

# Indice

<b>1</b>	<b>Introduzione</b>	<b>3</b>
1.1	Contesto e limiti delle onde elettromagnetiche .....	3
1.2	Motivazione dell'approccio acustico .....	4
1.3	Richiami teorici sulla propagazione acustica .....	4
1.4	Motivazioni sperimentali e obiettivi della tesi .....	5
1.5	Contributi attesi .....	6
<b>2</b>	<b>Stato dell'Arte</b>	<b>7</b>
2.1	Tecnologie IoT .....	7
2.2	Comunicazione in ambienti sotterranei .....	7
2.2.1	Radio frequenze .....	8
2.2.2	Onde acustiche .....	9
2.3	Applicazione della letteratura alla tesi .....	10
<b>3</b>	<b>Design del Protocollo</b>	<b>11</b>
3.1	Architettura generale del sistema .....	11
3.2	Livello Fisico .....	11
3.2.1	Ingresso .....	11
3.2.2	Uscita .....	17
3.3	Livello Bit .....	19
3.3.1	Decodifica .....	21
3.3.2	Codifica .....	21
3.4	Livello Link .....	22
3.4.1	Procedura di registrazione .....	22
3.4.2	Indirizzamento e instradamento .....	23
3.4.3	Gestione dei Token (TKN) .....	23
3.4.4	Comandi di controllo e di rete .....	25
3.4.5	Indirizzi di rete .....	25
3.5	Livello Applicazione .....	26
<b>4</b>	<b>Implementazione Hardware</b>	<b>28</b>

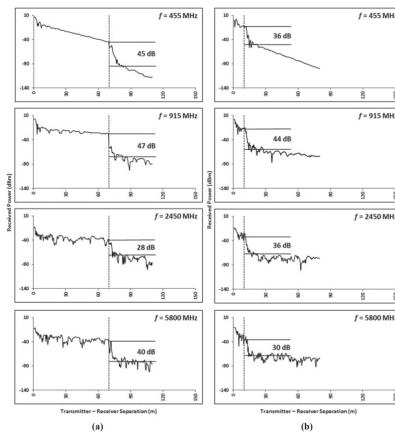
4.1	Tecnologie di base.....	28
4.1.1	Panoramica dell'ESP32 .....	28
4.1.2	Periferiche di Input/Output.....	30
4.2	Progettazione digitale .....	34
4.2.1	Collegamento del microfono digitale I <sup>2</sup> S.....	35
4.2.2	Collegamento dell'amplificatore digitale I <sup>2</sup> S .....	37
4.2.3	Sistema di segnalazione LED .....	38
4.2.4	Bottone di Hotspot .....	40
4.2.5	Circuito completo .....	40
4.2.6	Realizzazione del PCB .....	41
<b>5</b>	<b>Implementazione Software: Design del Protocollo</b>	<b>43</b>
5.1	Architettura generale del sistema .....	43
5.2	Emissione dei toni.....	44
5.3	Campionamento e bufferizzazione .....	45
5.4	Trasformata veloce di Fourier .....	47
5.5	Decodifica spettrale .....	48
5.6	Traduzione frequenza-bit .....	51
5.7	Ricevimento e trasmissione dei bit (Packers) .....	53
5.7.1	Ricezione dei bit .....	53
5.7.2	Trasmissione dei bit .....	54
5.8	Instrandamento e protocollo .....	55
5.9	Sensore di movimento .....	56
<b>6</b>	<b>Sperimentazione del Protocollo</b>	<b>57</b>
<b>7</b>	<b>Sperimentazione della propagazione del suono</b>	<b>60</b>
<b>8</b>	<b>Sviluppi futuri</b>	<b>63</b>
8.1	Evoluzione degli scenari applicativi .....	63
8.2	Revisione del formato di pacchetto.....	64
8.3	Sensoristica avanzata.....	65
8.4	Riduzione dei tempi di segnalazione .....	66
<b>9</b>	<b>Conclusioni</b>	<b>67</b>
<b>A</b>	<b>Repository del codice e risorse multimediali</b>	<b>75</b>

# 1. Introduzione

## 1.1 Contesto e limiti delle onde elettromagnetiche

Le reti di sensori in ambienti sotterranei pongono sfide peculiari alla comunicazione wireless. La propagazione delle *onde elettromagnetiche* nel sottosuolo soffre di attenuazioni elevate dovute alla permittività, alla conducibilità del terreno e all'umidità, con conseguente riduzione drastica della portata e dell'affidabilità dei link radio [Lee, 2000, pp. 24–59].

Alcune soluzioni proposte in letteratura includono l'uso di frequenze super-high ed ultra-high (SHF, UHF) con l'obiettivo di implementare sistemi di tracking e monitoraggio in miniere di carbone, gallerie o condotti [Jacksha and Zhou, 2016]. Tuttavia, queste frequenze a causa della loro natura fisica sono soggette a forti perdite di segnale e riflessioni multipath, limitando la copertura a range dai 10 ai 33 metri in condizioni ottimali con l'ausilio di antenne direzionali con un'altezza pari a 1.2 metri l'una; questa tecnologia inoltre mostra tutta la sua vulnerabilità in presenza di curve strette ( $90^\circ$ ) o ostacoli.



**Figura 1.1.** Corner Loss  $\sim 30$  dB in un condotto con angolo di  $90^\circ$  per radio frequenze HF/SHF [Jacksha and Zhou, 2016].

## 1.2 Motivazione dell'approccio acustico

Da questa problematica nasce la presente tesi, che si pone l'obiettivo di progettare e validare un **protocollo di comunicazione acustica** per reti **master–slave** in ambienti sotterranei, che dovrà essere: **efficiente, robusta ed economica**.

L'idea è sfruttare la *propagazione del suono nell'aria* presente in cavità, condotti o tunnel, trattando l'aria come un canale guida all'interno degli spazi confinati, e più in generale impiegare l'onda acustica come mezzo portante laddove il canale elettromagnetico è troppo penalizzato.

## 1.3 Richiami teorici sulla propagazione acustica

In effetti, l'aria può essere modellata come un fluido compressibile: le variazioni locali di pressione e densità generate da una sorgente si propagano come onde longitudinali, la cui dinamica è descritta dall'equazione delle onde acustiche. La velocità di propagazione, che in condizioni standard è circa 343 ms, dipende da temperatura, pressione e composizione del gas, secondo la relazione  $c = \sqrt{\gamma p_0 \rho_0}$ .

Inoltre «Schroeder ha definito l'omonima frequenza limite  $F_{sch}$ , diversa per ogni singolo ambiente, che distingue la regione a bassa frequenza da quella ad alta.»[Bianchi, 2023, pp. 1], pertanto il calcolo di tale frequenza risulta fondamentale per far sì che le frequenze utilizzate siano più alte della frequenza di Schroeder, in modo da poter sfruttare al meglio le proprietà statiche dell'ambiente. Questa visione permette di trattare l'aria non solo come “spazio vuoto”, ma come un vero e proprio *canale fisico*, caratterizzato da attenuazioni dovute a dispersione geometrica, assorbimento atmosferico, la letteratura non considera, quindi, il suono come un'onda, bensì come un **raggio sonoro**, in grado di riflettersi su una superficie [Bianchi, 2023, pp. 3–4].

## Frequenza di Schroeder nei tunnel minerari

Per delimitare il regime modale da quello statistico si adotta la frequenza di Schroeder [Bianchi, 2023, p. 1], stimabile come

$$f_s \approx 2000 \sqrt{\frac{T_{60}}{V}},$$

con  $T_{60}$  in secondi e  $V$  in  $\text{m}^3$ .

In gallerie o tunnel con superfici rigide si riscontrano comunemente tempi di riverbero  $T_{60}$  dell'ordine di 3 s. Considerando un tratto rappresentativo di galleria di sezione  $5 \times 5 \text{ m}$  e lunghezza 100 m, il volume è

$$V \simeq 2500 \text{ m}^3.$$

Ne segue:

$$f_s \approx 2000 \sqrt{\frac{3}{2500}} = 2000 \sqrt{0.0012} \approx 2000 \times 0.0346 \approx 69 \text{ Hz}.$$

Per sezioni più piccole (ad esempio  $4 \times 4 \text{ m}$  con lunghezza 50 m,  $V \simeq 800 \text{ m}^3$ ) si ottiene

$$f_s \approx 2000 \sqrt{\frac{3}{800}} \approx 122 \text{ Hz}.$$

In sintesi, nei tunnel minerari la frequenza di Schroeder è tra  $\sim 70$  e  $\sim 120$  Hz; il funzionamento di una rete che opera tra 1–9 kHz risulta, quindi, al di sopra della soglia, riducendo la sensibilità a risonanze modali locali e rendendo più uniforme la propagazione; facendo avvenire il fenomeno del **raggio sonoro** di cui si è parlato in precedenza.

## 1.4 Motivazioni sperimentali e obiettivi della tesi

Il fine ultimo della rete sarà quello di permettere lo scambio di dati tra nodi utilizzando frequenze sonore appartenenti al range 1–9 kHz, la cui applicazione principale è destinata al monitoraggio post-disastro, considerando che in altri contesti l'utilizzo di segnali acustici potrebbe risultare disturbante. I sensori saranno, inoltre, in grado di rilevare la presenza di corpi umani o gas tossici a seguito di crolli o esplosioni, e trasmettere queste informazioni a un nodo master situato in superficie o in una zona sicura.

La letteratura recente indica che, in scenari confinati, segnali acustici a media frequenza possono mantenere un rapporto segnale/rumore utilizzabile su distanze dell'ordine delle

decine di metri, con modalità di propagazione *lungo condotti* (ad esempio tubazioni o gallerie) e con modelli di attenuazione prevedibili [Al Moshi et al., 2024].

Queste evidenze motivano la definizione di un protocollo leggero e robusto (rilevazione spettrale, soglie adattive con ausilio di machine learning) che faccia uso di componenti economici e facilmente integrabili (microfoni/codec, amplificatori, altoparlanti) per creare un *layer fisico* acustico e il relativo *protocollo di accesso*.

## 1.5 Contributi attesi

In sintesi, i contributi attesi sono:

1. modellazione e scelta dei parametri del canale acustico in aria in ambienti confinati;
2. progettazione di un protocollo master–slave basato su pattern di frequenze, ACK e gestione del ritardo casuale;
3. implementazione hardware/software a basso costo;
4. validazione sperimentale mediante misure di SPL a diverse distanze, con stima dei livelli attesi per amplificazioni maggiori tramite traslazione dei valori rilevati.

## 2. Stato dell'Arte

### 2.1 Tecnologie IoT

L'Internet of Things (IoT) rappresenta una rete di dispositivi fisici, veicoli incorporati con elettronica, software, sensori e connettività di rete. Questi dispositivi sono progettati per raccogliere e scambiare dati, consentendo di comunicare con l'ambiente circostante. Negli anni diverse tecnologie sono state usate per far comunicare questi dispositivi, la maggior parte di queste si basano su radio frequenze, come LoRa, NB-IoT, ZigBee, Wi-Fi, BLE, fortemente indirizzate ad un uso outdoor o indoor.

Tuttavia, esistono scenari in cui la propagazione elettromagnetica incontra limiti strutturali, ad esempio in ambienti sotterranei, gallerie, miniere o condotti, dove l'attenuazione del segnale cresce a causa della composizione del terreno e dell'umidità. In questi contesti, l'affidabilità della comunicazione wireless tradizionale risulta fortemente compromessa.

### 2.2 Comunicazione in ambienti sotterranei

La comunicazione in ambienti sotterranei rappresenta una sfida significativa a causa delle caratteristiche uniche di questi ambienti. Le onde elettromagnetiche, comunemente utilizzate per la comunicazione wireless, soffrono di attenuazioni elevate dovute alla permittività del terreno e all'umidità, con una conseguente riduzione della portata e dell'affidabilità dei ponti radio [Akyildiz and Stuntebeck, 2006, sec. 2].

In letteratura è possibile individuare diverse soluzioni proposte per affrontare queste sfide, tra cui si riconoscono due linee principali di ricerca:

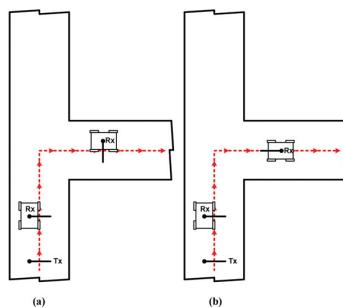
- l'uso di radio frequenze molto basse (VLF), medie (MF) o più alte (UHF, SHF), adattate a condizioni specifiche;
- l'uso di onde acustiche, che sfruttano la propagazione meccanica del suono attraverso solidi e fluidi.

## 2.2.1 Radio frequenze

Le ricerche sulle onde elettromagnetiche in gallerie e miniere hanno mostrato limiti significativi: la propagazione è fortemente influenzata dalla geometria degli ambienti e dalle proprietà elettriche del mezzo. Le ricerche effettuate sulle onde elettromagnetiche in gallerie e miniere mostrano limiti significativi riguardanti la propagazione, che risulta fortemente influenzata dalla geometria degli ambienti e dalle proprietà elettriche del mezzo.

Gli studi del U.S. National Institute for Occupational Safety and Health's (NIOSH) Office of Mine Safety and Health Research (OMSHR) [Jacksha and Zhou, 2016], avvenuti in seguito a crolli di alcune miniere di carbone in U.S. nel 2006 hanno piantato le fondamenta per l'uso di onde elettromagnetiche SHF/UHF in ambienti sotterranei.

Attraverso una serie di antenne direzionali dalla misura di 1.2 metri e ricevitori collocati prima in linea retta ad una distanza crescente, e poi in presenza di curve a 30°, 60°, 90° e 180° in condotti sotterranei hanno dimostrato che, in condizioni ottimali, la copertura può raggiungere i 33 metri, ma la presenza di curve strette, biforcazioni o ostacoli riduce drasticamente la portata [Jacksha and Zhou, 2016, sec. 3]; inoltre lo stesso fenomeno è stato osservato anche negli studi effettuati da Jeho Lee [Lee, 2000, p. 59].



**Figura 2.1.** Mappa Tunnel U.S. National Institute for Occupational Safety and Health's[Jacksha and Zhou, 2016, sec. 3].

Mentre [Akyildiz and Stuntebeck, 2006, sec. 2] descrivono le difficoltà intrinseche alla propagazione sotterranea a causa di permittività e conducibilità elevate del terreno.

Più in generale, le tecnologie wireless pensate per ambienti sotterranei devono fare i conti con una ridotta penetrazione delle frequenze tradizionalmente usate.

