KU LEUVEN

FACULTY OF ENGINEERING TECHNOLOGY

CAMPUS GROUP T LEUVEN

IB3

Baby Sleep Radio Monitor

Senne DE GREEF

Chieh Fei HSIUNG

Penelope REKKAS

Arthur TAVARES QUINTAO

Dries VAN RANST

Coach(es): Gil Vranken

Technical paper submitted to obtain the degree of Bachelor of Science in Engineering Technology: Electronics Engineering

Academic Year 2022 - 2023

KU LEUVEN

Faculty of Engineering Technology Campus GROUP T Leuven — Technical paper submitted to obtain the degree of Bachelor of Science in Engineering Technology: Electronics Engineering
Academic Year 2022 - 2023

IB3

Baby Sleep Radio Monitor

De Greef Senne, Hsiung Chieh Fei, Rekkas Penelope, Tavares Quintao Arthur, Van Ranst Dries

Bachelor in Electronics Engineering, Faculty of Engineering Technology, Campus GROUP T Leuven, Andreas Vesaliusstraat 13, 3000 Leuven, Belgium

Coach(es): Gil Vranken

Electronics Engineering, Faculty of Engineering Technology, Campus GROUP T Leuven, Andreas Vesaliusstraat 13, 3000 Leuven, Belgium, gil.vranken@KULeuven.be

ABSTRACT

The Baby Radio Monitoring System (BRMS) is a device designed to address Sudden Infant Death Syndrome (SIDS), the leading cause of death for one-month- to one-year-olds. The design integrates a room temperature and sleeping position monitor since both factors are thought to increase the risk of SIDS, according to the Centers for Disease Control and Prevention. The BRMS is furthermore capable of detecting whether a baby is crying. All temperature, position and audio data is captured at the baby's side, processed by a microcontroller, and then sent wirelessly to the parent's side.

The temperature monitor was built using a thermistor as its core component. An STM8S105K4 microcontroller closely measures the voltage across this temperature-dependent resistor. This way, the BRMS is capable of tracking the temperature in the baby's room with up to 0.5 $^{\circ}$C precision.

The position monitor was built using a Hall sensor as its core component. In addition, a magnet is placed onto the baby's chest which interacts with the Hall sensor. The voltage across the sensor changes according to the magnetic field strength applied to it. Again, the STM8 records this voltage and can hence determine in which position the baby is sleeping.

For the cry detector, audio signals arriving from a microphone were processed with an ESP32-WROOM-32 microcontroller using digital processing techniques. The cry detector achieved an 88.4% accuracy in distinguishing infant cries, showing the effectiveness of the cry detection algorithm.

Whenever crying, flipping or unsafe temperature is detected at the infant's side of the monitor, the parent is warned through an LCD display present at their side. Hence, the BRMS is an effective baby monitoring system.

1 INTRODUCTION

Sudden Infant Death Syndrome (SIDS) is the leading cause of death among babies aged one month to one year. In 2017, approximately 1360 babies died of SIDS according to the Centers for Disease Control and Prevention (1). While the exact cause of SIDS is unknown, research has shown that environmental and physical factors can increase the risk of SIDS. Factors such as sleeping position and temperature have been linked to SIDS, and monitoring these factors can reduce the risk according to the American Academy of Pediatrics (2).

To tackle these two risks, we developed a microcontroller-based Baby Radio Monitoring System (BRMS) that tracks the temperature inside the infant's sleeping room and alerts parents if the temperature reaches an unsafe level. In addition, the system tracks the baby's position in bed, as sleeping in the supine position, i.e., lying on the back, can reduce the risk of SIDS. Finally, the system also monitors whether or not the baby is crying, notifies parents accordingly and plays a lullaby to the baby.

The next section will analyze the requirements for building our system. Following the requirement analysis, this technical paper is divided into four main chapters: Design and Materials, Implementation, Evaluation and Validation, and Discussion. In "Design and Materials", the top-level design of the BRMS is explained, as well as the materials and components used in this project. The "Implementation" section elaborates on our hardware and software implementations for this project. The "Evaluation and Validation" provides tests and measurements on the performance of the BRMS, related to the requirements proposed in the coming paragraph "Requirements analysis". Finally, these test results will then be discussed in the "Discussion" section.

1.1 Requirements analysis

The AAP recommends that the temperature in the baby's room should be between 20 ℃ and 22.2 ℃ (2). Hence, our system must monitor this range with 0.1 ℃ precision. Moreover, because the CDC recommends placing babies on their backs (1), our system must be able to detect if the baby is in a prone position, i.e., lying on the stomach, and alert the caregiver. Finally, detecting crying sounds is crucial for caregivers to respond promptly to the baby's needs. To ensure the baby's well-being, we aim for an 80% cry detection rate. Even if the audio transmission is switched off, our system should accurately notify parents that the baby is crying and play a lullaby.

To ensure that parents can monitor their baby's sounds from different rooms in a house and achieve the required temperature precision and position detection, the system will be split into three modules: two in the baby's room and one at the parent's side. The first module, Baby Module A (green block in Figure 1.1), placed under the mattress, will monitor temperature and sleeping position, while the second, Baby Module B (blue block in Figure 1.1), positioned 1-2 meters away, will capture sounds in the room and detect crying.

The modules must communicate wirelessly with the Parent's Module (orange block in Figure 1.1) within a range of up to 15 meters. The Parent's Module will perform several functions, such as displaying the temperature in the baby's room, playing the sounds emitted at the baby's end of the monitor, and alerting the caregiver when the room temperature falls outside the allowed range or the baby is in a prone sleeping position. Baby Module A should not consume more than 2 Watts of power such that it could be battery-powered if implemented as a standalone product.

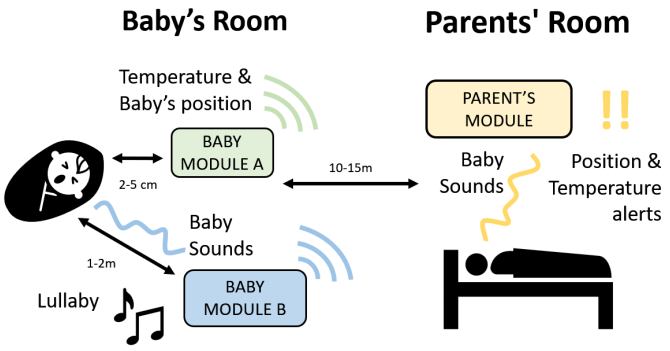

Figure 1.1: Top level block diagram

2 DESIGN AND MATERIALS

On the one hand, Baby Module A is designed to support an STM8S105K4 microcontroller chip (STM8) (3) from ST Microelectronics. On the other hand, Baby Module B and the Parent's Module are designed around an ESP32-WROOM-32 microcontroller development board (ESP32) from Espressif Systems (4).

The decision to use two different microcontrollers in the system is due to their varying capabilities. An 8-bit microcontroller like the STM8 is limited in its ability to perform complex tasks like Digital Signal Pro-

cessing (DSP) for cry detection and audio streaming between the Parent's Module and Baby Module B. In contrast, the ESP32 is equipped with a Wi-Fi module, a 12-bit built-in analogue-to-digital converter (ADC), and an 8-bit digital-to-analogue converter (DAC), making it a cost-effective solution for DSP and audio streaming. However, for tasks that allow high latency and require low power consumption, such as temperature and position detection, the STM8 is a more efficient choice.

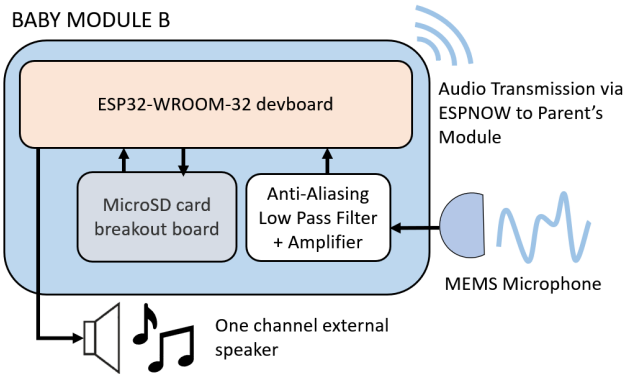

Figure 2.2: Block diagram of Baby Module B

2.1 Baby Module A

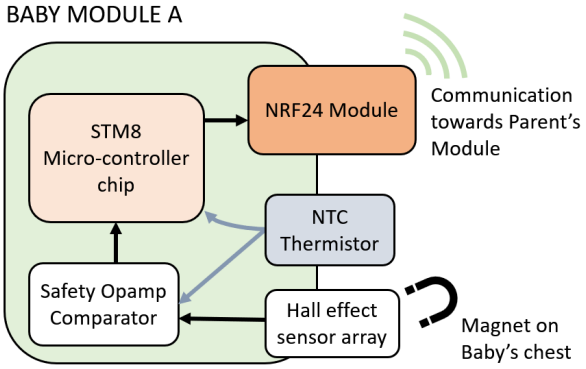

Figure 2.1: Block diagram of Baby Module A

Aside from the STM8, Module A was equipped with an NTCLE100E3222*B0 thermistor (5), an AH277A Hall Sensor (6), an LM311 comparator, and a Si24R1 RF module (which is a clone of the nRF24 module) for wireless communication (7). A 0.02m x 0.01m N42 neodymium magnet is used in combination with the Hall sensor.

