the error introduced by dithering and quantisation, moving the noise power to less critical frequency bands. There is a popular noise-shaping algorithm used in image processing which is known as "Floyd Steinberg dithering" [15]. Noise shaping algorithms for audio applications are simpler than those for image processing, because of the less dimensional data. This approach is currently evaluated and will be published in future work.

### REFERENCES

- [1] Texas Instruments Incorporated, *TAS5630B 300-W Stereo and 400-W Mono PurePath HD Analog-Input Power Stage*, TAS5630B, Rev. D, May 2015.
- [2] Sure Electronics, "AA-AB32192 - 2x8W at 4 Ohm TPA3110 Class-D Audio Amplifier Board", 2019, [Online]. Available: <https://store.sure-electronics.com/product/AA-AB32192> [retrieved: 2024-09-24].
- [3] Espressif Systems, *ESP32 Series Datasheet*, ESP32, Version 4.6, Aug. 2024.
- [4] PJRC, "Teensy 4.1 Development Board", [Online]. Available: <https://www.pjrc.com/store/teensy41.html> [retrieved: 2024-10-21].
- [5] STMicroelectronics N.V., *Arm Cortex-M7 32-bit 550 MHz MCU, up to 1 MB Flash, 564 KB RAM, Ethernet, USB, 3x FD-CAN, Graphics, 2x 16-bit ADCs*, STM32H723, Rev. 4, Nov. 2023.
- [6] STMicroelectronics N.V., *32-bit Arm Cortex-M7 280 MHz MCUs, up to 2-Mbyte Flash memory, 1.4-Mbyte RAM, 46 com. and analog interfaces, SMPS*, STM32H7A3, Rev. 8, May 2022.
- [7] STMicroelectronics N.V., *Multiprotocol wireless 32-bit MCU Arm-based Cortex® -M4 with FPU, Bluetooth 5.4 and 802.15.4 radio solution*, STM32WB55, Rev. 16, Aug. 2023.
- [8] J. Mulcahy, *Room Equalization Wizard (REW)*, version V5.31.3, Jul. 2024.
- [9] STMicroelectronics N.V., *N-channel 60 V, 4.7 mOhm typ., 100 A STRipFET F7 Power MOSFET in a TO-220 package*, STP100N6F7, Rev. 2, Dec. 2014.
- [10] Infineon Technologies AG, *IR2110 HIGH AND LOW SIDE DRIVER*, IR2110, Rev. V, Jun. 2019.
- [11] L. Miller, "EMV-Optimiertes Schaltreglerlayout", Mar. 2012, [Online]. Available: <http://www.lothar-miller.de/s9y/categories/40-Layout-Schaltregler> [retrieved: 2024-09-24].
- [12] R. Mann and J. Vanderkooy, "An Open-Source Electroacoustic Measurement System Part 2: Sound Card Setup, System Characterization and a few more Examples", *Linear Audio*, vol. 13, pp. 159–204, Apr. 2017.
- [13] Behringer, *U-CONTROL UCA202 User Manual*, 2021.
- [14] Infineon Technologies AG, *3x37 W, 10 to 20 V, Ultra-Low Idle Power, MERUS Multilevel Switching Class-D Audio Amplifier with Inductor-Less Operation*, MA2304PNS, Rev. 2.2, Aug. 2022.
- [15] R. W. Floyd and L. Steinberg, "An Adaptive Algorithm for Spatial Grayscale", in *Proceedings of the Society of Information Display*, vol. 17, 1976, pp. 75–77.