Per questo, l'uso di frequenze estremamente basse (VLF, tra 3–30 kHz) è stato studiato in scenari di emergenza e applicazioni militari, ma presenta limiti di banda e di miniaturizzazione delle antenne [Salam et al., 2023, sec. 5.1].

## 2.2.2 Onde acustiche

Un’alternativa promettente alla comunicazione elettromagnetica in ambienti sotterranei è l’impiego di *onde acustiche* (meccaniche), che si propagano attraverso solidi e fluidi (aria, acqua, fanghi di perforazione). A differenza delle onde EM, la propagazione acustica è spesso meno sensibile alla conducibilità elettrica del mezzo e può seguire percorsi guidati (ad es. lungo tubazioni o gallerie), con migliore resilienza a curve e biforcati [Fishta et al., 2023, Heifetz and et al., 2017, Farai et al., 2023].

**Modello di canale (cenni).** Nei **solidi** (roccia, calcestruzzo, metalli, polimeri), la propagazione è dominata da onde elastiche (longitudinali e di taglio) e, in geometrie guidate (tubi, piastre, travi), da *guided waves* (p. es. onde di Lamb o creep waves) con attenuazione che cresce fortemente con la frequenza e con il disaccoppiamento modale [Bianchi, 2023, p. 2].

Nei **fluidi** in condotti (acqua/aria), il canale è tipicamente a bassa frequenza (centinaia di Hz–pochi kHz), con riflessioni multiple e dispersione; in vasche o condotti lunghi si osservano fenomeni di riverbero e frequenze di taglio dei modi guida [Fishta et al., 2023].

**Hardware e trasduttori.** Sono impiegati trasduttori piezoelettrici o magnetostrettivi accoppiati meccanicamente al mezzo (collari su tubi, inserti su rocce/pareti, piastre di accoppiamento).

Nei tubi metallici, coppie di attuator/ricevitori permettono una comunicazione non invasiva (senza rompere il tubo), sfruttando onde di taglio o di Lamb [Heifetz and et al., 2017, sec. 4.1].

Inoltre, per la comunicazione sotterranea si usano sorgenti e ricevitori acustici per mandare segnali nel terreno, fino a circa 50 m ma con velocità molto bassa ( 20 bps) [Yang et al., 2020, sec. 2.1].

**Tecniche di modulazione e codifica.** Dato il canale fortemente dispersivo e soggetto a multi-percorso, si adottano modulazioni quali: FSK/MFSK e BFSK a bassa frequenza creando collegamenti bassa velocità [Fishta et al., 2023, Farai et al., 2023].

**Prestazioni riportate in letteratura.** Risultati sperimentali rappresentativi includono:

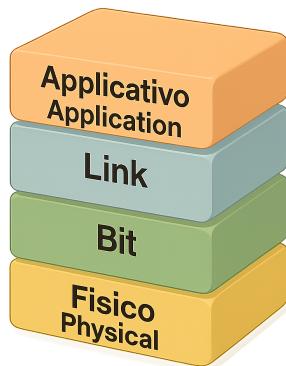
- **Suolo (through-soil).** Collegamenti fino a  $\sim$ 50 m con  $\sim$ 20 bps utilizzando portanti a bassa frequenza e trasduttori compatti; dimostrazione di comunicazione digitale robusta in campi prova agricoli [Yang et al., 2020, sec. 3].
- **Tubi metallici (acciaio).** Trasmissione di dati via onde elastiche lungo tubazioni esistenti in impianti industriali; dimostrate immagini e pacchetti a centinaia di bps–pochi kbps su decine–centinaia di metri a bassa potenza, sfruttando onde di taglio e di Lamb [Heifetz and et al., 2017].
- **Condotti idrici reali.** Revisione dei test in reti idriche urbane indica fattibilità di reti acustiche in-pipe per IoT, con vincoli di potenza, sincronizzazione e instradamento; lo stato dell’arte è pre-commerciale ma in rapida evoluzione [Fishta et al., 2023].

## 2.3 Applicazione della letteratura alla tesi

La tesi si occuperà quindi della creazione di un protocollo di comunicazione acustica per sistemi distribuiti in ambienti sotterranei, basandosi sui risultati e le tecniche di cui sopra, è possibile quindi dedurre alcuni requisiti chiave: Il protocollo opera nella banda 1–9 kHz, questa scelta è coerente con la letteratura, come visto nel Capitolo 1, inoltre l’utilizzo di una codifica FSK permetterà di avere un sistema robusto ma lento, coerentemente con le velocità di trasmissione viste in letteratura [Fishta et al., 2023].

# 3. Design del Protocollo

## 3.1 Architettura generale del sistema



**Figura 3.1.** Livelli del Protocollo

Il protocollo è strutturato in quattro livelli principali, ognuno con una funzione specifica.

## 3.2 Livello Fisico

Il livello fisico è responsabile della trasmissione dei bit precedentemente composti dal livello Bit. Questo sfrutta uno speaker per la trasmissione e un microfono per la ricezione dei bit che vengono trasmessi mediante coppie superimposte di frequenze.

### 3.2.1 Ingresso

Il microfono I2S (descritto nel Capitolo 4) è collegato al microcontrollore ESP32 tramite il bus I2S, sfruttando uno dei due canali disponibili. Il segnale audio viene campionato a 48 kHz con una risoluzione di 16 bit, per poi essere gestito mediante due array in swapping, concetto che verrà approfondito successivamente.

La scelta della frequenza di campionamento deriva dal **teorema di Nyquist-Shannon** [Shannon, 1949, pp.448–449], secondo il quale un segnale può essere ricostruito senza ambiguità se la frequenza di campionamento  $f_s$  è almeno doppia rispetto alla massima frequenza del segnale  $f_{\max}$ :

$$f_s \geq 2f_{\max}.$$

Ne consegue che la massima frequenza rappresentabile è

$$f_{\text{Nyquist}} = \frac{f_s}{2}.$$

Nel nostro caso:

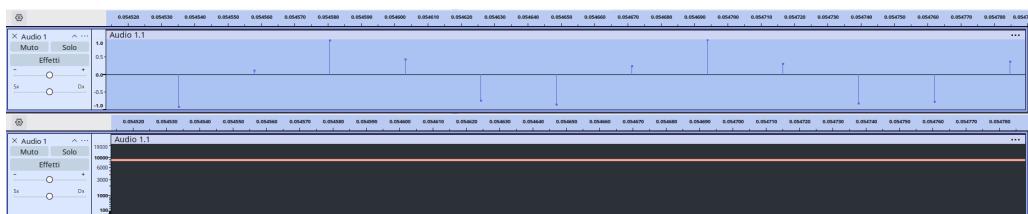
$$f_s = 48 \text{ kHz} \quad \Rightarrow \quad f_{\text{Nyquist}} = 24 \text{ kHz}.$$

Poiché l'orecchio umano percepisce frequenze fino a circa 20kHz [Zwicker and Fastl, 1999], la scelta di 48kHz garantisce la copertura dell'intero spettro udibile, con un margine di sicurezza di 4kHz. Frequenze prossime al limite teorico di Nyquist risulterebbero invece difficili da catturare senza aliasing, a causa dei limiti pratici dei filtri anti-alias.

Per l'elaborazione, i campioni vengono raccolti in blocchi di lunghezza  $N = 512$ . Con la frequenza di campionamento fissata:

$$T_{\text{blocco}} = \frac{N}{f_s} = \frac{512}{48 \cdot 10^3} \approx 0.01066 \text{ s}.$$

Ogni blocco di dati rappresenta quindi un intervallo di circa 10.7ms di segnale audio. Questa durata, successivamente a una attenta analisi effettuata sul software Audacity, si è rivelata sufficiente per poter catturare in modo affidabile i toni.



**Figura 3.2.** Verifica della cattura del tono a 9kHz con campionamento a 48kHz nel periodo 0-0.010 secondi, utilizzando finestra di Hann su 512 elementi

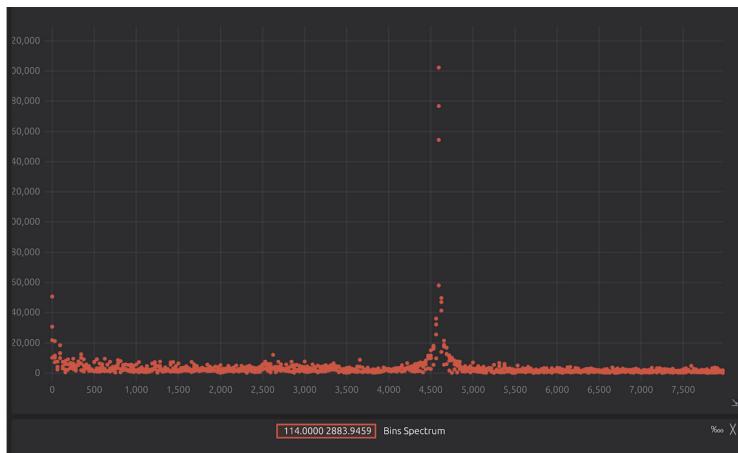
Una volta acquisito il blocco, viene calcolata la **trasformata veloce di Fourier (FFT)** [Cooley and Tukey, 1965, pp. 297–301] per analizzare lo spettro in frequenza. La complessità della FFT cresce come  $N \log_2 N$ , e per  $N = 512$  si hanno circa:

$$512 \cdot \log_2 512 = 512 \cdot 9 = 4608$$

operazioni complesse.

Sul microcontrollore ESP32 dual-core operante a 240MHz questo si traduce in un tempo di elaborazione di circa 0.21 ms per blocco, ovvero  $\sim 0.415 \mu\text{s}$  per campione. In termini di carico computazionale:  $\sim 20$  moltiplicazioni,  $\sim 29$  a ddizioni/sottrazioni e un’operazione di radice quadrata per campione, pari a  $\sim 99$  cicli di clock [esp, 2020, p. 30].

Questo significa che, a fronte di una finestra temporale di 10.7 ms, l’FFT viene calcolata in meno del 3% del tempo disponibile, consentendo un’elaborazione in tempo reale anche senza multi-threading.



**Figura 3.3.** Grafico contenente lo spettro di frequenza calcolato tramite FFT su un blocco di 512 campioni acquisiti a 48kHz.

**FFT, bin e asse delle frequenze** L’algoritmo FFT non calcola direttamente lo spettro in frequenza continua, bensì restituisce  $N$  valori discreti  $X_k$  detti *bin*, con indici

$$k = 0, 1, \dots, N - 1.$$

Questi bin rappresentano i coefficienti della trasformata discreta e contengono l’informazione spettrale su griglie di frequenza equispaziate.

Per  $N = 512$  e frequenza di campionamento  $F_s = 48$  kHz, ciascun bin corrisponde alla frequenza

$$f_k = k \frac{F_s}{N}, \quad \Delta f = \frac{F_s}{N} = \frac{48000}{512} = 93.75 \text{ Hz.}$$

Dunque l'asse  $x$  del grafico, che mostra le frequenze in Hz, è in realtà una *ridenominazione* dei bin calcolati dall'FFT. Ogni valore sull'asse delle ascisse non è una misura continua, ma la mappatura del bin discreto  $k$  nella corrispondente frequenza  $f_k$ .

**Simmetria dello spettro** Poiché il segnale è reale, vale la simmetria

$$XN - k = \overline{Xk} \quad \Rightarrow \quad |XN - k| = |Xk|.$$

Per questo motivo lo spettro viene spesso mostrato soltanto fino al bin  $N/2$ , ossia da  $k = 0$  a  $k = 256$ . In particolare, il bin  $k = 256$  corrisponde alla frequenza massima non ambigua, cioè la frequenza di Nyquist:

$$f_{256} = 256 \cdot \frac{48000}{512} = 24 \text{ kHz.}$$

Durante la creazione di questo protocollo sono emerse diverse difficoltà legate alla presenza di rumore ambientale, questa condizione ha reso necessario l'uso di tecniche di filtraggio.

### 3.2.1.1 Filtraggio e riconoscimento dei toni

Una parte cruciale del processo di design del protocollo è stata la scelta delle frequenze da utilizzare per la trasmissione dei dati. I picchi che vengono selezionati dalla fase di filtraggio devono appartenere a un insieme di frequenze predefinite, così da poter essere riconosciuti dal **Livello Bit**.

La seguente tabella riporta l'insieme delle frequenze adottate, distinguendo quelle di tipo *carrier* da quelle destinate al trasferimento dei dati:

Frequenza [Hz]	Tipo
1000	Carrier
1400	Data
1800	Data
2200	Data
2600	Data
3000	Data
3400	Data
3800	Data
4200	Data
4600	Data
5000	Data
Frequenza [Hz]	Tipo
5400	Data
5800	Data
6200	Data
6600	Data
7000	Data
7400	Data
7800	Data
8200	Data
8600	Carrier
9000	Carrier

**Tabella 3.1.** Frequenze e tipi di segnale

Il filtraggio avviene nel dominio della frequenza, dopo l'applicazione della trasformata FFT a 512 punti. A differenza dei filtri digitali convenzionali (IIR/FIR), la selezione dei toni non si basa su maschere statiche ma su un insieme di procedure che rendono il sistema adattivo al rumore e preciso nell'identificazione dei picchi. Inizialmente, quando il protocollo era ancora nelle sue fasi embrionali, era stato deciso di utilizzare una semplice soglia fissa (filtro passa-basso) per discriminare i picchi delle frequenze predefinite dal rumore. Questa soluzione, tuttavia, si è rivelata inefficace: i microfoni presentano una risposta in frequenza che degrada sulle alte frequenze, rendendo i livelli in dB di queste ultime attenuati al punto da non superare la soglia prefissata. Inoltre, il rumore ambientale non è mai costante ma varia nel tempo e nello spettro, il che rendeva difficile definire un limite statico in grado di funzionare in tutte le condizioni.

In un'evoluzione successiva si è quindi ipotizzato l'impiego di soglie fisse ma differenziate per ciascuna frequenza, così da compensare la risposta non piatta del microfono.

Per calcolare tali soglie è stato utilizzato un algoritmo di regressione lineare, in grado di stimare la retta che meglio approssima l'andamento della soglia minima:

$$y = \beta \beta x \varepsilon \rightarrow y = -301.751324 \times x + 48531.689491$$

dove  $x$  rappresenta il numero del bin. La Figura 3.4 mostra lo spettro in frequenza di tre toni (basso, medio e alto) e mette in evidenza l'attenuazione introdotta dalla risposta del microfono.