As seen in Figure 2.1, a Hall effect sensor is used to detect a change in the polarity of a magnetic field. In the BRMS, a two-dimensional array of Hall sensors is integrated into a soft mattress that is placed under the baby. Additionally, a neodymium magnet is attached to the infant's pyjamas at chest level. This way, in case the baby flips onto its belly, the position and orientation of the magnet change. This is detected by the Hall sensors.

2.2 Baby Module B

To meet the requirements outlined in section 1.1, a real-time wireless transmission system with cry detection signals is necessary. In addition to an ESP32, the proposed Module B, depicted in Figure 2.2, includes an S15OT421 MEMS microphone (8) with a signal-to-noise ratio of 58 dB for input signal pre-conditioning. A 6th-order Butterworth anti-aliasing filter with monotonic and linear delay is used to condition the microphone signal. Moreover, a MicroSD card breakout board (9) is included for storing pre-recorded lullabies and evaluating the DSP performance. The system, in addition, includes a female AUX connector to attach a one-channel external loudspeaker to play lullabies.

2.3 Parent's Module

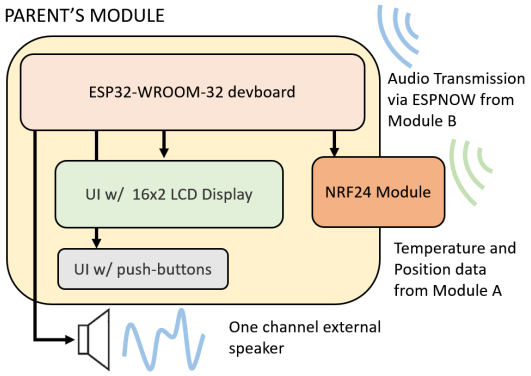

Figure 2.3: Block diagram for the Parent's Module

As shown in Figure 2.3, on the Parent's Module, the off-the-self components connected to the ESP32 development board are a 16x2 LCD display, four push-buttons, a Si24R1 RF breakout board and an AUX female connector. The Radio Frequency module is used to receive the temperature and baby position data from Module A. This data is presented to the user with the LCD display. Using the interface, the user can navigate using the four push buttons added to the module. Finally, the AUX connector allows the audio coming from Module B to be played to the parents or other caregivers.

3 IMPLEMENTATION

The off-the-shelf components used in the three BSRM modules have been integrated into their respective Printed Circuit Boards (PCBs) along with the necessary electrical circuits for their optimal performance. In addition to the hardware implementation, extensive software development was required for both the STM8 and ESP32 microcontrollers, with programming conducted in C language.

3.1 Hardware implementation

In terms of hardware development, the temperature and prone position detection on the STM8 required the most work. On the modules with ESP32s, the circuits related to audio pre-processing were the most complex hardware features.

It is worth mentioning that given the similarities between Baby Module B and the Parent's Module, we were able to design the same PCB for the two modules but solder only the components used by each module. This way, we reduced the number of different designs to be manufactured since most suppliers require at least five boards to be ordered per design. The PCB board designs can be found at Appendix G

3.1.1 Baby Module A: Temperature monitoring

The temperature sensor used is the NTCLE100E3222*B0 thermistor. This module is placed in the same room as the baby's cradle. On the circuit board itself, the NTC circuit is driven by a 5 V supply. The thermistor is placed in series with a 2.2 k-Ohm resistor. The node between the resistor and thermistor is directly connected to an LM311 comparator.[1]

When the temperature rises above 25 °C, or below 20 °C, the comparator will set the input pin of the STM. Both temperature thresholds are manually set by picking the right resistor values.

According to its datasheet (5), the NTC has a resistance of 2.2 k-Ohm at 25 °C. When the temperature increases, the resistance of the NTC decreases, thereby reducing the voltage over the NTC. Consequently, the high-temperature threshold comparator IC3 will set its output high. The low-temperature threshold circuit applies the same principles, albeit with different resistor values. At 20 °C, the NTC has a resistance of 2.747 k-Ohm, thus corresponding to a voltage drop of 2.78 V in the voltage divider. By picking a 2.76 k-Ohm resistance, the temperature needs to drop slightly below 20 °C before the comparator will set its output high.

A third feature of the thermistor circuit is the Analog-to-Digital Conversion (ADC) to measure the actual temperature, instead of monitoring a threshold. This is obtained by measuring the voltage across the NTC and inputting this value into one of the STM8's ADC channels (3). More about this feature in section 3.2.1.

3.1.2 Baby Module A: Prone position detection

Figure B.1 in the Appendix represents a simplified model of the Hall sensor circuit with a 2x2 array of Hall sensors. In practice, the dimensions of this array need to be adapted to the size of the baby's bed to ensure there are no blind spots where the baby cannot be detected. To adapt the array dimensions, one would simply need to put an additional amount of Hall sensors in parallel to the circuit shown in the figure. [2]

The circuit is driven by a 5 V supply and works as follows: the monitor (i.e. the microcontroller) is connected to an LM311 comparator which, in turn, is connected to the Hall sensor array. The Hall sensors detect the changing magnetic field strength of the magnet on the baby's chest. This causes a change in voltage across the sensors. Consequently, this triggers the comparator to send a signal to one of the STM8 I/O pins where it is used to activate an alarm.

3.1.3 Baby Module B: Audio signal conditioning

For a real-time signal conditioning stage, a flat magnitude and specifically a linear group delay offered by the Butterworth low-pass filter was an ideal choice. The quality factors for the cascaded three stages low-pass filter were selected as 1.932, 0.707, and 0.518. The cut-off frequency was designed 8,000 Hz, away from the upper bound of 22,050 Hz constrained by the 44,100 Hz ADC sampling rate (f_s). The design is outlined in Appendix F.1.

3.1.4 Parent's Module: Audio output

The audio captured at the baby's side is received from the Baby Module B via the Wi-Fi module's I2S

[1] For an overview of the NTC sensor circuit, refer to Appendix B (NTC Sensor Circuit A.1).

[2] For an overview of the Hall sensor circuit, refer to the Appendix (C.1).

interface. It is then processed by the ESP32 microcontroller, specifically utilizing its built-in 8-bit DAC which provides sufficient resolution to discern if the input signal corresponds to crying or other sounds.

3.2 Software implementation

In order to process the input from the sensors and transmit data between Module A, B and the Parent's Module, the STM8 and the ESP32 had to be programmed. In particular, algorithms for temperature threshold, prone position detection and SPI communication with the radio frequency module had to be programmed on the STM8.

The most important software implementations in the ESP32s were, the DSP for the baby cry detection, the audio streaming via ESPNOW, the connection to Module A via Si24R1 and the User interface.

3.2.1 Baby Module A: Temperature monitoring

As explained in 3.1.1, an analogue voltage serves as the input of the STM8 where it is converted into a digital value. Next, the digital result needs to be scaled to the appropriate range. Then, the resistance of the thermistor needs to be calculated based on the voltage just obtained. Subsequently, the calculated resistance needs to be equated to the corresponding temperature for the given thermistor. Finally, the temperature needs to be converted from Kelvin to Celsius. All these steps are programmed into the STM8 microcontroller.

Since the possible voltage range across the thermistor is from 0 V to the supply voltage of 5 V and ADC is 10-bit long for this microcontroller, the ADC result V_{ADC} is first scaled in order to obtain the voltage across the thermistor V_{NTC} in Volts.

$$V_{\text{NTC}} = \frac{V_{\text{ADC}}}{2^{10}} \cdot 5$$

Assuming R_{NTC} is the temperature-dependent resistance of the thermistor, and R_{REF} is the resistance of the resistor in series with the NTC, it is known that

$$V_{\text{NTC}} = V_{\text{DD}} \cdot \left(\frac{R_{NTC}}{R_{REF} + R_{NTC}} \right)$$

The following formula is derived from the previous one.

$$R_{\text{NTC}} = \frac{V_{\text{NTC}} \cdot R_{\text{REF}}}{V_{\text{DD}} - V_{\text{NTC}}} = \frac{V_{\text{NTC}} \cdot 2200}{5 - V_{\text{NTC}}}$$

That is how the temperature-dependent resistance across the thermistor can be calculated from the voltage across the thermistor.