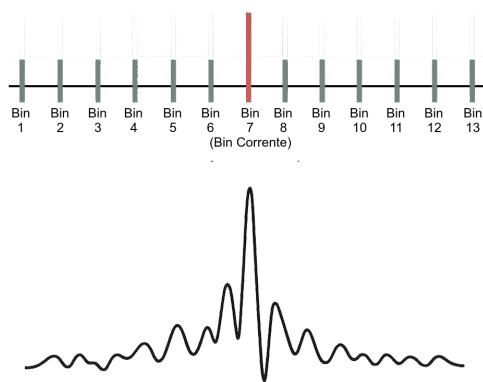


**Figura 3.4.** Grafico contenente lo spettro di frequenza di tre frequenze (bassa, media, alta), utilizzato per identificare l'attenuazione in frequenza del microfono.

Il sistema attuale ha superato queste limitazioni introducendo una catena di elaborazione più robusta. In primo luogo viene stimato il livello medio di rumore calcolando la media delle ampiezze spettrali:

$$\text{noise\_floor} = \frac{1}{N} \sum_{k=0}^{N-1} |X_k|$$

dove  $X_k$  è il modulo del  $k$ -esimo bin. A partire da questa misura viene definita una soglia dinamica, proporzionale al rumore, che permette di adattarsi alle condizioni del segnale: un picco viene considerato valido solo se la sua ampiezza supera di almeno otto volte il livello medio di rumore. Per ridurre ulteriormente i falsi positivi, ogni frequenza candidata deve corrispondere a un **massimo locale**, ossia il **bin identificato** deve avere ampiezza superiore rispetto ai sei **bin adiacenti a sinistra e a destra**.

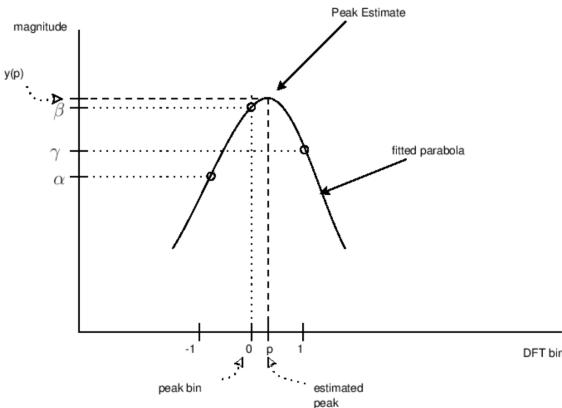


**Figura 3.5.** Picco Locale nell'intorno dei sei Bins

**Interpolazione parabolica** A questo punto, per migliorare la risoluzione in frequenza oltre i limiti del singolo bin FFT, viene applicata un'**interpolazione parabolica** basata sui valori  $X_{k-1}$ ,  $X_k$  e  $X_{k+1}$ , secondo la formula

$$p = \frac{1}{2} \frac{\alpha - \gamma}{\alpha - 2\beta},$$

dove  $\alpha = |X_{k-1}|$ ,  $\beta = |X_k|$  e  $\gamma = |X_{k+1}|$ . La frequenza stimata diventa così  $f_k + p\Delta f$ , con  $\Delta f = 93.75 \text{ Hz}$ , ottenendo una risoluzione sub-bin.



**Figura 3.6.** Interpolazione Parabolica [Smith and Serra, 2020]

È inoltre disponibile, sebbene disattivato di default, un modulo di regressione lineare che consente di compensare eventuali pendenze dello spettro in condizioni particolarmente critiche. Infine, i **picchi confermati vengono incapsulati in una struttura dati (`struct_tone_frequencies`) e inoltrati al *Livello Bit***, che li utilizza per ricostruire i bit associati ai canali *master*, *slave* e *config*.

### 3.2.2 Uscita

Il livello fisico si occupa anche della trasmissione dei dati, convertendo coppie di frequenze ricevute dal Livello Bit in segnali audio. Questa operazione viene eseguita mediante la sintesi digitale di due toni sinusoidali, che vengono sommati, sistemati in fase, sistemati in ampiezza ed infine inviati attraverso il bus I2S ad un amplificatore di potenza (descritto nel Capitolo 4).

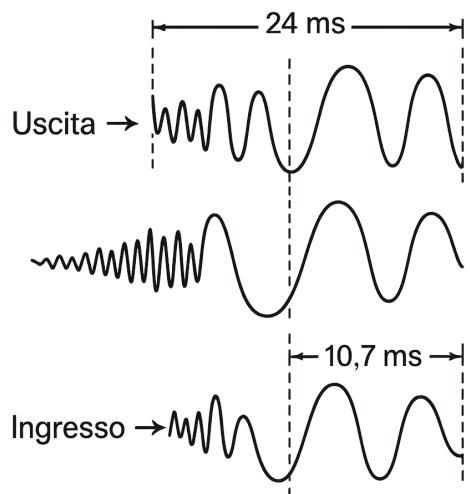
**Sintesi dei toni** L'emissione di una coppia di frequenze superimposte è un'operazione che coinvolge l'utilizzo di diverse tecniche di sintesi digitale.

In primo luogo viene calcolato il numero necessario di campioni che compongono la **sinusoide**, per far ciò è necessario utilizzare la stessa frequenza di campionamento adottata per l'acquisizione, ovvero 48 kHz, ed la durata d'emissione che è stata fissata a 0.024 s. La scelta di questa durata deriva dalla finestra temporale di 10.7 ms utilizzata per l'analisi FFT, quindi 24 ms è un compromesso che consente di avere un segnale sufficientemente lungo per essere percepito e demodulato, evitando così che la **finestra di ascolto (FFT)** possa essere in mezzo tra un tono e l'altro, rendendo così difficile la demodulazione. Con questa durata, quindi, si avrà sempre una serie di 512 campioni in ingresso che conterranno, sicuramente, i toni.

Il numero di campioni necessari per la sintesi è quindi

$$N = f_s \cdot T \quad (3.1)$$

$$N = 48\,000 \text{ Hz} \cdot 0.024 \text{ s} = 1152 \text{ campioni} \quad (3.2)$$



**Figura 3.7.** Finestra d'ascolto della FFT sul segnale emesso

A seguito di diversi test, dove è stato riscontrato un "fruscio" tra l'emissione di un tono e l'altro, è stato implementato un sistema di allineamento di fase.

Questo funziona attraverso una variabile che tiene traccia della fase dell'ultimo campione emesso, in questo modo il primo campione del nuovo tono sarà sempre in fase con l'ultimo campione del tono precedente. Ciò consente di evitare discontinuità nel segnale che si traducono in rumore udibile.

La sintesi di ciascuna sinusoide avviene quindi secondo la formula

$$x_n = A \cdot \frac{\sin\phi_1 \sin\phi_2}{2} \quad (3.3)$$

formula semplificata dove  $A$  è l'ampiezza del segnale,  $\phi_1$  e  $\phi_2$  sono le fasi delle due sinusoidi calcolate come

$$\phi = \frac{2\pi f}{f_s} \quad (3.4)$$

### 3.3 Livello Bit

Il Livello Bit si occupa di convertire le frequenze ricevute dal Livello Fisico in bit, e viceversa. Il pacchetto con cui il livello Bit comunica con il Livello Fisico è la seguente struttura dati

Master			Slave			Config		
Frequenza 1	Frequenza 2	Frequenza 3	Frequenza 1	Frequenza 2	Frequenza 3	Frequenza 1	Frequenza 2	Frequenza 3

**Tabella 3.2.** Struttura di comunicazione tra Livello Bit e Livello Fisico

Ogni gruppo (master, slave, config) rappresenta un gruppo a sè stante, che può essere utilizzato per diversi scopi, questo avviene al fine di evitare conflitti tra i nodi, in quanto il ruolo è già indicato dalle frequenze che usano. Il livello Fisico quindi restituisce per ogni gruppo una struttura dati contenente 2 frequenze, se le 3 frequenze fossero tutte presenti allo stesso momento si avrebbe, allo stato attuale della tecnologia un errore Multi-Tone. In ogni gruppo, le frequenze seguono una logica definita come segue:

- La prima frequenza rappresenta quale **Signal Code** utilizzare dal lato sinistro della tabella che segue
- La seconda frequenza è la portante ed è sempre la stessa per ogni gruppo, questa serve successivamente per la decodifica
- La terza frequenza rappresenta quale **Signal Code** utilizzare dal lato destro della tabella che segue

Frequenza	Signal Code	Portante	Signal Code
1000	(0) Bits: 0		
1400	(1) Bits: 00		
1800	(2) Bits: 000		
2200	(3) Bits: 0000		
2600	(4) Bits: 00000		
3000	(5) Bits: 000000		
3400	(6) Bits: 0000000		
3800	(7) Bits: 0000000 0000000		
4200	(8) Bits: 0000000 0000000 0000000		
4600	(9) <Libero>		
5000			(10) Bits: 1
5400			(11) Bits: 11
5800			(12) Bits: 111
6200			(13) Bits: 1111
6600			(14) Bits: 11111
7000			(15) Bits: 111111
7400			(16) Bits: 1111111
7800			(17) Bits: 1111111 1111111
8200		Master Carrier	
8600		Config Carrier	
9000		Slave Carrier	

**Tabella 3.3.** Mappatura frequenze, codici e portanti.

L'utilizzo di questo schema di codifica, viene dopo l'esecuzione di diversi test, in quanto in origine i dati venivano trasmessi utilizzando 3 frequenze per ogni gruppo:

- La prima frequenza rappresentava se era uno 0
- La seconda frequenza era la portante
- La terza frequenza rappresentava se era un 1

L'utilizzo della Tabella 3.3 ha permesso di comprimere i pacchetti di dati a livello di bit, in quanto ora è possibile trasmettere più bit con una sola frequenza.

Questo ha permesso di aumentare la velocità di trasmissione fino a 5[codes per second], che nel caso più favorevole rappresentano **30bps**, che seppur bassa è comunque un miglioramento rispetto alla versione precedente che raggiungeva al massimo 5bps.

Inoltre, grazie a questa codifica, è possibile trasmettere il codice 8 (21 volte 0) con estrema facilità, questo corrisponde alla ripetizione del carattere ASCII NUL per tre volte, che, nel protocollo corrisponde all' End of Packet (EOP).

### 3.3.1 Decodifica

In fase di decodifica, come preannunciato, il Livello Bit riceve dal Livello antecedente una struttura dati contenente fino a 2 frequenze per ogni gruppo (master, slave, config). Questo associa la frequenza data al corrispondente **Signal Code**, attraverso

$$sn = \frac{f - base}{step} = \frac{f - 1000}{400} \quad (3.5)$$

Dove **base** è la frequenza minima (1000Hz), **step** è la differenza tra una frequenza e l'altra (400Hz) e **f** è la frequenza ricevuta.

Questo permette di ottenere il **Signal Code** che può essere compreso tra 0 e 17, che viene successivamente convertito in bit attraverso la [Tabella 3.3](#).

Durante la fase di conversione tra Signal Code e bit, nel momento in cui viene identificato il Codice 8 (21 volte 0) viene interpretato come End of Packet (EOP), avviene quindi la **memorizzazione del pacchetto in bits nell'apposito buffer**, differenziato per tipo di pacchetto (master, slave, config), in attesa che il Livello Link li svuoti.

### 3.3.2 Codifica

La codifica avviene in maniera opposta alla decodifica, il Livello Bit riceve dal Livello Link un buffer di caratteri ASCII da trasmettere. Ogni carattere ASCII (7 bits) viene convertito in bit, successivamente i bit vengono convertiti in **Signal Code** attraverso la [Tabella 3.3](#). Infine i **Signal Code** vengono convertiti in frequenze attraverso la formula inversa

$$f = base code * step = 1000 * code * 400 \quad (3.6)$$

Dove **base** è la frequenza minima (1000Hz), **step** è la differenza tra una frequenza e l'altra (400Hz) e **code** è il Signal Code da convertire.

Il buffer di frequenze viene successivamente inviato al Livello Fisico per la trasmissione.

## 3.4 Livello Link

Il **Livello Link** riceve le sequenze ASCII dal Livello Bit e realizza l’incapsulamento logico dei messaggi. Le comunicazioni vengono suddivise in tre canali distinti: *config*, *master* e *slave*. Ogni pacchetto segue lo schema generale:

**ID:destinatario{OP}k{mittente}**

dove i campi rappresentano, rispettivamente, l’ID del destinatario, l’operazione da eseguire e l’ID del mittente.

I pacchetti ASCII in arrivo dal Livello Bit vengono inseriti in una coda FIFO di Input e letti dal Livello Link. Una volta estratti, vengono interpretati in base al canale di appartenenza e al comando ricevuto. Analogamente, quando il Livello Link deve inviare dati, i pacchetti vengono preparati ed inseriti nella coda FIFO di Output, in attesa di essere trasformati in frequenze dal Livello Bit.

### 3.4.1 Procedura di registrazione

Quando un nodo si collega per la prima volta, o ha perso il proprio ID, deve eseguire la registrazione sul canale *config*. In questa fase, un nodo senza ID (e non in modalità *hotspot*) invia una richiesta all’hotspot, generando un identificativo provvisorio di **3 cifre alfanumeriche**. La richiesta assume la forma:

**ID:0000{REQ}1{idGenerato}**

Il carattere 1 sostituisce la k perché il nodo non dispone ancora di un ID valido. L’hotspot, ricevuta la richiesta, assegna un ID definitivo di **4 cifre numeriche univoche** e risponde con:

**ID:idAssegnato{SET}k{0000}**

Il nodo registra l’ID in memoria non volatile e conferma con:

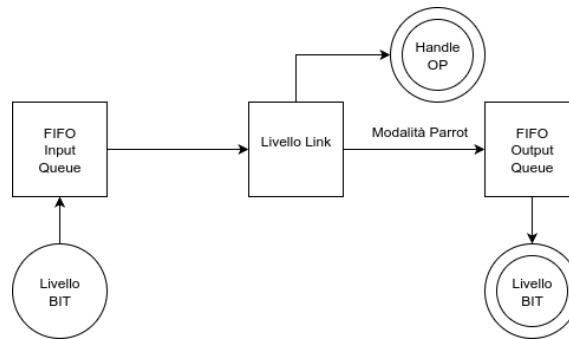
**ID:0000{OK}k{idAssegnato}**

### 3.4.2 Indirizzamento e instradamento

Ogni pacchetto ricevuto viene analizzato per verificare l'indirizzo di destinazione:

- Se l'ID non coincide con quello locale, il nodo entra in modalità *parrot*: memorizza il pacchetto nella coda FIFO di Output e lo ritrasmette, in attesa di un acknowledgment. In assenza di ACK, il pacchetto viene ritrasmesso sullo stesso canale (ad esempio, un messaggio ricevuto in frequenze *master* viene riemesso in frequenze *master* anche se il nodo non è un hotspot).
- Se invece l'ID corrisponde a quello del nodo o all'indirizzo broadcast 1111, il pacchetto viene elaborato in base al comando ricevuto. Se è richiesta una risposta, questa viene generata tramite gli *handle* interni del Link e inserita nella coda di Output *slave*, pronta per essere trasmessa in frequenze *slave*.

Per garantire ordine e separazione dei flussi, il Livello Link mantiene tre code circolari dedicate ai canali *master*, *slave* e *config*, ognuna delle quali gestisce il traffico relativo.



**Figura 3.8.** Schema del flusso dei pacchetti tra FIFO Input, Livello Link e FIFO Output

### 3.4.3 Gestione dei Token (TKN)

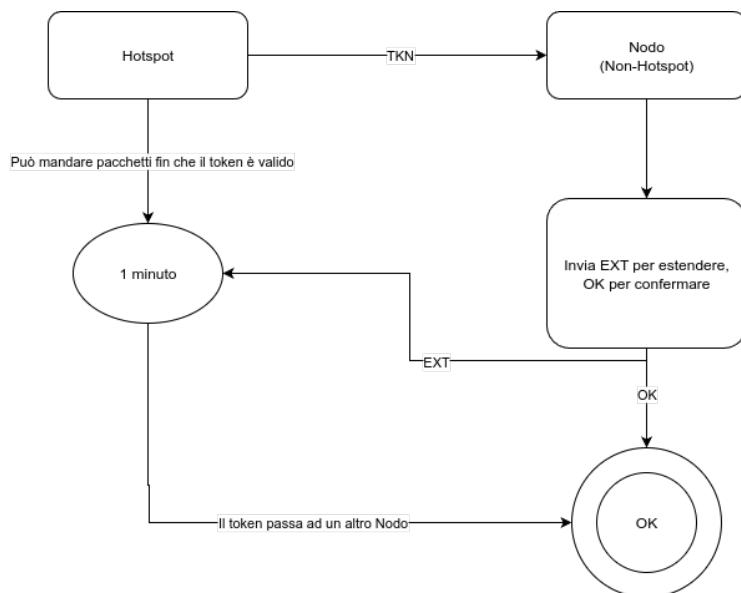
Un aspetto centrale riguarda la possibilità, per i nodi non-hotspot, di trasmettere i pacchetti contenuti nella propria coda di Output solo quando l'hotspot concede loro un *token* (operazione TKN).

Quando un nodo riceve un pacchetto TKN in frequenze *master*, acquisisce il diritto di trasmissione per un tempo massimo di **1 minuto**. Durante questo intervallo, il nodo può:

- inviare un pacchetto con operazione EXT per richiedere l'estensione del token. In tal caso l'hotspot risponde con un acknowledgment (OP:OK) e il tempo si prolunga di un ulteriore minuto;
- inviare un semplice OP:OK per confermare la gestione del token entro il minuto stabilito.

Se entro il tempo limite il nodo non invia alcun messaggio, il token decade automaticamente e l'hotspot lo assegna ad un altro dispositivo. È dunque responsabilità del nodo non-hotspot calcolare i tempi necessari per le proprie trasmissioni future e decidere se richiedere estensioni o accordarsi con l'hotspot per la durata del token.

Questa rappresenta attualmente la parte più delicata del protocollo: al momento della stesura il funzionamento è ancora limitato, ma in prospettiva si prevede la possibilità che anche i nodi non-hotspot possano generare autonomamente token, così da estendere la rete su più livelli e aumentare la flessibilità del sistema. Questa apertura costituisce un'opportunità interessante, benché richieda un'implementazione più solida per evitare l'accavallamento di comunicazioni su canali multipli quando più nodi tentano di gestire i token in parallelo.



**Figura 3.9.** Schema del flusso dei Token tra hotspot e nodi non-hotspot

### 3.4.4 Comandi di controllo e di rete

Il Livello Link gestisce alcune operazioni speciali:

- **ABORT**: comando broadcast (ID: 1111) inviato dall'hotspot e ripetuto due volte in frequenze *master* da tutti i nodi non-hotspot. Non richiede ACK e serve a interrompere le operazioni, svuotando i buffer di Input e Output. Viene tipicamente generato quando la piattaforma *Blynk.io* invia il comando *abort* all'hotspot.
- **OK**: rappresenta l'ACK e viene inviato dai nodi non-hotspot all'hotspot per confermare la ricezione di un pacchetto. Solo in questo momento l'hotspot può eliminare il pacchetto dalla sua coda di Output *master*.

Infine, il Livello Link supporta un comando di rete dedicato, **conns**, che la piattaforma *Blynk.io* invia all'hotspot per ottenere la lista dei nodi attualmente connessi.

### 3.4.5 Indirizzi di rete

Gli indirizzi usati dal protocollo sono riassunti nella seguente tabella.

Indirizzo	Descrizione
0000	ID dell'hotspot
1111	Indirizzo broadcast (tutti i nodi)
0001 – 1110	Range di indirizzi assegnabili ai nodi

**Tabella 3.4.** Indirizzi di rete del Livello Link

### 3.5 Livello Applicazione

Dato che si conoscono già gli usi futuri di questo protocollo e non è destinato al grande pubblico, il dispositivo *hotspot* che riceve in ingresso dal loop **Blynk** un comando nella forma

Comando -> IdDestinatario

si occupa, con un codice "saldato" al livello Link, sia dell'incapsulamento dell'identificativo del destinatario nel formato ID:IdDestinatario, sia della traduzione semantica attraverso la seguente tabella.

Comando	Operazione	Non-Hotspot	Hotspot
movement_sensor_on	MOVEMENT_ON	Esegue misurazione PIR per 5s; restituisce MOVEMENT_YES o MOVEMENT_NO	ACK->OP:OK
movement_sensor_on_[tempo_in_ms]	MOVEMENT_ON_[tempo_in_ms]	Esegue misurazione PIR per l'intervallo specificato; restituisce MOVEMENT_YES o MOVEMENT_NO	ACK->OP:OK

**Tabella 3.5.** Mappatura comandi Blynk, operazioni interne e azioni eseguite dai nodi

Questa tabella mostra come l'hotspot agisca da traduttore e compressore semantico: ogni comando testuale viene ridotto a un codice operativo interno e ogni dispositivo non-hotspot può eseguire le relative funzioni senza dover gestire sintassi testuali complesse.

Il ritorno dell'esecuzione è anch'esso standardizzato (nel caso illustrato, MOVEMENT\_YES / MOVEMENT\_NO), mentre l'hotspot conferma localmente con un messaggio di tipo ACK->OP:OK, garantendo così sia l'affidabilità della comunicazione che la compatibilità futura con ulteriori estensioni già previste.

Quindi l'utente finale che sia l'industria mineraria, soccorritori o tecnici possono digitare un comando tramite l'interfaccia **Blynk.io**, corrisponde ad un'**operazione più compatta** che viene effettivamente interpretata dai nodi della rete.

La **compressione** avviene quindi su due livelli: non solo a livello di codifica dei bit, ma anche a **livello semantico**, poiché i comandi accettati non appartengono a categorie generiche di protocolli di uso comune (web come HTML/PHP, o trasferimento file come PNG/JPG nei protocolli FTP/HTTP), bensì ad un insieme chiuso e proprietario di istruzioni standardizzate.

In questo modo l'interpretazione da parte dei dispositivi risulta rapida e priva di ambiguità, mentre l'interfaccia utente mantiene la leggibilità e la semplicità d'uso.

```
<
| HotSpot Device ON |
\_____|_
Segnale ricevuto: (REQ)l{860
Invio: ID:860{SET:{0001}}k{0000}
Segnale ricevuto: ID:0006{OK}k{0001}
Invio: Nuovo dispositivo registrato con successo, ID: 0001
> conns
< 0001
> h
< Comando non riconosciuto digita help per tutti i comandi
> help
< Comandi disponibili:
CONN: lista dispositivi connessi
ABORT: interrompe la trasmissione corrente
movement_sensor_on[_durata_ms]->ID: attiva il sensore di movimento (default 5000 ms)
REQ->ID, SET->ID, OK->ID, MOVEMENT->ID, EXT->ID: invia il comando indicato al dispositivo
HELP: mostra questo messaggio
> movement_sensor_on->0001
< Invio: ID:0001{MOVEMENT:ON}k{0000}
Segnale ricevuto: ID:0006{MOVEMENT:YES}k{0001}
Invio: ID:0001{OK}z{0000}
Risposta movimento da 0001: YES
Segnale ricevuto: ID:0006{MOVEMENT:YES}k{0000}
Invio: ID:0001{OK}z{0000}
Risposta movimento da 0001: YES
> abort
< Invio: ID:1111{ABORT}k{0000}
Invio: ID:1111{ABORT}k{0000}
> movement_sensor_on_20000->0001
< Invio: ID:0001{MOVEMENT:ON}k{0000}
Invio: ID:0001{MOVEMENT:ON}k{0000}
Segnale ricevuto: ID:0006{MOVEMENT:YES}k{0001}
Invio: ID:0001{OK}z{0000}
Risposta movimento da 0001: YES
> abort
< Invio: ID:1111{ABORT}k{0000}
Invio: ID:1111{ABORT}k{0000}
```

**Figura 3.10.** Utilizzo del Livello Applicazione tramite l'interfaccia Blynk.io

# 4. Implementazione Hardware

## 4.1 Tecnologie di base

In questo capitolo tratteremo tutte le tecnologie utilizzate nel progetto, come i microcontrollori e i dispositivi di output, concentrando ci sull'ESP32.

### 4.1.1 Panoramica dell'ESP32

#### 4.1.1.1 Generalità

I microcontrollori (MCU) sono circuiti integrati che combinano nello stesso chip un microprocessore con una serie di componenti periferici come: memoria, timer, convertitori analogico-digitali e soprattutto pin di I/O.

A differenza di un microprocessore (CPU), i microcontrollori non sono pensati per svolgere onerose attività di calcolo ma per interagire il più possibile con l'ambiente esterno, gestendo numerosi scambi fra le periferiche di Input/Output.

L'ESP32, sviluppato da Espressif Systems, rappresenta un microcontrollore avanzato con architettura dual-core Tensilica LX6 a 32 bit (alcune varianti hanno anche core single-core o LX7), integrando al suo interno Wi-Fi e Bluetooth. Questa caratteristica lo rende particolarmente adatto per applicazioni IoT (Internet of Things), domotica e sistemi embedded complessi [esp, 2020].

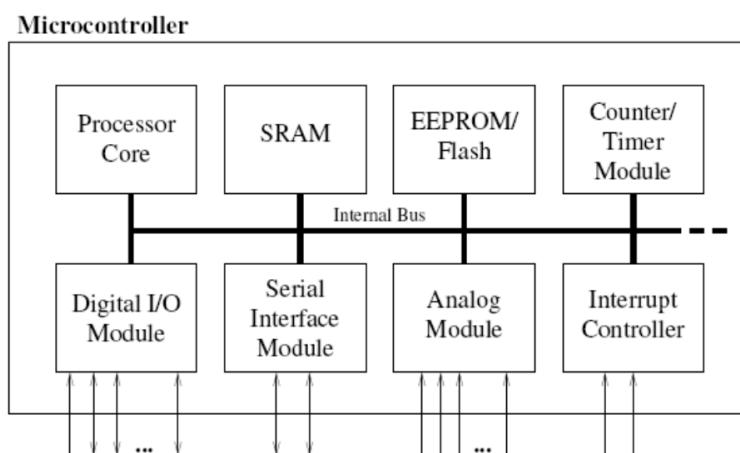
#### 4.1.1.2 Famiglie di ESP32

I microcontrollori sono suddivisi in famiglie a seconda del processore e delle periferiche integrate. Nel caso dell'ESP32, troviamo diverse varianti come ESP32-WROOM-32, ESP32-WROVER, ESP32-C3, ESP32-S2 ed ESP32-S3. Queste versioni condividono la stessa architettura di base ma differiscono per risorse di memoria, numero di GPIO disponibili, supporto a periferiche aggiuntive o funzionalità avanzate come acceleratori AI.

La scelta del modulo da utilizzare deve essere fatta non sulla base delle prestazioni massime, ma dell'ottimalità rispetto al progetto in questione [Patti, 2024].

#### 4.1.1.3 Componenti hardware

Ogni ESP32 integra al suo interno non solo la CPU, ma anche le memorie e una grande quantità di periferiche digitali e analogiche, oltre ai moduli di comunicazione wireless (Wi-Fi 802.11 b/g/n e Bluetooth 4.2/5.0). Questa elevata integrazione consente di ridurre i costi, semplificare il design e garantire maggiore velocità di accesso ai dati.



**Figura 4.1.** Componenti hardware principali di un microcontrollore ESP32.

#### 4.1.1.4 Memorie

Le memorie di un microcontrollore possono essere di tipo volatile e non volatile. Le volatili includono SRAM e DRAM; le non volatili comprendono ROM, EEPROM, Flash e, in alcuni casi, NVRAM.

L'ESP32 tipicamente dispone di:

- 448 KB di ROM, che contiene il bootloader e alcune librerie di base;
- 520 KB di SRAM interna, divisa tra cache, stack e heap per le applicazioni;
- Flash esterna fino a 16 MB (a seconda del modulo), utilizzata per memorizzare codice e file system;
- opzionalmente, PSRAM esterna fino a 8 MB, utile per applicazioni che richiedono elaborazioni multimediali o server web embedded.

## 4.1.2 Periferiche di Input/Output

Le periferiche di Input/Output (I/O) sono componenti essenziali che permettono al microcontrollore di interagire con il mondo esterno. Per la realizzazione del progetto sono state utilizzate diverse periferiche, in questa sezione se ne approfondiscono le loro caratteristiche.

Tutta la componentistica impiegata in questo progetto è stata selezionata dal catalogo **Adafruit Industries**. La scelta è motivata dall'elevata qualità dei loro moduli, dalla cura del layout dei circuiti stampati e dalla ricca documentazione tecnica a corredo, che garantiscono affidabilità in fase di prototipazione e semplicità di integrazione. L'utilizzo di componenti a marchio Adafruit riduce i rischi di incompatibilità o malfunzionamenti tipici dei cloni economici e assicura un supporto didattico di livello professionale.