When the resistance across the thermistor is known, the Steinhart-Hart equation is applied in order to obtain the corresponding resistance in Kelvin.

$$T_{\text{K}} = \frac{1}{A + B \cdot \ln\left(\frac{R_{\text{NTC}}}{R_{\text{REF}}}\right) + C \cdot \left(\ln\left(\frac{R_{\text{NTC}}}{R_{\text{REF}}}\right)\right)^2 + D \cdot \left(\ln\left(\frac{R_{\text{NTC}}}{R_{\text{REF}}}\right)\right)^3}$$

In the above formula, A = 0.003354016, B = 0.0002569850, C = 0.000002620131 and D = 0.0000000638309 according to the datasheet of the thermistor (5).

Finally, the temperature is converted from Kelvin to degrees Celsius as follows.

$$T = T_{\text{K}} - 273.15$$

Once these calculations are performed by the STM8, the final result (i.e. a temperature reading in degrees Celsius) is sent to the ESP32 microcontroller of the Parent's Module.

The ESP32 is programmed to follow the Finite State Machine (FSM) in figure 3.1. The ESP32 updates the temperature displayed on the LCD display with the latest temperature recorded by the STM8 and checks whether that temperature is within the accepted range. If the temperature doesn't fall within that range, it activates an alarm at the parent's side of the monitor.

When considering the bounds of the temperature range, the ESP32 takes into account those set by the user via the LCD unless they have not been set, in which case the default range will be used. The default range is set between 20 and 25 ℃ since the ideal room temperature for an infant below the age of 1, according to experts, is 20 to 22.2 ℃ (10).

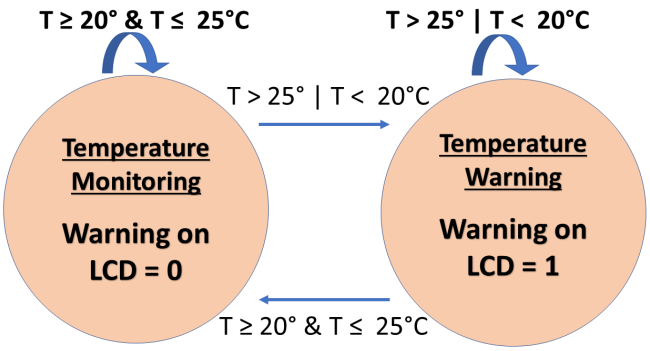

Figure 3.1: FSM for temperature monitoring

3.2.2 Baby Module A: Position monitoring

The FSM in Figure 3.2 explains the logic behind the software code, which is programmed into the STM8 microcontroller. Whenever a specific GPIO pin detects a high signal coming from the comparator, the STM8 notifies the RF module to send a signal to the master node where it will trigger an alarm. This way, the caretakers are notified when the baby is sleeping in a potentially dangerous position.

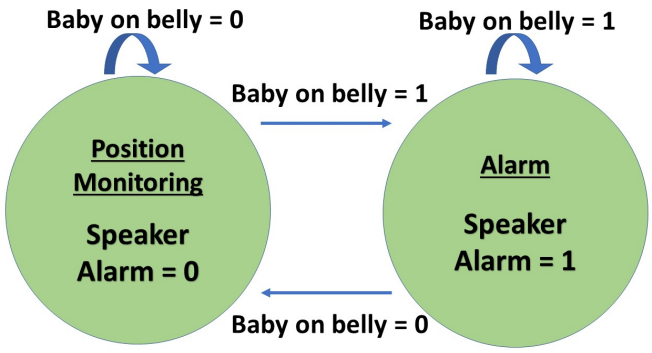

Figure 3.2: FSM for position monitoring

3.2.3 Baby Module B: Cry detection

During their first year, babies undergo four main language milestones: crying, cooing, laughing, and babbling. While crying and laughing are self-explanatory, cooing represents the stage where babies produce vowel and gurgling sounds, leading up to the next milestone of babbling, where babies introduce consonants in their vocabulary, allowing them to produce repeated syllables.

Hence, our monitor was programmed to not only differentiate between infant-generated sounds and environment-generated sounds but also to distinguish crying from other types of infant-generated sounds.

The distinction between crying and other types of infant-generated sounds is established upon identifying a set of features that are simultaneously present in a crying sample but never (or rarely) simultaneously present in an infant-generated signal other than crying. In order to come up with an efficient set of features for that purpose, insights from both the literature and our analysis, which also serve to validate the literature findings, were used.

3.2.3.1 Literature review

An important feature of sound in the frequency domain is the fundamental frequency (F_0) and its harmonics (F_n).

Thanks to the researchs including frequency spectrum analysis of categorized infant voice signals (Van Dyke et al. 2017 (11), Ji et al (12)), the following is known. The F_0 range of cry signals for infants aged within 12 months falls between 400 Hz and 550 Hz whereas that of cooing is between 300 Hz and 400 Hz. In addition, the dominant first formant (F_1) of babbling signals associated with the production of vowel phoneme utterance falls within the range of 675 Hz to 1000 Hz. Besides, the F_0 of adult speech signals is typically between 85 Hz to 250 Hz.

Therefore, an overlap between the range associated with crying and that of another type of infant-generated sound is likely and identifying crying based on the fundamental frequency only would not suffice. That is why both fundamental frequency and additional sound features are investigated in the following section.

3.2.3.2 Data analysis

A MATLAB script analysis was applied to 25 infant crying audio samples and 15 cooing, babbling, or laughing audio samples in order to extract their zero-crossing rate (ZCR), Fast-Fourier-Transform (FFT), Welch-window Power Spectral Density (PSD), and root mean square (RMS) intensity.

The computation of the FFT, followed by the Welch-windowed PSD, enables the identification of the F_0 and its F_n in a given sample. By normalizing the amplitude and calculating the energy for each frequency sweep,the dominant frequency among F_0, F_1, and F_2 could be determined.

Our insights from that analysis are presented in what follows.

3.2.3.2.1 Crying versus cooing/babbling

The crying signals were observed to have two dominant peaks, F_0 and F_2, respectively at frequencies 439.926 Hz and 1500 Hz. Hence, the fundamental frequency obtained is consistent with the research. Additionally, the second peak occurring at the second formant F_2 was found between 1150 Hz and 1500 Hz.

Based on the samples analyzed, the first formant, F_1, is considered irrelevant as it contains low power.

The mean values for the F_0 of the cooing and babbling envelope also corresponded to the findings of the reviewed literature (13). Moreover, cooing and babbling signals in a short-timed frame displays only one dominant peak, which is one peak less than in a typical crying pattern. The location at which that single peak occurs is dependent on the vowel pronounced within the babbling/cooing sound. An /o/ causes a dominant frequency to occur at 357 Hz, an /i/ at 619.08 Hz and an /a/ at 1000 Hz.

Another important distinction can be established between crying and cooing on the basis of the ZCR. The average ZCR for crying signals was measured to be 0.100, whereas the average ZCR for cooing signals was 0.058. Thus, a ZCR higher or equal to 0.1 can be considered as a characteristic feature of an infant crying signal. Figure 3.3 illustrates a time-domain input signal comparison between crying and cooing.

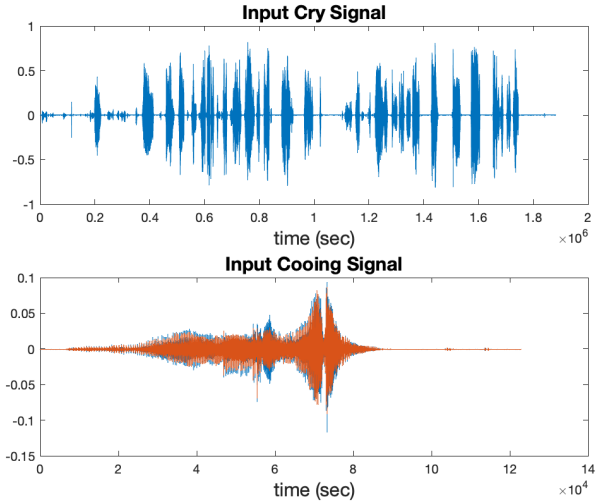

Figure 3.3: Infant crying and cooing signals in the time domain

Figure 3.3 depicts the time-domain signal comparison between a cooing signal with a crying signal. The cooing signal exhibited a smoother amplitude envelope, indicating a lower rate of amplitude transitions between consecutive signals. The characteristic aligned with the fluency nature of infant speech that is resembled by cooing. In contrast, the crying signal displayed more significant amplitude swings more frequently, indicating a higher magnitude of amplitude transitions. The discrepancy was due to the fundamental characteristics of a cry signal. As crying would comprise short bursts of energy-intensive signals, a higher acceleration of signal amplitudes was expected.

3.2.3.2.2 Crying versus laughing

The laughing signals were found to have a fundamental frequency range between 400 and 500 Hz, which overlaps with a crying signal's F_0 range. Additionally, the ZCR range of laughing signals ([0.05,0.09]) was found to have its upper bound overlapping with the lower bound of the crying signal ZCR ([0.07,0.22]).