### 4.1.2.1 Microfono MEMS I2S digitale Adafruit – SPH0645LM4H

Il breakout Adafruit monta il microfono MEMS digitale **SPH0645LM4H-B**, bottom-port, con uscita I<sup>2</sup>S [ada, 2015, p. 1]. È un dispositivo *slave* che richiede dal master le linee di clock (BCLK, LRCLK) [nor, 2019]. Fornisce direttamente campioni PCM a 24 bit (con 18 bit effettivi), evitando stadi analogici esterni [ada, 2015, p. 3].



**Figura 4.2.** Microfono MEMS I2S digitale Adafruit SPH0645LM4H-B.

## Caratteristiche principali.

- Alimentazione: 1.62 V a 3.6 V, logica 3.3 V, tipico consumo  $600 \mu\text{A}$  [ada, 2015, p. 2].
- Interfaccia: I<sup>2</sup>S, pin BCLK, LRCLK, DOUT, SEL, più 3V e GND.
- Formato dati: parole a 24 bit (MSB first, complemento a due), 18 bit utili, oversampling ratio (OSR) = 64 [ada, 2015, p. 3].
- Prestazioni: SNR tip. 65 dBA, AOP 120 dB SPL, sensibilità  $\sim -26$  dBFS @ 94 dB SPL [ada, 2015, p. 2].

**Pinout e collegamenti.** Due moduli possono essere usati in stereo condividendo BCLK, LRCLK, DOUT; SEL basso/alto definisce L/R. In modalità mono basta un solo modulo con SEL a GND (Left di default) [ada, 2015].

**Interfaccia I<sup>2</sup>S.** Il microfono trasmette in formato *Philips*: ogni semiperiodo di LRCLK contiene una parola da 24 bit. La relazione fondamentale è  $\text{BCLK} = 64f_s$  [nor, 2019].

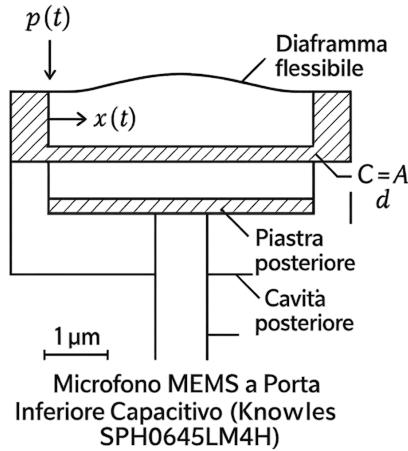
## Catena del segnale interna.

1. Trasduttore MEMS capacitivo: diaframma + backplate  $\rightarrow$  capacità  $C_t$ .
2. Amplificatore a carica: traduce  $\Delta C$  in tensione.
3. Modulatore  $\Sigma\Delta$ : noise-shaping fuori banda audio.
4. Filtro di decimazione: riduce a  $f_s$  mantenendo 18 bit effettivi.
5. Serializer: prepara le parole I<sup>2</sup>S.

**Fondamenti fisici.** La pressione  $p(t)$  agisce sul diaframma (massa  $m$ , molla  $k$ , smorzamento  $b$ ) [Industries, 2016b, pp. 192–195]:

$$mx \ddot{b}x - kx = S p(t)$$

La variazione di capacità è  $\Delta CC_0 \approx xd_0$ , con risonanza a  $\omega_0 = \sqrt{km}$ . Il rumore complessivo deriva da contributi termico–meccanici, elettronici e quantizzazione.



**Figura 4.3.** Sezione verticale del microfono Adafruit MEMS.

**Limiti.** Non tollera 5 V, banda garantita  $\sim 20 \text{ Hz} - 10 \text{ kHz}$  (utilizzabile fino a  $\sim 15 \text{ kHz}$  secondo Adafruit) [ada, 2015, p. 2].

**Relazione dBFS–SPL.**  $94 \text{ dB SPL} = 1 \text{ Pa} \Rightarrow \approx -26 \text{ dBFS RMS}$  [ada, 2015, p. 2]. L'AOP a 120 dB SPL ( $\sim 20 \text{ Pa}$ ) indica il massimo senza distorsione significativa.

#### 4.1.2.2 Adafruit MAX98357A I<sup>2</sup>S Amplificatore Classe-D

Il breakout Adafruit basato sul chip **MAX98357A** è un amplificatore audio digitale in classe D, in grado di convertire direttamente un flusso I<sup>2</sup>S in segnale audio analogico amplificato per pilotare altoparlanti a bassa impedenza. La caratteristica principale è l'eliminazione della necessità di un DAC esterno: il MAX98357A integra al suo interno la conversione digitale-analogica, il filtraggio, e la sezione di potenza [Industries, 2016a].



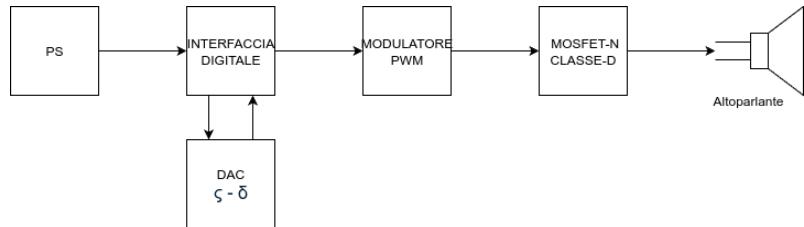
**Figura 4.4.** Amplificatore digitale Classe-D Adafruit MAX98357A.

#### Caratteristiche essenziali.

- **Alimentazione:** 2.5 V a 5.5 V, con tensione tipica 3.3 V o 5 V.
- **Potenza d'uscita:** fino a 3.2 W su 4 Ω a 5 V, THD+N ≈ 1 %.
- **Interfaccia:** I<sup>2</sup>S standard (BCLK, LRCLK, DIN); pin aggiuntivi per modalità (GAIN, SD, DMP).
- **Efficienza:** superiore all'85 % grazie all'architettura in classe D.
- **Gamma dinamica:** ≈ 98 dB, con rumore di fondo contenuto [Industries, 2016a, p. 16].

**Funzionamento.** Il MAX98357A riceve un flusso I<sup>2</sup>S in formato PCM (16, 24 o 32 bit, standard Philips). L'architettura interna comprende:

1. **Interfaccia I<sup>2</sup>S:** decodifica il flusso digitale e gestisce il framing (left/right).
2. **DAC sigma-delta:** converte i campioni PCM in una modulazione a densità di impulsi.
3. **Driver in classe D:** due MOSFET push-pull pilotano direttamente l'altoparlante in configurazione *bridge-tied load (BTL)*.
4. **Filtro LC virtuale:** la stessa induttanza del carico acustico e la risposta del driver fungono da filtro passa-basso, eliminando la componente PWM ad alta frequenza.



**Figura 4.5.** Funzionamento interno del MAX98357A.

**Motivazioni fisiche del funzionamento.** L'amplificatore in classe D si basa su modulazione a *duty cycle* dei MOSFET: il segnale analogico ricostruito corrisponde al valore medio del PWM. L'elevata efficienza deriva dal fatto che i transistor lavorano quasi sempre in saturazione o interdizione, riducendo le perdite per dissipazione. L'integrazione del DAC sigma-delta consente un noise-shaping che sposta il rumore di quantizzazione fuori banda audio, garantendo un'ottima qualità percepita [Industries, 2016a].

## 4.2 Progettazione digitale

Per la progettazione digitale del sistema è stato utilizzato il software **Fritzing**, un tool open-source per la progettazione di circuiti elettronici, con la possibilità di creare moduli personalizzati.

Nel Capitolo 3 è stato descritto il protocollo di comunicazione, facendo riferimento al **Livello Fisico**, come il livello più basso del modello. L'interazione I/O con il mondo circostante che avviene nel Livello Fisico, è realizzata con due componenti: una software, che tratteremo a breve nel Capitolo 5, e una hardware, che vedremo in questo capitolo.

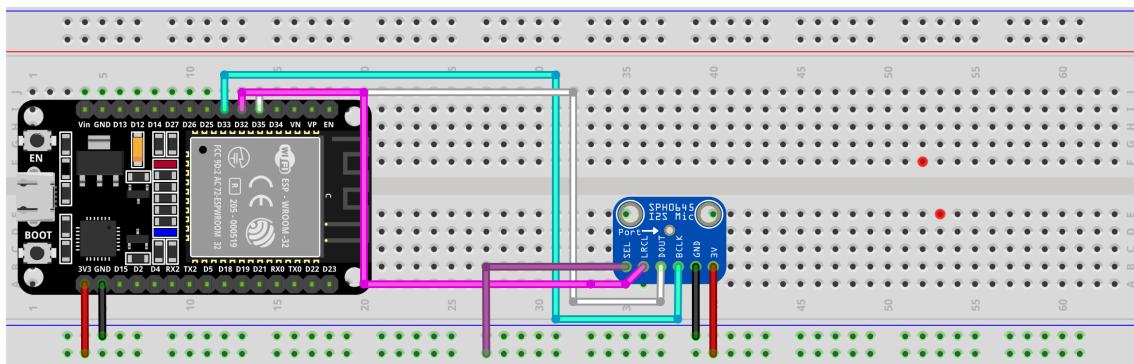
Il Livello Fisico è realizzato con due periferiche di Input/Output, il microfono MEMS digitale e l'amplificatore Classe-D, che comunicano con il microcontrollore ESP32 tramite il protocollo I<sup>2</sup>S.

Per la realizzazione del circuito elettronico sono stati utilizzati due breakout board Adafruit, che integrano i componenti necessari per il corretto funzionamento del microfono e dell'amplificatore.

L'utilizzo del microfono MEMS I<sup>2</sup>S digitale Adafruit SPH0645LM4H viene da una selezione accurata, basata su test di qualità, il progetto, che all'inizio prevedeva un microfono analogico, è stato modificato per utilizzare un microfono digitale, in quanto i test effettuati con il microfono analogico non hanno dato risultati soddisfacenti.

#### 4.2.1 Collegamento del microfono digitale I<sup>2</sup>S

La Figura 4.6 mostra lo schema di cablaggio realizzato in **Fritzing** tra la breakout Adafruit basata sul microfono MEMS digitale **SPH0645LM4H** e la **ESP32**, mediante il bus I<sup>2</sup>S (Inter-IC Sound). L'interfaccia prevede tre linee digitali di segnale (BCLK, LRCLK, DOUT) e due linee di alimentazione (3V3, GND), oltre al pin SEL per la selezione del canale logico (sinistro/destro).



**Figura 4.6.** Collegamento fisico tra ESP32 e microfono I<sup>2</sup>S (Adafruit SPH0645LM4H) su bread-board.

**Pinout del modulo microfonico.** Sulla scheda Adafruit, da sinistra a destra (come serigrafato) sono presenti i pin: SEL, LRCL (o LRCLK/WS), DOUT (o SD), BCLK (o SCK), GND, 3V. DOUT è un'uscita a logica 3.3 V; LRCLK e BCLK sono ingressi di clock; SEL

è un ingresso statico per selezionare su quale slot (sinistro o destro) il microfono presenta i campioni sul bus I<sup>2</sup>S.

**Mappatura dei segnali sull'ESP32.** La seguente tabella riassume la corrispondenza dei pin usati, come nella figura. Le scelte rispettano i vincoli dell'ESP32 (evitando pin con funzioni di strapping all'avvio come GPIO15-12 o limitazioni di sola-lettura).

Pin modulo	Segnale	Pin ESP32 (sigla serigrafata)
3V	Alimentazione 3.3 V	3V3
GND	Riferimento di massa	GND
BCLK	Bit Clock I <sup>2</sup> S	GPIO14 (D14)
LRCL	Word-Select / LRCLK	GPIO25 (D25)
DOUT	Serial Data (uscita microfono)	GPIO32 (D32)
SEL	Selezione canale	GND → canale sinistro

**Tabella 4.1.** Mappatura hardware dei pin I<sup>2</sup>S tra ESP32 e microfono SPH0645LM4H.

Il segnale SEL determina in quale slot vengono emessi i campioni audio: se portato a massa (SEL = GND) i dati escono sul canale sinistro, mentre a livello alto (SEL = 3V3) sul destro.

Nel nostro schema è stato fissato a 3V3, quindi l'uscita è forzata a destra. Questo a seguito di alcuni test dove si è verificato che per qualche difformità nel IC il canale destro garantiva maggiore precisione.

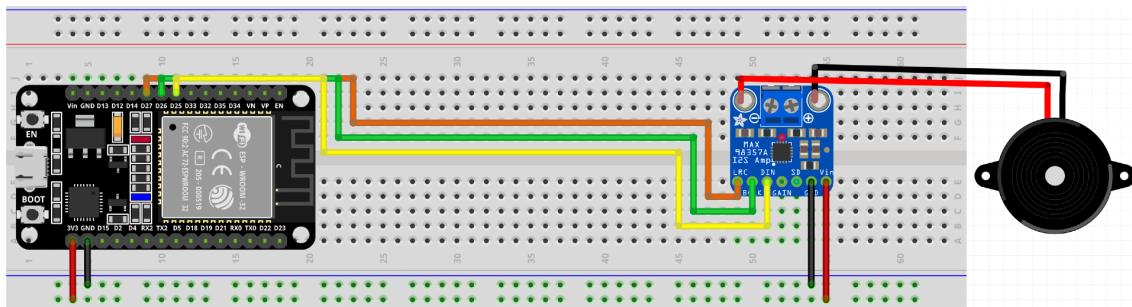
**Dettagli elettrici e motivazioni delle scelte.** Il microfono Adafruit SPH0645LM4H funziona con una tensione fino a 3.6 V e un'alimentazione a 5 V lo danneggerebbe [ada, 2015, p. 2].

Per l'ESP32 si è scelto **GPIO14 come BCLK, GPIO25 come LRCLK e GPIO32** per il dato in ingresso.

Infine, il microfono trasmette campioni a 24 bit allineati su frame da 32 come discusso precedentemente nella sottosottosezione 4.1.2.1.

## 4.2.2 Collegamento dell'amplificatore digitale I<sup>2</sup>S

La Figura 4.7 mostra lo schema di cablaggio realizzato in **Fritzing** tra la breakout Adafruit basata sul chip **MAX98357A** (amplificatore in classe D con interfaccia I<sup>2</sup>S) e la **ESP32**. L'interfaccia prevede tre linee digitali di segnale (BCLK, LRCLK, DIN) e due linee di alimentazione (3V3, GND). Il modulo espone inoltre i pin di configurazione **GAIN** e **SD** (shutdown), che in questo caso non vengono utilizzati e restano flottanti.



**Figura 4.7.** Collegamento fisico tra ESP32 e amplificatore I<sup>2</sup>S (Adafruit MAX98357A) su breadboard.

**Pinout del modulo amplificatore.** Sulla scheda Adafruit, da sinistra a destra (come serigrafato) sono presenti i pin: GAIN, SD, DIN, BCLK, LRCLK, GND, Vin. Il pin **DIN** è l'ingresso dati audio seriali (logica 3.3 V); **BCLK** e **LRCLK** sono segnali di clock forniti dall'ESP32; **Vin** riceve l'alimentazione (3–5 V, compatibile col regolatore onboard). I pin **GAIN** e **SD** sono opzionali: **GAIN** permette di configurare il guadagno interno tramite resistenze, mentre **SD** forza il chip in modalità risparmio energetico.

**Mappatura dei segnali sull'ESP32.** La seguente tabella riassume la corrispondenza dei pin usati, come nello schema. Anche in questo caso la scelta dei GPIO tiene conto delle raccomandazioni di Espressif, evitando pin con funzioni critiche all'avvio.

Pin modulo	Segnale	Pin ESP32 (sigla serigrafata)
Vin	Alimentazione 3.3 V / 5 V	3V3
GND	Riferimento di massa	GND
BCLK	Bit Clock I <sup>2</sup> S	GPIO14 (D14)
LRCLK	Word-Select / LRCLK	GPIO25 (D25)
DIN	Serial Data (ingresso audio)	GPIO26 (D26)
SD	Shutdown (non connesso)	—
GAIN	Configurazione guadagno (non connesso)	—

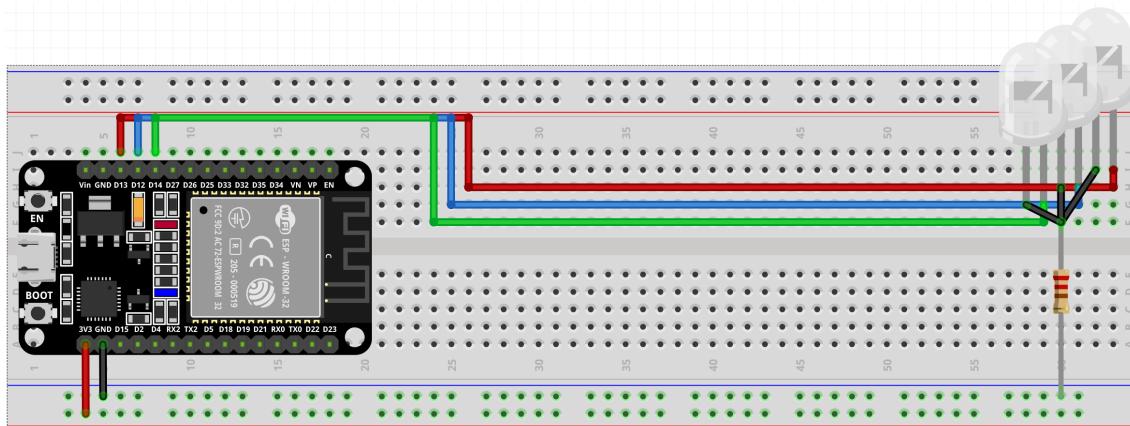
**Tavella 4.2.** Mappatura hardware dei pin I<sup>2</sup>S tra ESP32 e amplificatore MAX98357A.

Il collegamento all'altoparlante avviene direttamente dai terminali di uscita del modulo (OUT, OUT–), come mostrato in figura. Essendo un amplificatore in classe D capace di pilotare piccoli altoparlanti ( $4\text{--}8\ \Omega$ , fino a circa 3 W a 5 V), non è necessaria ulteriore elettronica per il collegamento all'altoparlante.

**Dettagli elettrici e motivazioni delle scelte.** Il MAX98357A accetta alimentazione sia a 3.3 V che a 5 V [Industries, 2016a, p. 4]: in questo schema si è scelto di usare il **3V3 dell'ESP32** per semplicità, tenendo conto che ciò limita la potenza massima erogabile all'altoparlante e considerando che in una reale applicazione a scopo industriale si utilizzerebbero amplificatori da  $\tilde{2}30$  V. Il consumo è nell'ordine di poche decine di mA a volume medio, quindi l'uscita 3V3 dell'ESP32 è sufficiente in prototipazione.

#### 4.2.3 Sistema di segnalazione LED

La Figura 4.8 mostra lo schema di cablaggio realizzato in **Fritzing** tra la scheda **ESP32** e tre LED collegati su breadboard. Il sistema fornisce un feedback visivo sullo stato di decodifica del segnale audio, a supporto del protocollo discusso nella sottosottosezione 3.2.1.1.



**Figura 4.8.** Collegamento fisico tra ESP32 e LED di segnalazione su breadboard.

**Funzionamento del sistema.** Sono stati utilizzati tre LED di colore diverso (verde, blu e rosso) per rappresentare in modo immediato lo stato di elaborazione:

- LED verde: il segnale è stato decodificato correttamente;
- LED blu: il segnale è in fase di decodifica;
- LED rosso: il segnale non ha superato il controllo dell'interpolazione parabolica e non è stato decodificato correttamente.

**Mappatura dei segnali sull'ESP32.** La seguente tabella riassume la corrispondenza dei pin usati per il pilotaggio dei tre LED, come nello schema in figura. Ogni LED è collegato in serie a una resistenza di limitazione, necessaria a ridurre la corrente entro i limiti di sicurezza per i GPIO dell'ESP32 (massimo 12 mA consigliati).

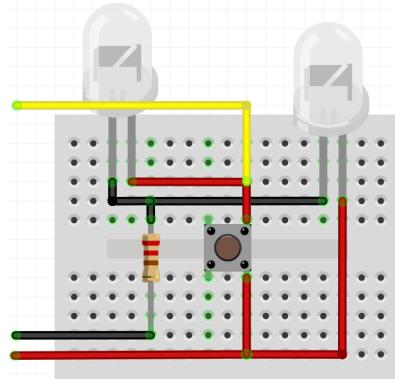
LED	Funzione	Pin ESP32 (sigla serigrafata)
Verde	Decodifica corretta	GPIO14 (D14)
Blu	Decodifica in corso	GPIO27 (D27)
Rosso	Errore di decodifica	GPIO26 (D26)

**Tabella 4.3.** Mappatura hardware dei pin di uscita dell'ESP32 verso i LED di stato.

**Dettagli elettrici e motivazioni delle scelte.** Ciascun LED è collegato al rispettivo GPIO tramite una resistenza da  $220\Omega$ , sufficiente a limitare la corrente senza ridurre eccessivamente la luminosità. I pin **GPIO14**, **GPIO27** e **GPIO26** sono stati scelti in quanto liberi da vincoli di boot e comunemente usati per uscite digitali.

#### 4.2.4 Bottone di Hotspot

La Figura 4.9 mostra il collegamento su breadboard di un pulsante e due LED all'ESP32. Il pulsante è collegato al pin **GPIO34 (D34)** e consente di abilitare la modalità *hotspot* prevista dal protocollo (sezione 3.4).



**Figura 4.9.** Schema di collegamento del bottone e dei LED di stato per la modalità hotspot.

Il LED giallo, a sinistra, indica l'attivazione dell'hotspot e si accende solo alla pressione del pulsante. Il LED a destra segnala invece lo stato di alimentazione della scheda, acceso quando l'ESP32 è in funzione e spento in assenza di alimentazione.

#### 4.2.5 Circuito completo

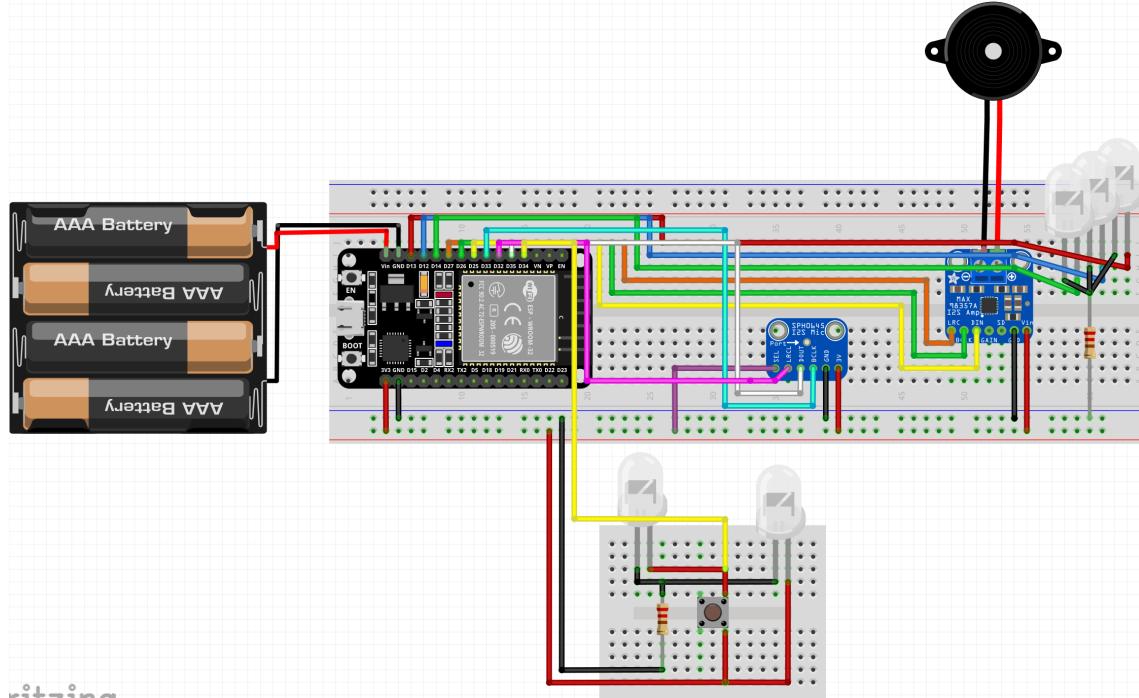
La Figura 4.10 riassume l'intero sistema realizzato su breadboard, integrando in un unico schema tutte le funzionalità descritte nelle sezioni precedenti. L'**ESP32**, alimentato da un pacco batterie AAA, costituisce il cuore del circuito e gestisce sia la parte di acquisizione sia quella di elaborazione e segnalazione.

Il microfono **I<sup>2</sup>S** raccoglie il segnale audio e lo invia al microcontrollore, che a sua volta lo riproduce attraverso l'amplificatore digitale **MAX98357A** collegato a un piccolo altoparlante.

In parallelo, il sistema integra un'interfaccia visiva con tre LED che indicano lo stato della decodifica, insieme a un pulsante dedicato all'attivazione della modalità *hotspot*, il cui stato è reso evidente da un ulteriore LED giallo.

Tutti i collegamenti sono realizzati su breadboard, con le linee di alimentazione distribuite ai vari moduli e i segnali instradati verso i pin dell'ESP32, tra cui il **GPIO34**, correttamente utilizzato come ingresso per il pulsante.

In questo modo il circuito offre un esempio completo di integrazione tra acquisizione, elaborazione, riproduzione e interfaccia utente, consentendo il passaggio fluido dall’ascolto alla decodifica, fino alla segnalazione visiva e al controllo della connessione wireless.



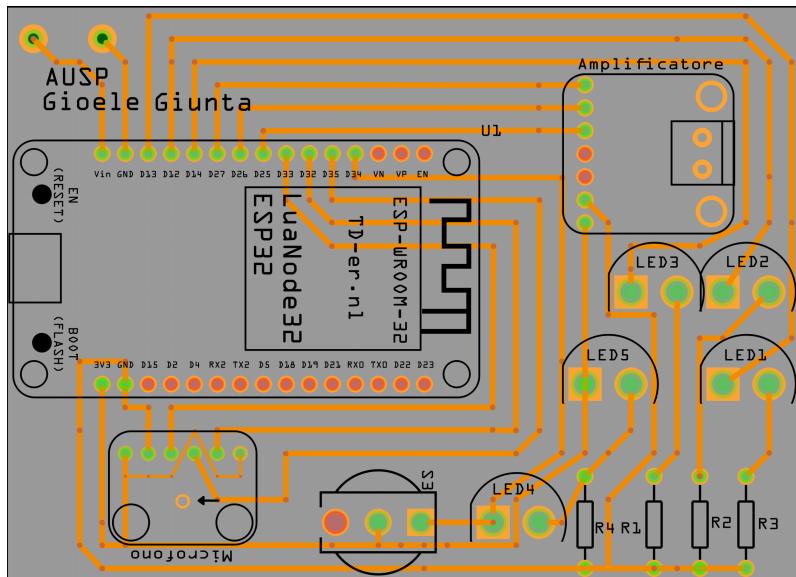
**Figura 4.10.** Schema del circuito completo con ESP32, microfono I<sup>2</sup>S, amplificatore MAX98357A, LED di stato e pulsante di hotspot.

#### 4.2.6 Realizzazione del PCB

Dopo aver completato lo schema su breadboard, il progetto è stato trasferito in **vista PCB** all’interno di Fritzing, con le dovute modifiche, ottenendo così una scheda stampata ordinata e compatta, come mostrato nella Figura 4.11.

In questa configurazione l’**ESP32** rimane al centro della scheda, con le piste di segnale distribuite verso i moduli periferici: i LED di stato con le relative resistenze sono disposti lungo il bordo destro per facilitare la visibilità [LED-1-2-3], mentre il pulsante di hotspot è collocato nella parte inferiore [S2] insieme al LED [LED4] che ne segnala l’attivazione. Il layout delle piste è stato ottimizzato per ridurre le interferenze e mantenere separati i percorsi di potenza da quelli di segnale.

Inoltre il PCB è stampato su un solo strato, questo ha richiesto un’attenta disposizione dei componenti e delle piste per evitare incroci, ma ha diversi vantaggi come la riduzione dei costi di produzione e la semplificazione del processo di fabbricazione.



**Figura 4.11.** Layout del circuito stampato (PCB), con ESP32, LED di stato, pulsante di hotspot e connettori di alimentazione.

# 5. Implementazione Software: Design del Protocollo

## 5.1 Architettura generale del sistema

Il software è stato sviluppato in linguaggio **C**, scelta che ha permesso di avere un controllo diretto sull'hardware dell'ESP32, ottenendo esecuzioni prevedibili ed un uso minuzioso delle risorse.