Despite this similarity, we observed a 139.72 Hz difference between the second frequency peak of laughing signals and that of crying signals. Specifically, both crying and laughing signals showed a second peak in the PSD, but the second peak for crying occurred at a mean frequency of 1100 Hz, and the second peak for laughing at a mean frequency of 960 Hz. The distinction in the second peak suggested that a more rigid secondary frequency peak value should be set in the algorithm-defined cry signal.

3.2.3.3 Threshold and algorithm design

In light of the aforementioned findings, the acoustic feature extraction and threshold detection system for infant crying was designed as follows.

The frequency threshold settings for the cry identification algorithm include a required F_0 in the range of 400 Hz to 550 Hz and a F_2 in the range of 1150 Hz to 1500 Hz, as observed in Figure 3.4. In addition, the temporal-based threshold setting involves a ZCR greater than or equal to 0.10 and an amplitude 20 dB higher than the surrounding environmental noise level according to (11).

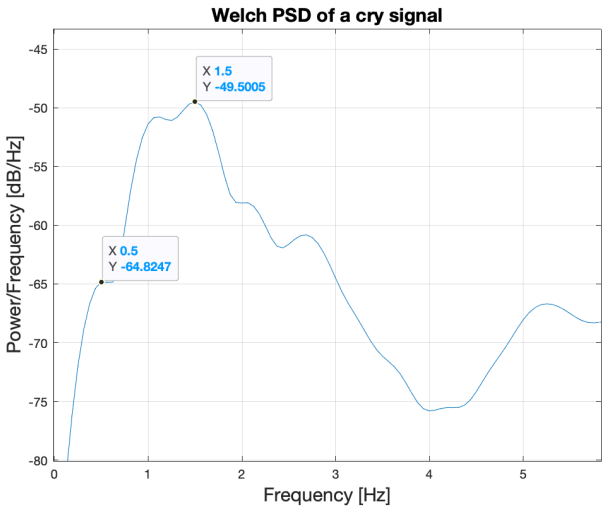

Figure 3.4: Welch-windowed PSD of a cry signal

When the filtered and digitized signal arrives on the

ESP32 of Module B, it is sent to the FFT in frames of 10 milliseconds for audio feature extraction. By employing block-computing with rectangular windowing and pre-computed twiddle factors, a radix-two FFT algorithm is utilized to generate an almost real-time frequency spectrum with minimal smearing.

The twiddle factors are computed in the following manner.

$$W_N^k = \cos\frac{2\pi \cdot k}{N} + j\sin\frac{2\pi \cdot k}{N} \tag{3.1}$$

The frequency spectrum of the signals is then normalized, and the peaks of the spectrum are distinguished in PSD. The first peak is considered as the fundamental frequency F_0, and the second peak as the second-harmonic pitch (F_2).

Once the F_0 and F_2 satisfy the threshold, the system enters the classification stage. Time-domain feature analysis ZCR is triggered and computed. A frame is classified as a cry if its ZCR is above 0.1 and its amplitude is 20 dB greater than pre-set environmental value.

If a cry signal is confirmed, the classifier enables the speaker to play a cry-soothing lullaby at the baby's end of the monitor.

3.2.4 Wireless communication

3.2.4.1 Baby Module A: Si24R1 on STM8

The software needed to control the Si24R1 via STM8 was developed according the STM8S105K4 (3) and the nRF24 (7) datasheets. They explain the specific pinout mapping that had to be used. The needed pinout mapping can be found in Appendix D.1. Without the correct connection, transmission of data is impossible.

Moreover, all 30 registers had to be set separately to the correct values specified in the nRF24 datasheet (7). The most important registers to make the transmission work are the ones specified in Appendix D.2. It was important that both the transmitting and receiving sides had the same registers set.

The software written follows the flowchart depicted in Figure 3.5. It starts by reading the temperature and position status of the sensors at every cycle, as explained in sections 3.2.1 and 3.2.2. Then, the data is transmitted to the Parent's Module as an array `value[3]`, which comprises the lower byte of the temperature, the upper byte of the temperature, and a position status bit. The `nR24L01_write_payload()` function employs the SPI interface to send the data from the STM8 to the nRF24 module, which then transmits the data to the Parent's Module.

Following the transmission, a delay is introduced using the `delay_ms()` function before the loop restarts.

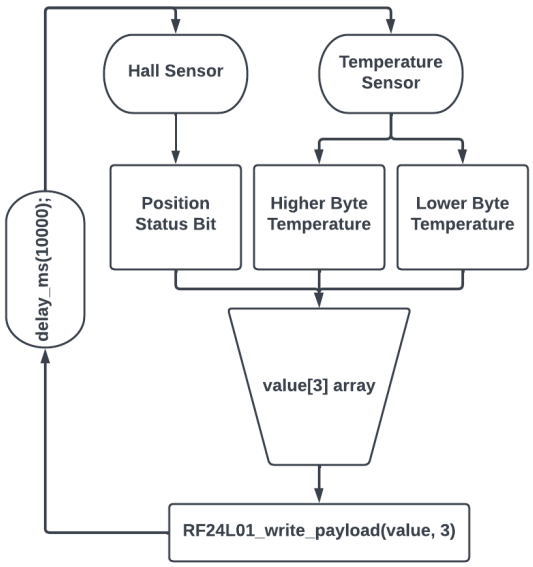

Figure 3.5: STM8 Flowchart

3.2.4.2 Parent's Module: Si24R1 on ESP32

Regarding the software used to interface an ESP32 with an STM8 microcontroller via nRF24, the code utilizes the MIRF library (14). It consists of a C implementation of the nRF24L01+ radio module SPI toolpack that also works for the Si24R1 RF module. The library provides an abstraction layer for communication with the radio module, making it easier to use.

Two functions were programmed: a transmitter and a receiver. The ESP32 transmitter was utilized during development, while the receiver code was implemented in the final version.

The receiver function continually monitors the radio channel for incoming data, which is then stored in a C struct containing three data bytes and a time variable. The received data is sent to a StreamBuffer for further processing.

Conversely, the transmitter function utilizes the same struct to send data through the radio channel, using MIRF library's `nRF24getData()` function. Once the data is received by the receiver, the data is sent to

the StreamBuffer.

The main function creates the StreamBuffer with a specific buffer size and trigger level. If the creation of the StreamBuffer fails, an error message is logged. The StreamBuffer is then passed to two initialization tasks, one for receiving the Si24R1 data and the other for User Interface, which establishes a communication link between the two tasks and initializes the nRF24L01+ module.

3.2.4.3 Baby Module B and Parent's Module: ESP-NOW

To optimize the data transfer rate and achieve low latency for audio signals, a wireless communication protocol ESP-NOW based on IEEE802.11 and a rate of 72 Mbps at 20 MHz were used. As the 12-bit ADC one-channel audio signal sampled at 44100 Hz produces a bit rate of 529.2 Kbps, this wireless transmission rate provides sufficient bandwidth for high-quality audio streaming without disrupting communication. The ESP-NOW wireless transmitter sends 250 bytes per packet and an identical transmission rate is set on the receiver end to prevent aliasing in signal sampling.

3.2.5 Parent's Module: User interface

By means of an I2C-enabled 16x2 LCD display(15) and four push buttons, a User Interface was developed to allow the user to interact with the monitor. The first two buttons are used to select and cancel state changes. The two other buttons allow the user to navigate through the four menu screens.

The different states of the FSM on Figure 3.6 correspond to different screens on the LCD1602 display and allow the user to easily navigate through the system's features. David Antliff's `esp32-i2c-lcd1602-example` library (16) was used as a reference to control the LCD display in the FSM. The library provides an example code to interface with the LCD1602 display using the I2C protocol. Additionally, `esp32-smbus` (17) and `esp32-i2c-lcd1602` (18) libraries were also needed as dependencies to handle the I2C communication and control the LCD1602 display.

To allow user input, the code initializes four GPIO pins to receive input from buttons and the debounce time is set to 300 ms. One interrupt handler was written for each button. The functions update the current

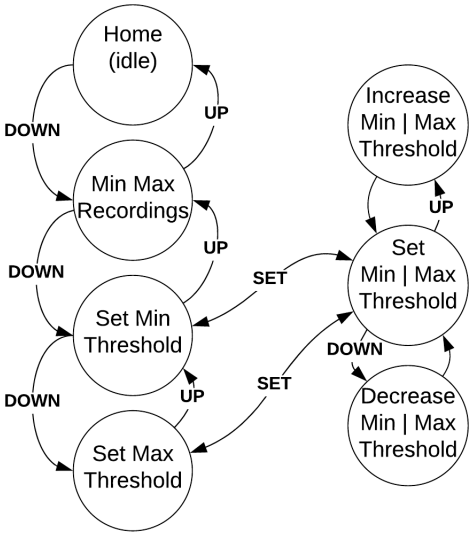

Figure 3.6: LCD display menu FSM

state of the system depending on the current state and button press.

In the main function of the source code written for the Parent's Module, an initialization function sets up the I2C interface and starts the LCD1602 display while a "datetime" function logs the current date and time on the home screen. The `updateMinMaxTemp()` function is a helper function used by the home task is used to update the minimum and maximum temperature readings while the `homeTask()` function reads the stream of data from the `nrfDataXStream` and displays the temperature readings and baby position information on the LCD1602 display.

A switch case statement is used to perform actions based on the current state of the system. The first screen displays the measured temperature inside the baby's room and a temperature and baby-flip alert in the appropriate circumstances. The second screen displays the maxima and minima temperatures recorded in the baby's room. The third and fourth screens allow the user to set the maximum and minimum temperature thresholds that condition the alarm on the first screen.

4 EVALUATION AND VALIDATION

4.1 Power consumption

To estimate the power consumption of the temperature and position monitor, all input currents into the circuit were measured. By multiplying those with the supplied power, the consumed power is calculated.

Appendix E.1 summarises the measurements and maximum power ratings. The table shows that the power consumption of the temperature and position applications stays below 1.8 W under normal operating conditions.

4.2 Temperature precision

The precision of the NTC was assessed by comparing its calculated values with the corresponding measured values obtained using a thermocouple connected to a multimeter. Our measurements indicate that the NTC has a mean error of 0.5 °C.

Graph 4.1 presents the temperature profiles of the calculated values (red), measured values (black), and the corresponding error (green).

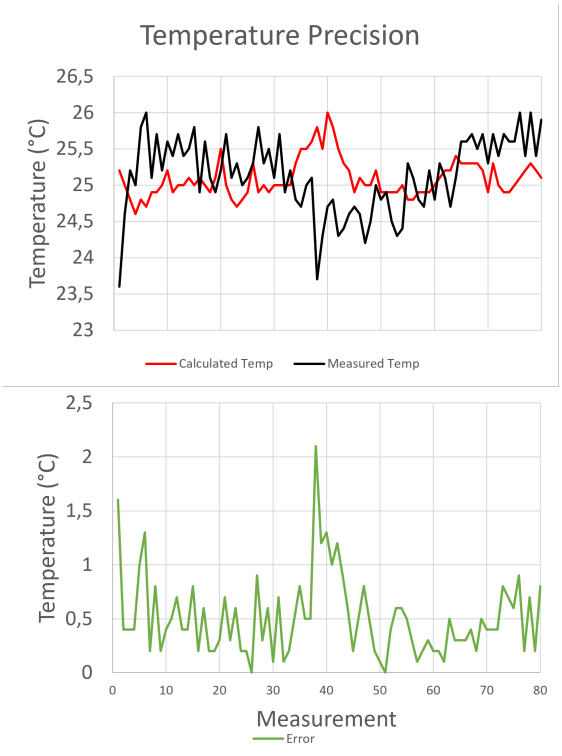

Figure 4.1: Temperature Precision

4.3 Hall sensor detection

The maximum allowed distance between Hall sensor and N42 neodymium magnet was tested with the help of an oscilloscope. We checked until what distance changing magnetic fields could be detected by the Hall sensor. It was observed that detection was possible until a distance of 4 cm.

4.4 Infant cry classification

The performance metrics, such as the precision of the frequency value extracted and the accuracy of the cry segmentation algorithm, were studied to evaluate the reliability of the infant cry detection system. A total of six types of signals, encompassing both infant-generated and environment-generated sounds, were collected and applied to the cry detection system. The infant-generated sounds included were crying, cooing, babbling, and laughing. In addition, pop-music and adult speech signals were considered as environment-generated sounds.

The precision of the F_0 was calculated in response to the PSD envelope of each data set and was recorded into a spreadsheet generated by the testing program. During testing, a signal recording task thread was running concurrently with the frequency spectrum analyzing sequence. The concurrent analysis approach ensured the coherence of the data frame analyzed.

The mapping analysis between F_0 computed by the cry detection system and the F_0 analyzed by the MATLAB script can be found in Figure 4.2. The figure described in frequency domain a one-to-one mapping relationship. A F_0 computed by the cry detection system was expected to deviate no more than 43.06 Hz from the F_0 analyzed by MATLAB script.

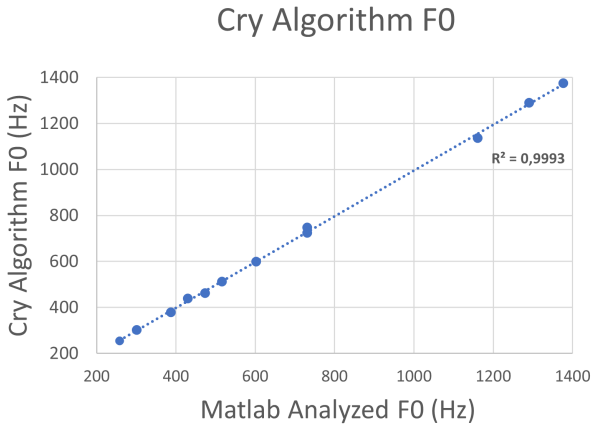

Figure 4.2: Cry detection system F_0 in function of MATLAB detect script F_0

From the collected data set, a maximum of frequency delta -25 Hz was observed, resulting in a minimum of 97.8 % accuracy or an error maximum of deviation 2.2 %.

The effectiveness of the infant cry classification was tested by the six types of signals, a sum of 26 samples. The samples used for evaluation included ten infant crying samples, three cooing samples, three babbling samples, three laughing samples, four adult

talking samples, and three pop-music samples. By introducing such variations in infant sound patterns and environmental noise conditions, the algorithm's ability to distinguish cry signals from other types of signals crying patterns could be learned.

The testing results yielded an 88.4% accuracy in the algorithm's ability to detect infant cries correctly. All ten crying samples were justly identified. The computed crying F_0 and the second dominant peak, F_2, detected in system algorithm PSD, were respectively identified in the range of 413 Hz to 531 Hz and 1053 Hz to 1341 Hz The extracted frequency range again validated the threshold setting and literature findings mentioned in section 3.2.3.3 and section 3.2.3.1.

However, one of the three laughing samples was mistakenly classified as crying as being exclamatory laughter. The laughter contained a F_0 between 380 Hz to 575 Hz and a first formant up to a maximum of 1020 Hz as mentioned in section 3.2.3.2.2, overlapping with the second formant range of cry signals.

Furthermore, all three cooing and babbling samples were correctly classified as non-cry signals, indicating the algorithm's ability to distinguish between these infant vocalizations. The cooing and babbling samples collected for evaluation again enforced the corresponding frequencies studied in section 3.2.3.2.1. For infants under a 12-months of age, the frequency components of the most pronounced syllabuses were found abundantly in the testing samples. The F_0s were 387 Hz, 602 Hz, and 1033 Hz. Similarly, three of four adult talking samples and two of three pop-music signals were correctly discerned as non-cry signals.

4.5 Audio streaming range

The range of the audio streaming was evaluated by moving the Parent's Module away from the Baby Module B until the audio was no longer intelligible. The distance between the modules was increased in steps of 50 cm in a closed environment. The audio streaming range was found to be unintelligible after 4.7 m.

5 DISCUSSION

5.1 Power consumption performance

As shown in 4.1, Baby Module A approximately consumes 1.8 W of power. As this unit should be able to function as a standalone unit close to the baby, an estimation of the battery capacity is desirable. To calculate the battery capacity, the following formula is used:

$$Q = \frac{P}{S} \cdot t \tag{5.1}$$

Where Q is the battery capacity in units of [Ah], P is the power consumption in units of [W], S is the power supply voltage in units of [V], and t is the battery life in hours.

For a 5 V supply, this means that in order to obtain a battery life of 30 days, a battery of at least 260 Ah is required.

5.2 Temperature measurement performance

The calculated values align closely with the measured values, indicating the NTC's ability to provide accurate temperature estimates. However, the tests showed that we failed to achieve the desired precision of 0.1°C as mentioned in section 1.1.

It should however be mentioned that significant fluctuations were observed on the multimeter during testing. Hence, the temperature measurement tests show a precision of only 0.5 °C. Therefore, better precision is expected if more precise measuring equipment was available.

5.3 Hall sensor performance

The Hall sensors can detect changing magnetic field of the N42 neodymium magnet until a distance of 4 cm. This means that detection of the prone position is only possible when the baby is not further away from the Hall sensor than 4 cm.

5.4 Cry segmentation performance

False positives occurred when non-target signals produced peaks in their frequency spectrum that overlapped with the range of crying signals, such as a burst of exclamatory infant laughter. Similarly, a female adult producing sounds with a sound pattern resembling an infant's cry would also be recognized as a cry signal. In the false positive female adult laughing case, the F_1 of the recorded audio frame was at 465.6 Hz, whereas the detected second peak was at 1481.8 Hz. This phenomenon was also observed when a female singer was singing the energized and high-frequency phonetics. These findings highlight the need for further algorithm refinement to reduce

false positives when encountering vocal characteristics resembling infant cries.