L'utilizzo del C ha inoltre garantito maggiore leggerezza rispetto al C++ e una buona portabilità, grazie alla disponibilità di librerie già consolidate in ambito firmware.

Alcune librerie Arduino originariamente scritte in C++ sono state adattate in C, questo per permettere la compilazione del codice.

Allo stesso tempo, l'uso del framework Arduino su ESP32 ha offerto vantaggi importanti tra cui: documentazione estesa, supporto diretto alle periferiche del microcontrollore e una community ampia, che ha velocizzato la fase di sviluppo.

Per lo sviluppo è stato utilizzato *Visual Studio Code*, scelto per l'ampia disponibilità di estensioni. In particolare, l'integrazione con **PlatformIO** ha reso più semplice la gestione del progetto: compilazione, caricamento e debugging sono stati unificati in un unico ambiente; la gestione automatica delle librerie e delle dipendenze ha permesso uno sviluppo più rapido ed organizzato.

L'architettura complessiva si basa su moduli separati e specializzati, ognuno responsabile di una parte precisa del protocollo. Questo approccio modulare ha garantito una migliore manutenibilità del codice e ha semplificato l'integrazione tra i diversi livelli funzionali descritti in Capitolo 3, il cui codice sorgente completo è riportato in Appendice A.

## 5.2 Emissione dei toni

L'uscita audio discussa nella sottosezione 3.2.2 è gestita dal modulo **audio\_driver**, responsabile dell'inizializzazione del bus I<sup>2</sup>S e della generazione in tempo reale delle sinusoidi che rappresentano i bit.

La funzione **play\_two\_tones** produce due toni simultanei, richiedendo come parametri le due frequenze delle quali dovrà generare le sinusoidi. La generazione del segnale avviene con persistenza di fase, evitando discontinuità tra un'emissione e la successiva.

```
1 void play_two_tones(int freq1, int freq2) {
2     if(freq1 == 0 && freq2 == 0){
3         delay(80);
4     } else {
5         const float tone_duration = 0.024f;
6         const int tone_samples = (int)(G_SAMPLE_RATE *
7             tone_duration);
8         const int tone_buffer_size = tone_samples * 2;
9         int16_t tone_buffer[tone_buffer_size];
```

Le variabili sotto garantiscono la continuità di fase tra chiamate successive della funzione:

```
1 static float phase1 = 0.0f;
2 static float phase2 = 0.0f;
3 const float inc1 = 2.0f * PI * freq1 / G_SAMPLE_RATE;
4 const float inc2 = 2.0f * PI * freq2 / G_SAMPLE_RATE;
```

Questo frammento di codice mostra la generazione del buffer di campioni, che viene poi scritto sul bus I<sup>2</sup>S:

```
1 for (int i = 0; i < tone_samples; i++) {
2     float mixed = sinf(phase1) + sinf(phase2);
3     int16_t sample = (int16_t)(3000 * (mixed / 2.0f));
4     tone_buffer[2 * i] = sample;
5     tone_buffer[2 * i + 1] = sample;
6     phase1 += inc1;
7     if (phase1 >= 2.0f * PI) phase1 -= 2.0f * PI;
8     phase2 += inc2;
```

```

9      if (phase2 >= 2.0f * PI) phase2 -= 2.0f * PI;
10     }
11     size_t bytes_written = 0;
12     i2s_write(I2S_NUM, tone_buffer, sizeof(tone_buffer),
13               &bytes_written, portMAX_DELAY);
14   }

```

L'ampiezza massima del segnale è limitata a un valore inferiore a 32767 per evitare saturazioni sul DAC, questa conclusione deriva da test che hanno mostrato come il fattore di scaling a **3000** garantisce **headroom sufficiente**. Quando entrambe le frequenze sono nulle, la routine introduce una pausa di **80 ms**, che equivale a una “inattività” progettata per separare gruppi di simboli o pacchetti.

### 5.3 Campionamento e bufferizzazione

Il percorso inverso, ovvero l'acquisizione dei toni ricevuti, discusso nella sottosezione 3.2.1 come l'ingresso del Livello Fisico è affidato al **reader**.

Il modulo crea due buffer circolari tra cui effettua un meccanismo di **swapping sincronizzato**: mentre uno viene riempito dal task di I<sup>2</sup>S, l'altro è disponibile per l'elaborazione FFT, questo avviene attraverso un insieme di variabili rese **extern** e mediante l'uso di un **puntatore** con attributo **extern** che punta sempre al buffer pronto, così da renderlo disponibile alla funzione richiedente senza che questa entri nella logica del reader.

```

1 extern volatile int data_ready;
2 extern volatile int status_flag;
3 extern complex_g3_t *array_ready;

```

La funzione **reader\_task** legge i campioni dal bus I<sup>2</sup>S, li converte in numeri complessi e li memorizza nel buffer corrente.

```

1 static void reader_task(void *param) {
2     size_t bytes_read;
3     int32_t dma_buffer[DMA_BUFFER_SIZE / 4];

```

```

4     static float dc_mean = 0.0f;
5     const float alpha = 1.0f / 1024.0f;
6     while (1) {
7         i2s_read(I2S_PORT, (void*)dma_buffer, DMA_BUFFER_SIZE,
8             &bytes_read, portMAX_DELAY);
9         int samples = bytes_read / 4;
10        for (int i = 0; i < samples; i++) {
11            int32_t s32 = dma_buffer[i] >> 8;
12            float x = (float)(s32 >> 12);
13            dc_mean += alpha * (x - dc_mean);
14            float val = x - dc_mean;
15            current_data[counter].re = (double)val;
16            current_data[counter].im = 0.0;
17            counter++;
18            if (counter >= ARRAY_ELEMENTS) {
19                data_ready = 1;
20                swap_array();
21                counter = 0;
22            }
23        }
24    }

```

L'eliminazione della componente DC avviene con una media mobile parametrizzata da **alpha**,

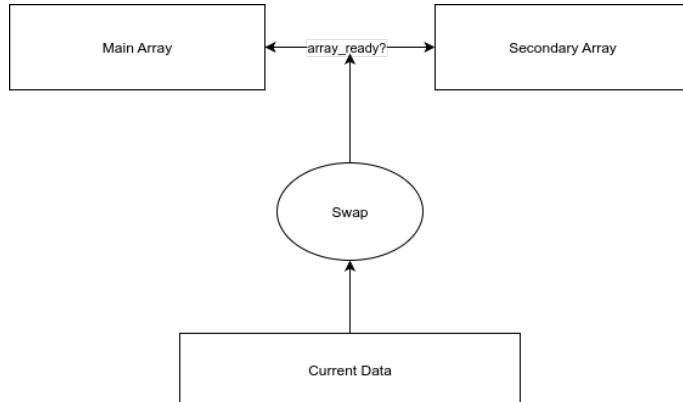
il quale viene; i campioni centrati vengono convertiti in numeri complessi con parte immaginaria nulla per l'elaborazione spettrale.

L'uso di **swap\_array** consente di non perdere alcun campione e di garantire un'elaborazione a flusso continuo.

```

1     static void swap_array(void) {
2         array_ready = current_data;
3         current_data = (current_data == main_array) ? secondary_array
4             : main_array;
5     }

```



**Figura 5.1.** Swapping degli array di campioni

## 5.4 Trasformata veloce di Fourier

Altro aspetto critico del Livello Fisico è la trasformata rapida di Fourier (FFT), che converte i campioni temporali in rappresentazioni spettrali. Così come discusso nella , l’implementazione adottata è una versione personalizzata dell’algoritmo di Cooley–Tukey, ottimizzata per il contesto real-time e per la risoluzione richiesta.

L’analisi frequenziale viene eseguita nel modulo **fft.c**, questa include due funzioni principali **FFT\_get\_twiddle\_factors** e **FFT\_calculate**.

```

1 void FFT_calculate (complex_g3_t *x, long N, complex_g3_t *X,
2                     complex_g3_t *scratch, complex_g3_t *twiddles){
3     int k, m, n;
4     int skip;
5     boolean evenIteration = N & 0x55555555;
6     complex_g3_t* E;
7     complex_g3_t* Xp, *Xp2, *Xstart;
8     if (N == 1){
9         X[0] = x[0];
10        return;
11    }
12    E = x;
13    for (n = 1; n < N; n = n * 2){
14        Xstart = evenIteration ? scratch : X;
15        skip = N / (2 * n);

```

```

16    Xp = Xstart;
17    Xp2 = Xstart + N / 2;
18    for (k = 0; k < n; k++){
19        double tim = twiddles[k * skip].im;
20        double tre = twiddles[k * skip].re;
21        for (m = 0; m < skip; ++m){
22            complex_g3_t* D = E + skip;
23            double dre = D->re * tre - D->im * tim;
24            double dim = D->re * tim + D->im * tre;
25            Xp->re = E->re + dre;
26            Xp->im = E->im + dim;
27            Xp2->re = E->re - dre;
28            Xp2->im = E->im - dim;
29            ++Xp; ++Xp2; ++E;
30        }
31        E += skip;
32    }
33    E = Xstart;
34    evenIteration = !evenIteration;
35}
36}

```

Lo schema fa uso di array globali per l'uscita e per lo spazio temporaneo, così da evitare allocazioni dinamiche che in un contesto real-time sarebbero troppo costose.

Il flag **evenIteration** serve a gestire il ping-pong tra l'array risultato e quello di appoggio: in questo modo si alternano i ruoli dei due buffer a ogni passo, senza dover copiare dati inutilmente.

## 5.5 Decodifica spettrale

Le frequenze individuate dalla trasformata passano al modulo **decoder**, che ha il compito di capire se un picco spettrale corrisponde davvero a un tono del protocollo oppure se si tratta di rumore. Questa logica è implementata nella funzione **check\_active\_frequencies**, che lavora in tre fasi: calcolo della soglia dinamica, controllo dei massimi locali e interpolazione del picco confermato.

```

1 struct_interpolated_frequency check_active_frequencies(
2     complex_g3_t *data, int bin_1, int bin_2,
3     int id, double noise_floor){
4     int i, j;
5     struct_interpolated_frequency detected_freq;
6     detected_freq.work = 0;
7     detected_freq.frequency = -1.0;
8     detected_freq.estimated_amplitude = -1.0;
9     detected_freq.dynamic_amplitude_threshold = -1.0;

```

La funzione crea una struttura **detected\_freq** che conterrà il risultato. Tutti i campi sono inizializzati a valori “non validi” (ad esempio -1.0) così che, se nessuna frequenza viene trovata, il chiamante può accorgersene immediatamente.

```

14     for (j = bin_1; j <= bin_2; j++){
15         double freq = (double)(FS * j) / NN;
16         double amp = complex_magnitude(data[j]);

```

Il ciclo scorre i bin tra **bin\_1** e **bin\_2** che, come discusso nel sottosezione 3.2.1.1 i bin sono la rappresentazione discreta della frequenza nello spettro ottenuto dalla FFT.

$$bini = FS * iNN = 48000 * 512 = 93.75 * i. \quad (5.1)$$

La formula calcola la frequenza corrispondente a ogni bin, dove **FS** è la frequenza di campionamento e **NN** il numero di punti della FFT.

Il BIN è quindi un indice che rappresenta una specifica frequenza nell’array risultante dalla FFT.

La funzione **complex\_magnitude** calcola il modulo del numero complesso associato al bin corrente, che rappresenta l’ampiezza del segnale a quella frequenza.

```

18     if(G_LINEAR_REGRESSION_MODE == 0){
19         double dynamic_amplitude_threshold = noise_floor *
20             8.0;
21         if (amp > dynamic_amplitude_threshold){
22             for(i = j-6; i < j + 6 &&
23                 complex_magnitude(data[i]) <= amp; i++) {}

```

```

22     if (i == j+6){
23         detected_freq =
24             → interpolate_peak_frequency(data, j, FS,
25             → NN);
26         detected_freq.dynamic_amplitude_threshold =
27             → dynamic_amplitude_threshold;
28         detected_freq.work = 1;
29         return detected_freq;
30     }
31 }
```

La soglia dinamica (dynamic\_amplitude\_threshold) è otto volte il **noise\_floor**. Una volta superata la soglia, viene controllata una finestra di 12 bin, come rappresentato in Figura 3.5 (sei a sinistra e sei a destra) per assicurarsi che il valore corrente sia davvero un massimo locale.

Solo in questo caso si invoca **interpolate\_peak\_frequency**, le quali motivazioni e funzionamento sono spiegate nel Figura 3.2.1.1, che con una formula di interpolazione parabolica calcola la frequenza “vera”, per sopperire alla bassa risoluzione spettrale, questa tecnica, quindi fornisce una precisione sub-bin, cioè di una frequenza intermedia tra due bin adiacenti.

Infine, se nessuna frequenza valida è trovata, la funzione ritorna la struttura con i valori -1.0.

## 5.6 Traduzione frequenza-bit

Una volta isolati i picchi, i moduli riguardanti il Livello Fisico lasciano spazio ai **moduli del Livello Bit** il primo passo è tradurli in **Signal Codes** come nella seguente tabella. Questa operazione è svolta dal modulo **bit\_freq\_codec**, e in particolare dalla funzione **interpret\_bits**.

Frequenza	Signal Code (0)	Signal Code (1)
1000	(0)	
1400	(1)	
1800	(2)	
2200	(3)	
2600	(4)	
3000	(5)	
3400	(6)	
3800	(7)	
4200	(8)	
4600	(9)	
5000		(10)
5400		(11)
5800		(12)
6200		(13)
6600		(14)
7000		(15)
7400		(16)
7800		(17)

**Tabella 5.1.** Mappatura frequenze con divisione dei Signal Code in 0 e 1

```
1 int interpret_bits(int freqs[3]){
2     int count = 0;
3     int active[3] = {0};
4     for (int i = 0; i < 3; ++i) {
5         if (freqs[i] > 0) {
6             active[i] = 1;
7             count++;
8         }
9     }
```

La funzione riceve in ingresso un array di tre frequenze, dove la prima posizione è riservata alla **frequenza dello zero**, la seconda alla **frequenza della portante** questa

serve ad identificare il canale tra (master, slave ed config) più che alla creazione del Signal Code, e la terza alla **frequenza dell'uno**. Il contatore **count** tiene traccia del numero di toni attivi: se è diverso da due, oppure i toni ricevuti sono quelli esterni, quindi senza tono che indica il canale, allora il simbolo non è valido.

```

10     if (count == 2) {
11         if (active[0] && active[1] && !active[2]){
12             if(freqs[0] < MASTER_BASE + (TONE_STEP*19) && freqs[1]
13                 => == MASTER_BASE){
14                 return (freqs[0]-MASTER_BASE)/(TONE_STEP);
15             }
16             if(freqs[0] < SLAVE_BASE + (TONE_STEP*19) && freqs[1]
17                 => == SLAVE_CARRIER){
18                 return (freqs[0]-SLAVE_BASE)/(TONE_STEP);
19             }
20             if(freqs[0] < CONFIG_BASE + (TONE_STEP*19) && freqs[1]
21                 => == CONFIG_CARRIER){
22                 return (freqs[0]-CONFIG_BASE)/(TONE_STEP);
23             }
24         }
25     }

```

Se ci sono due frequenze attive, si controlla quale delle tre posizioni è spenta e si determina il ruolo: **master**, **slave** o **config**.