The current frequency precision comparison between the F_0 computed by the cry detection system and F_0 analyzed by the MATLAB script showed an expected deviation value within the frequency steps. The frequency step was computed as f_s over the extracted FFT frame length of 1024 (N). In spite of merely a maximum of 2.2 % error being found, such an error rate cannot be tolerable when the objectives are designated to be the analysis of the higher frequency components. It was also evident from the exclamatory laughing case that a minor difference in the F_0, F_1, or F_2 could have introduced an incorrect identification in a system.

In a future revision, we believe the accuracy of the cry detection system could be improved with three approaches. The first could be to adopt another controller offering a more significant heap memory. Due to heap capacity limitation, the current number of samples grouped for each FFT frame is limited to a size 43 times smaller than the f_s. Such a small size yields a large frequency step and a narrow-band FFT. Such a small number yields a large frequency step and a narrow-band FFT. A narrow-band FFT results in a poor frequency resolution and limits the system's ability to detect non-stationary signals given that the allowed time interval is small. However, the sample size was chosen to prevent triggering a heap capacity overflow error in the MCU. By switching to an MCU with greater heap capacity, a more extensive set of data samples could be segmented into a frame. Given a greater samples size, data encompassing a more sufficient amount of information can be analyzed at once, yielding a wide-band FFT. A wide-band FFT is, however, more suitable for a transient-signal monitoring system such as a cry detector. The frequency resolution accordingly can also decrease, mitigating spectral leakage.

The second approach would be to use an adaptive ZCR algorithm. An adaptive ZCR calculation periodically captures and computes the ZCR of surrounding signal, feedback to the system for threshold re-parsing and thus constantly deducting noise level from interested signals. An adaptive ZCR design would strengthen the performance of a cry detection system.

The last approach would be to train a Neural Network to detect baby cries. This would lead to a system that contains more detection features than we would be able to include in a hard-coded algorithm. The drawbacks of this solution are that Machine Learning would need high computation power and extensive heap memory capability.

5.5 Audio streaming range performance

It was observed that the audio streaming range reached its limit at a distance of 4.7 meters. Beyond this distance, the transmitted audio was no longer comprehensible. This finding implies that within a typical indoor setting, the Baby Radio Monitoring System (BRMS) is not capable of effectively transmitting audio signals between the two modules. One of the possible explanations is the radio interference caused by the Si24R1 RF module.

We believe the performance of the audio streaming could have been improved by using UDP packets to transmit the audio via a local network rather than using the ESPNOW protocol. Using UDP would have granted us larger packet sizes and lower latency. Moreover, it outsources the range issue to the coverage of the wifi network within a household.

6 CONCLUSION

The evaluation and validation of the Baby Radio Monitoring System (BRMS) showed that the system can effectively help parents prevent Sudden Infant Death Syndrome (SIDS), the primary cause of death among infants aged one month to one year. However, the BRMS' wireless audio streaming failed to achieve the desired range.

The thermistor-based temperature monitoring functionality of the BRMS proved to be effective in monitoring the temperature within the baby's room with a precision of up to 0.5°C. Although this precision exceeds the proposed requirement of 0.1°C, it remains sufficiently accurate to assess whether the room temperature is within a safe range for the baby.

The Hall sensor position monitoring feature was also proven to be effective in detecting the magnet on the baby's chest within a distance of 4 cm. The system almost immediately reacted to the prone position and sent an alert from Module A to the Parent's Module.

The cry detection algorithm was designed to have 80% accurate and it exhibited a remarkable accuracy of 88.4% in quiet environments. When a cry alert was triggered, immediate playback of a lullaby was activated to soothe the infant and facilitate timely caregiver response.

Finally, the audio streaming feature was the only requirement that was not met. We initially aimed for a range of 10-15 m but the audio transmitted from Module B and the Parents Module was only intelligible within 4.7 m. This demonstrated that the way we implemented our audio streaming via ESPNOW protocol was not effective.

Despite this drawback, the BRMS remains a cost-effective and reliable solution for baby monitoring. The temperature and position monitoring on the 8-bit STM8 show that simple microcontrollers can help counter SIDS. Moreover, the cry detection on the ESP32-WROOM with a small set of features such as F_0, F_2, and ZCR demonstrates that a robust algorithm for cry detection can be developed with little data acquisition. This technical report provides a useful foundation for future projects related to infant cry monitoring and SIDS prevention.

ACKNOWLEDGEMENTS

We would like to extend our appreciation to Yuri Cauwerts, Stijn Langendries, Gorik Stevens, Ruben Vanhoof, Gaël Carels, and Gil Vranken. The coaches of the Engineering Experience 3 (EE3) course at Groep T, who provided support and guidance throughout the course of this project.

LIST OF SYMBOLS

Symbols

Acronyms

ADC	Analogue-to-Digital Converter/Conversion
BRMS	Baby Radio Monitoring System
DAC	Digital-to-Analogue Converter/Conversion
DSP	Digital Signal Processing
ESP32	ESP32-WROOM-32
FFT	Fast Fourier Transform
FSM	Finite State Machine
PCB	Printed Circuit Board
PSD	Power Spectral Density
SIDS	Sudden Infant Death Syndrome
STM8	STM8S105K4
ZCR	Zero-Crossing Rate

BIBLIOGRAPHY

[1] K. Kochanek, S. Murphy, J. Xu, and E. Arias, "Deaths: Final data for 2017," *National Vital Statistics Reports*, vol. 68, no. 9, 2019. [Online]. Available: https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_09-508.pdf?deliveryName=USCDC_371-DM3857

[2] S. Jullien, "Sudden infant death syndrome prevention," *BMC Pediatrics*, vol. 21, no. S1, Sep. 2021, doi: https://doi.org/10.1186/s12887-021-02536-z.

[3] STMicroelectronics, "Stm8s105k4 - stmicroelectronics," STMicroelectronics website, Sep. 2015. [Online]. Available: https://www.st.com/en/microcontrollers-microprocessors/stm8s105k4.html

[4] E. Systems, "Esp32-wroom-32 datasheet," https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32_datasheet_en.pdf, Feb. 2023.

[5] Vishay, "Ntcle100e3," Vishay website, Mar. 2021. [Online]. Available: https://www.vishay.com/docs/29049/ntcle100.pdf

[6] BCD Semiconductors, "AH277A," Diodes Incorporated website, Oct. 2009. [Online]. Available: https://www.diodes.com/part/view/AH277A/

[7] N. Semiconductor, "nrf24l01 - single chip 2.4ghz transceiver," Mouser website, July 2007. [Online]. Available: https://www.mouser.com/datasheet/2/297/nRF24L01_Product_Specification_v2_0-9199.pdf

[8] M. Electronics, "Fermion mems microphone module," Mouser website, Jan. 2023. [Online]. Available: https://www.mouser.com/pdfDocs/Productoverview-DFRobot-SEN0487.pdf

[9] Transcend, "microsdhc card series." [Online]. Available: https://cdn-shop.adafruit.com/datasheets/TS16GUSDHC6.pdf

[10] Healthline. (2020, Feb.) What is the best room temperature for baby? Healthline. [Online]. Available: https://www.healthline.com/health/baby/room-temperature-for-baby#dressing-baby

[11] K. B. Van Dyke, R. Lieberman, A. Presacco, and S. Anderson, "Development of phase locking and frequency representation in the infant frequency-following response," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 9, p. 2740â2751, 2017.

[12] C. Ji, T. Mudiyanselage, Y. Gao, and et al., “A review of infant cry analysis and classification,” *J AUDIO SPEECH MUSIC PROC.*, August 2021. [Online]. Available: https://doi.org/10.1186/s13636-021-00197-5

[13] a. C. T. Lina Abou-Abbas and H. A. Fersaie, “A fully automated approach for baby cry signal segmentation and boundary detection of expiratory and inspiratory episodes,” *J Acoust Soc Am*, vol. 142, no. 3, pp. 1318–1331, Sep 2017. [Online]. Available: https://doi.org/10.1121/1.5001491

[14] nopnop2002, “mirf library for esp-idf,” 2021, accessed: April 16, 2023. [Online]. Available: https://github.com/nopnop2002/esp-idf-mirf

[15] H. Technology, “I2c serial interface 1602 lcd module,” Handson Technology. [Online]. Available: https://www.handsontec.com/dataspecs/module/I2C_1602_LCD.pdf

[16] D. Antliff, “esp32-i2c-lcd1602-example,” https://github.com/DavidAntliff/esp32-i2c-lcd1602-example, 2021.

[17] ——, “esp32-smbus,” https://github.com/DavidAntliff/esp32-smbus, 2021.

[18] ——, “esp32-i2c-lcd1602,” https://github.com/DavidAntliff/esp32-i2c-lcd1602, 2021.

APPENDICES

APPENDIX A : NTC SENSOR CIRCUIT

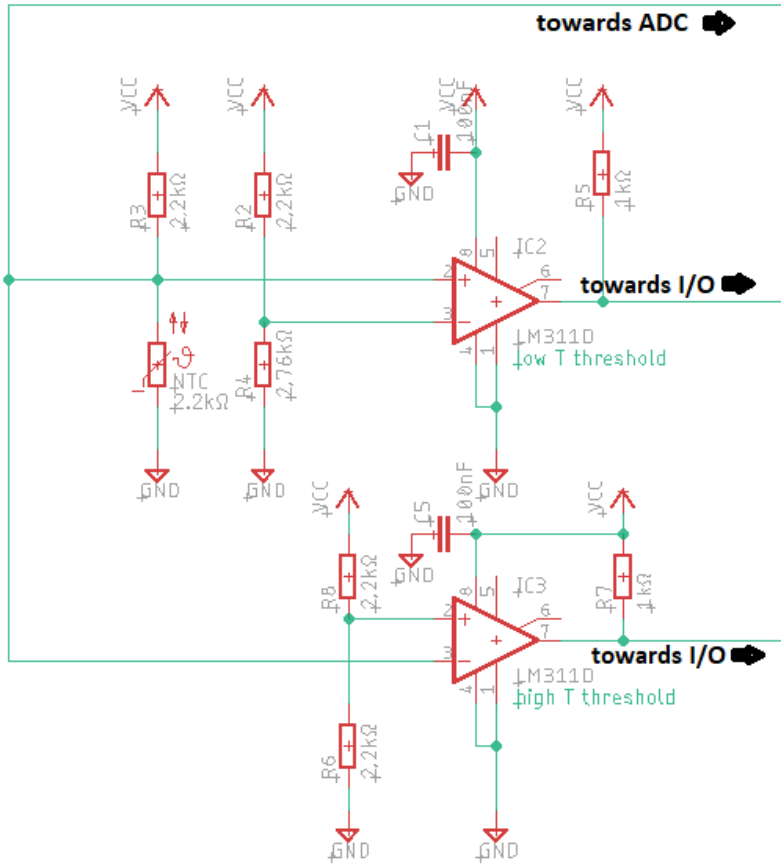

Figure A.1: NTC sensor circuit

APPENDIX B : HALL SENSOR 2D ARRAY

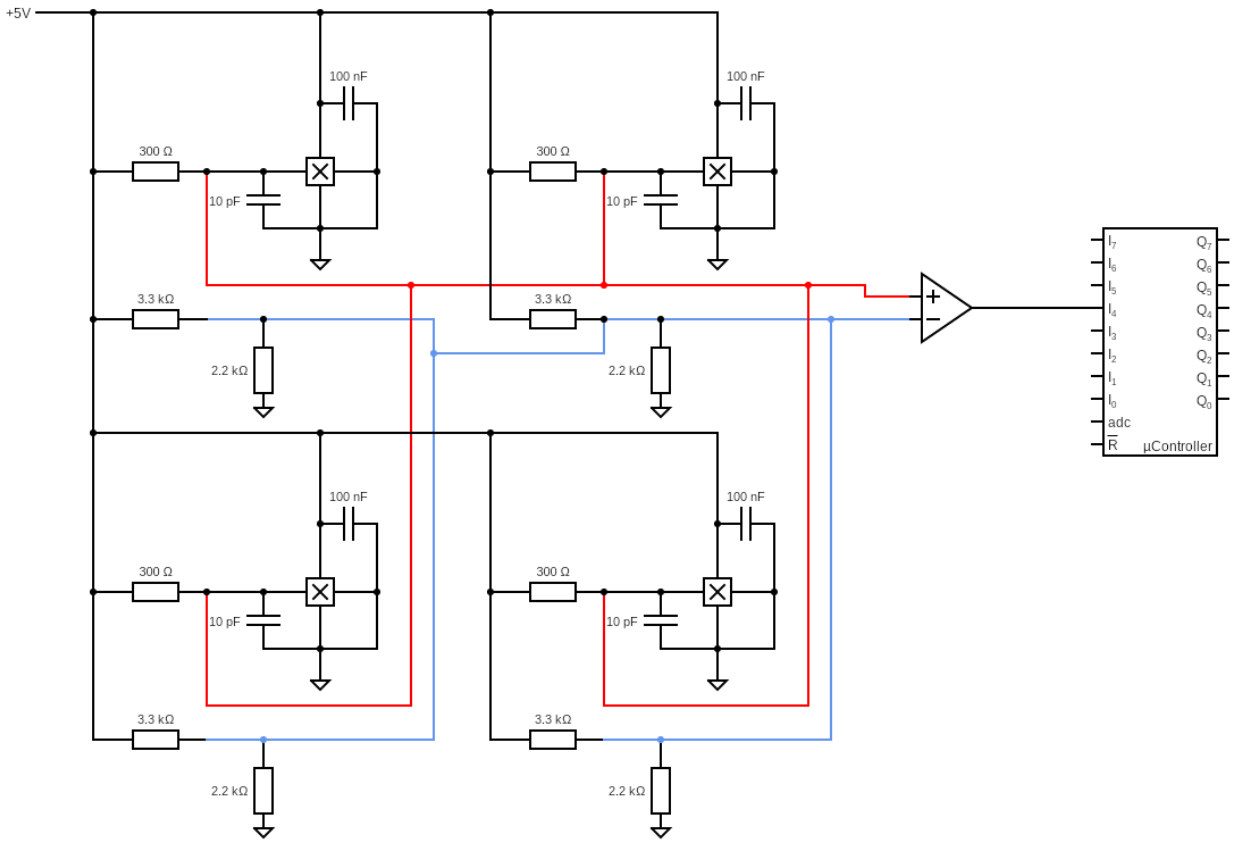

Figure B.1: Hall sensor 2D array

APPENDIX C : HALL SENSOR CIRCUIT

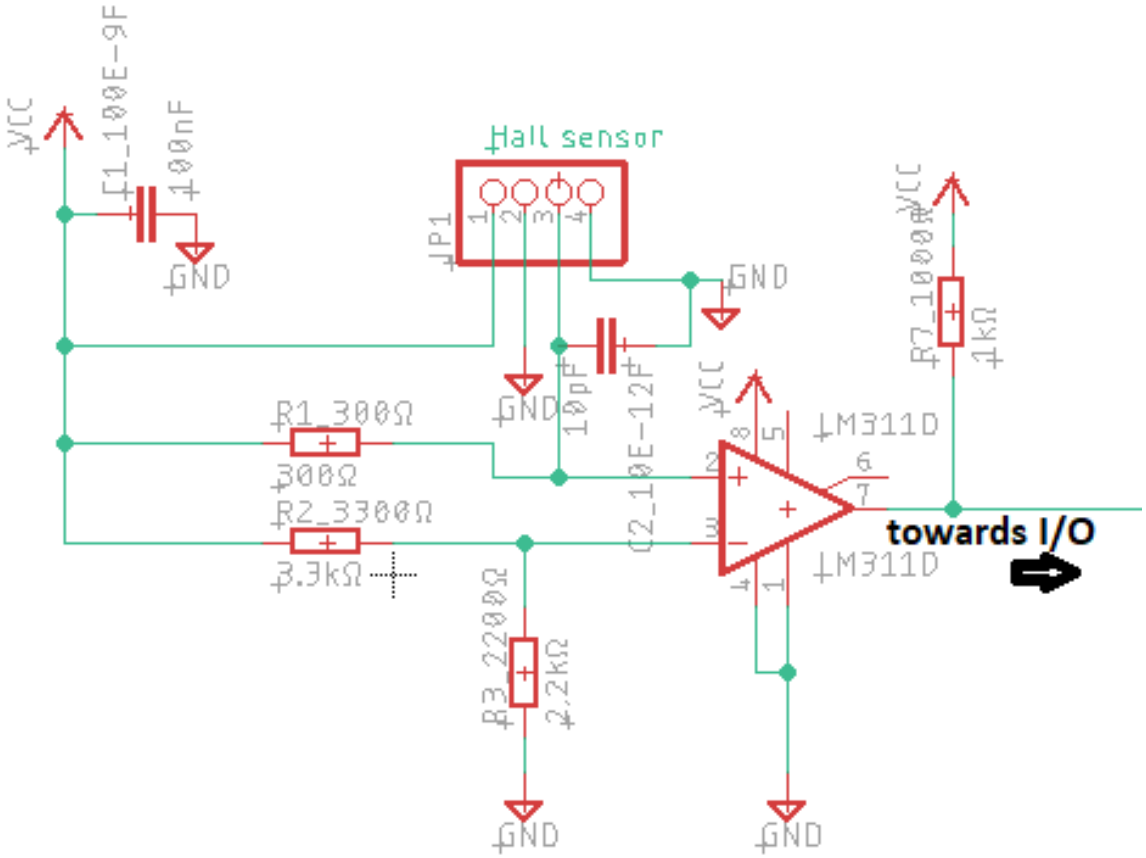

Figure C.1: Hall sensor circuit

APPENDIX D : NRF24

Chip Enable (CE)	PIN B0
Chip Select Not (CSN)	PIN E5
SCK	PIN C5 (SPI_SCK)
MOSI	PIN C6 (SPI_MOSI)
MISO	PIN C7 (SPI_MISO)

Table D.1: Pinout Mapping nRF24.

CONFIG	0x0A (TX), 0x0B (RX)
EN_AA	0x3F
EN_RXADDR	0x03
SETUP_AW	0x03
RF_CH	0x1C (Channel 28)
RF_SETUP	0x05 (1Mbps, -12dB)
STATUS	0x0E

Table D.2: Important Registers.

APPENDIX E : POWER CONSUMPTION

Hall Sensor			
Measurement Location	**Applied voltage (V)**	**Current (mA)**	**Supplied Power (mW)**
VCC to Hall pin 1	5	1,55	7,75
VCC to Hall pin 3, through 300 Ω		134	670
VCC to GND, through 3300 Ω and 2200 Ω		2,33	11,65
VCC to LM311 pin 8		0,7	3,5
VCC to LM311 pin 7		0,7	3,5
Power Consumption Hall circuit			696,4
NTC			
Measurement Location	**Applied voltage (V)**	**Current (mA)**	**Supplied Power (mW)**
VCC to GND through 2200 Ω and NTC	5	2,45	12,25
VCC to GND through 2200 Ω and 2760 Ω		1,78	8,9
VCC to LM311 pin 8		0,7	3,5
VCC to LM311 pin 7		125	625
VCC to GND through 2200 Ω and 2200 Ω		2,3	11,5
VCC to LM311 pin 8		0,7	3,5
VCC to LM311 pin 7		0	0
Power Consumption NTC circuit			664,65
nRF24			
Absolute Maximum Ratings			
	Minimum	**Maximum**	**Total Power Dissipation (mW)**
		60	60
STM8			
General Operating Conditions			
	Minimum	**Maximum**	**General Power Dissipation (mW)**
		360	360
Total Power Consumption			1781,05

Figure E.1: Power consumption of the Hall sensor circuit, NTC circuit, nRF24, and STM8

APPENDIX F : ACOUSTIC ANTI-ALIASING CIRCUIT

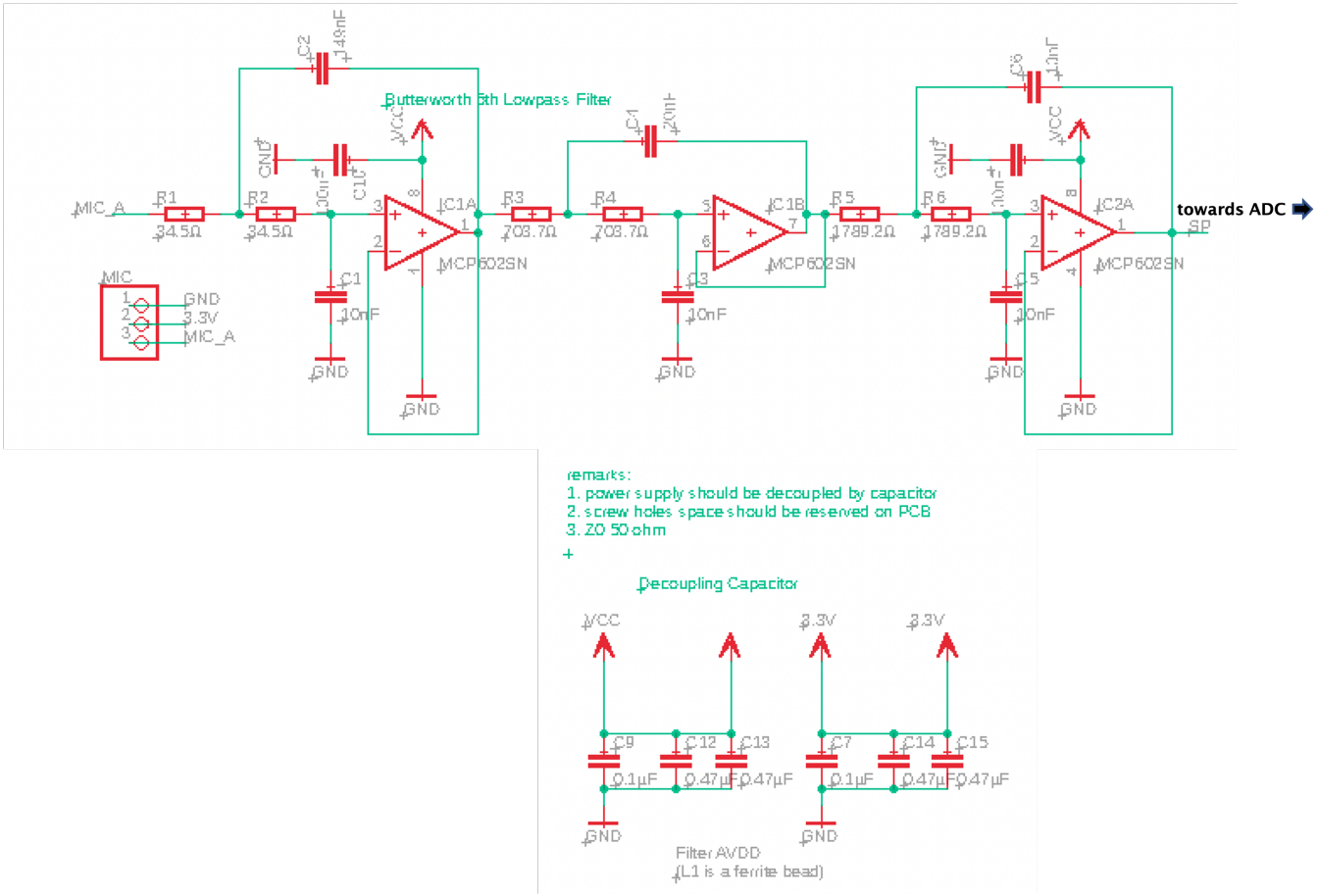

Figure F.1: Butterworth low-pass filter circuit

APPENDIX G : PCB BOARD DESIGNS

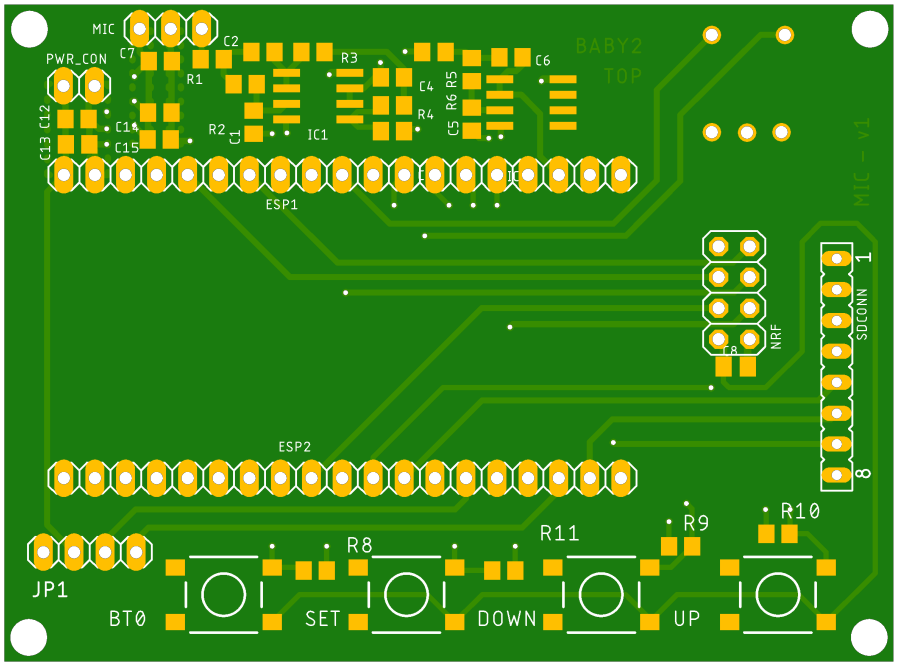

Figure G.1: PCB Parent's Module and Baby Module B

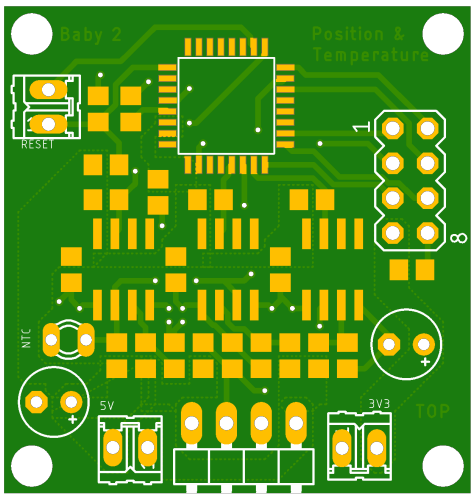

Figure G.2: PCB Baby Module A