Il calcolo restituisce un il **Signal Code** che rappresenta **quanti zeri o uno consecutivi** sono codificati in quel simbolo come è stato illustrato nella Tabella 3.3.

```

23     if (!active[0] && active[1] && active[2]){
24         if(freqs[2] < MASTER_BASE + (TONE_STEP*19) && freqs[1]
25             => == MASTER_BASE){
26             return (freqs[2]-MASTER_BASE)/(TONE_STEP);
27         }
28         if(freqs[2] < SLAVE_BASE + (TONE_STEP*19) && freqs[1]
29             => == SLAVE_CARRIER){
30             return (freqs[2]-SLAVE_BASE)/(TONE_STEP);
31         }
32         if(freqs[2] < CONFIG_BASE + (TONE_STEP*19) && freqs[1]
33             => == CONFIG_CARRIER){
34             return (freqs[2]-CONFIG_BASE)/(TONE_STEP);
35         }
36     }

```

```

31         return (freqs[2]-CONFIG_BASE)/(TONE_STEP);
32     }
33     }
34     return -2;
35 } else if (count == 3) {
36     return -2;
37 } else {
38     return -3;
39 }
40 }
```

I valori negativi gestiscono gli errori: **-2** significa che la combinazione di frequenze non è valida, mentre **-3** indica che non ci sono abbastanza toni per interpretare un simbolo.

Infine, la conversione inversa è svolta da **frequency\_coder**, che genera le coppie di frequenze a partire da un **Signal Code** e dal ruolo del canale.

## 5.7 Ricevimento e trasmissione dei bit (Packers)

### 5.7.1 Ricezione dei bit

Terminata la decodifica, i **Signal Code** vengono accumulati in buffer a sette bit dal modulo **bit\_input\_packer**.

Un aspetto sofisticato di questo componente è la gestione del flush: questo aspetta l'arrivo di un Signal Code 8 che rappresenta 21 zero consecutivi, ma è condizionato da un timeout, utile quando i toni cessano di arrivare ma il Code 8 non è arrivato, a causa di qualche interferenza.

L'algoritmo controlla costantemente l'istante dell'ultimo bit per ogni canale, e se per più di un secondo non sopraggiungono nuovi simboli effettua la conversione in ASCII. Questa ne è l'implementazione:

```

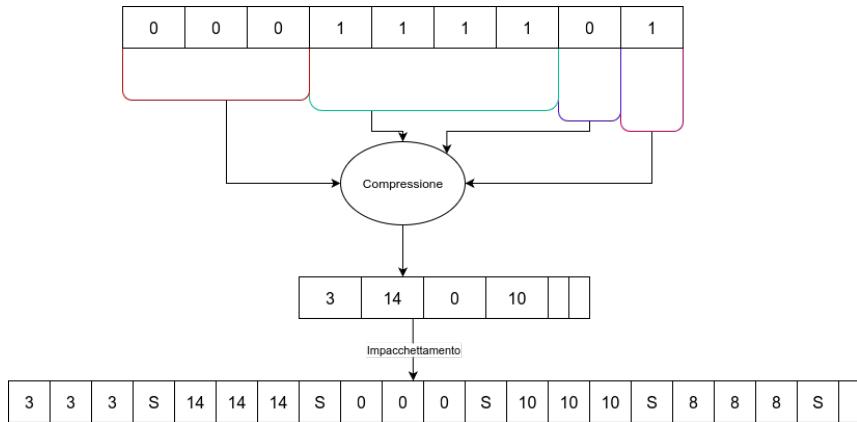
1 static bool timeout_flush_if_needed(BitPacker* packer,
2                                     const char* label,
3                                     bool* timeout_armed,
4                                     uint64_t last_bit_ms,
5                                     bool no_new_bit_this_tick){
6     if (!*timeout_armed) return false;
7     if (!no_new_bit_this_tick) return false;
8     uint64_t tnow = now_ms();
9     if ((tnow - last_bit_ms) < TIMEOUT_MS) return false;
10    bool ok = flush_and_convert_to_ascii(packer, label);
11    *timeout_armed = false;
12    return ok;
13 }
```

Quando il flush viene eseguito, la routine **flush\_and\_convert\_to\_ascii** trasferisce il contenuto della struttura in array di caratteri, inoltre ne verifica la qualità del pacchetto accertandosi che non ci siano caratteri non stampabili.

La funzione ritorna **true** se il pacchetto è valido, **false** altrimenti.

### 5.7.2 Trasmissione dei bit

L'inversa della fase di ricevimento dei bit, è la fase di trasmissione dei bit, ovvero l'impacchettamento dei bit in uscita in **Signal Code**, è svolto dal modulo **bit\_output\_packer**, che comprime i bit in **Signal Code** applicando una semplice forma di **run-length** encoding e replicando ogni codice più volte per incrementare la robustezza in presenza di rumore, così che venga trasmesso dal livello fisico per 3 volte consecutivamente, come illustrato in figura Figura 3.7 ogni codice verrà trasmesso per un tempo di 24ms, quindi ritrasmettendo per 3 volte ogni codice, si avrà una finestra di ascolto di 72ms per ogni codice.



**Figura 5.2.** Compressione ed impacchettamento mediante Algoritmo Run-Length S= Silenzio

Nel seguente frammento è mostrata la parte conclusiva di **bit\_output\_packer\_convert** che aggiunge il Code terminatore 8 (21 zeri) ed il silenzio finale che verrà interpretato come un silenzio di 80ms dal modulo **emit\_tones**:

```

1  for(int b = 6; b >= 0; --b){
2      printf(" 8 ");
3      packer->pairs[packer->pair_count++] = frequency_coder(8, role);
4  }
5  packer->pairs[packer->pair_count++] = silent;

```

## 5.8 Instradamento e protocollo

Il livello superiore, il Livello Link, del protocollo è rappresentato da **protocol.c**, dove vengono gestite l’assegnazione degli identificativi, l’emissione di comandi e la ricezione delle risposte.

I messaggi, incapsulati con la sintassi **ID:{dest}{OP}k{src}**, come discusso nella sezione 3.4 ed nella sottosezione 3.4.1 vengono convogliati all’uscita del canale appropriato tramite **send\_master**, **send\_slave** o **send\_config**.

Ogni funzione di trasmissione sfrutta **bit\_output\_packer** ed **emit\_tones**, ma si differenzia per la portante adottata.

Il frammento seguente illustra la procedura di invio sul canale master:

```
1 static unsigned long send_master(const char *msg){  
2     log_send(msg);  
3     BitOutputPacker packer;  
4     bit_output_packer_init(&packer);  
5     unsigned long duration = 0;  
6     if(bit_output_packer_compress(&packer, msg)){  
7         if(bit_output_packer_convert(&packer, 0)){  
8             duration = estimate_send_duration(&packer);  
9             emit_tones(packer.pairs, packer.pair_count);  
10        }  
11    }  
12    bit_output_packer_free(&packer);  
13    return duration;  
14 }
```

La funzione calcola anche il tempo necessario alla trasmissione attraverso **estimate\_send\_duration**, valore impiegato per ritardare eventuali retransmission in caso di mancato ACK. All'interno della routine **protocol\_handle\_message** sono implementate le sequenze di registrazione, il meccanismo del token di trasmissione e la mappatura dei comandi provenienti dall'interfaccia esterna.

Quando un nodo privo di identificativo invia una richiesta {REQ}l{xxx}, l'hotspot genera un ID univoco e risponde con il messaggio **ID:idAssegnato{SET}k{0000}**; il nodo memorizza l'ID non volatile e conferma con {OK}. Le stesse routine gestiscono anche le operazioni **ABORT** e **MOVEMENT**, delegando al modulo **movement\_sensor** l'effettiva lettura del sensore PIR.

## 5.9 Sensore di movimento

Infine per mostrare come il protocollo possa essere esteso con funzionalità pratiche, è stato sviluppato il modulo **movement\_sensor**, che non appartiene al protocollo in senso stretto, ma è la prima applicazione del **Livello Applicativo**. La routine **movement\_sensor\_detect** attiva il sensore PIR per un intervallo prefissato e restituisce **true** o **false** a seconda della rilevazione. Alla ricezione del comando **MOVEMENT:ON\_xxx**, il protocollo esegue la chiamata e risponde con **MOVEMENT:YES** o **MOVEMENT:NO**.

# 6. Sperimentazione del Protocollo

Il collaudo finale del protocollo è stato condotto con due nodi basati su ESP32-WROOM-32, uno equipaggiato con l’altoparlante pilotato dal MAX98357A e l’altro con il microfono MEMS SPH0645LM4H descritti in [Capitolo 4](#); i dispositivi sono stati alimentati da un’unica linea regolata e poggiati distanziati di 20cm per evitare riflessioni ravvicinate, mentre il firmware in [Capitolo 5](#) è stato compilato con gli stessi parametri su entrambe le schede. Nell’arco di una settimana sono stati effettuati tre blocchi di prove serali ognuna da circa 2 ore, rispettivamente da 180, 220 e 160 trasmissioni consecutive, per un totale di 560 pacchetti analizzati: ogni campagna prevedeva l’avvio automatico della sequenza di configurazione, la pubblicazione del comando "movement\_sensor\_on" oppure "movement\_sensor\_on\_" susseguito da un tempo casuale in millesimi di secondo; ogni richiesta quindi prevedeva l’impiego durante la fase di configurazione ma anche durante l’emissione del comando, l’utilizzo di un ID sequenziale, quindi diverso ad ogni invio, nel range tra 0000 e 0560, e infine l’invio di un messaggio telemetrico che riproduceva un timestamp decimale. Per ogni scambio il timer integrato nel processore dell’ESP32 ha registrato la finestra di trasmissione e i log generati da **estimate\_send\_duration**, successivamente in automatico uno **script Python** ha ricalcolato convertendo le stringhe ASCII in caratteri a sette bit e comprimendole con la tabella run-length di [Tabella 3.3](#), verificando che il calcolo corrispondesse.

Inoltre il software di ricezione ha salvato anche il conteggio delle ritrasmissioni a seguito di una trasmissione non confermata.

Scenario	Prove	Signal code medi	Tempo medio [ms]	Bitrate utile [bps]	Ritrasmissioni [%]
Configurazione (REQ/SET/OK)	180	171	26232	14.7	0.9
Movement sensor ON	220	239	36568	14.0	0.6
Telemetria ASCII “1738450195”	160	155	23720	14.5	0.8

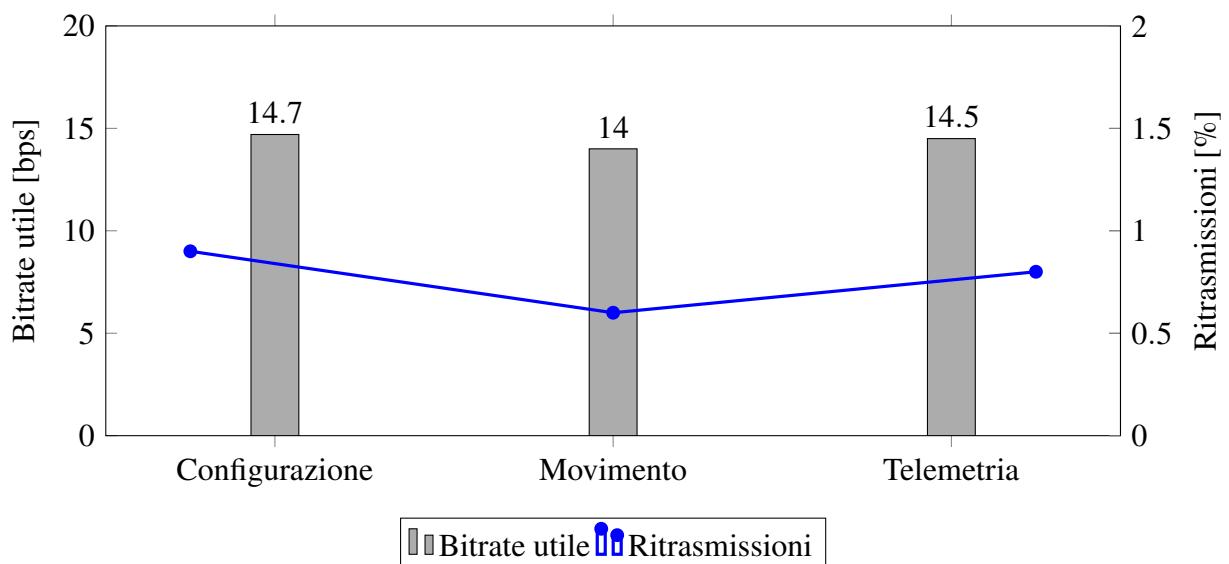
**Tabella 6.1.** Sintesi numerica delle tre campagne sperimentali condotte sul protocollo audio

I dati aggregati in [Tabella 6.1](#) mostrano che la registrazione (sequenza **REQ/SET/OK** del Livello Link descritta in [sezione 3.4](#)) produce in media 171 signal code per scambio, pari a 385 bit utili trasmessi in 26.232 s con un bitrate che si attesta sui  $14.7 \text{ bits}^{-1}$ .

Il comando "movement\_sensor\_on" codificato come **ID:0471{MOVEMENT\_ON}k{0000}** e la successiva risposta **ID:0000{MOVEMENT\_YES}k{0471}** totalizzano 239 signal code, equivalenti a  $36.568\text{ s}$  di trasmissione e a un'efficienza di  $14.0\text{ bits s}^{-1}$ ; la variante con parametro temporale **MOVEMENT\_ON\_12000** ha generato in laboratorio 107 signal code solo per il comando, confermando quanto l'assenza di coppie ripetute nelle cifre ASCII penalizzi la compressione run-length descritta in Capitolo 5.

Anche la trama telemetrica "1738450195", inviata con il prefisso **ID:0471{DATA}k{0000}**, si ferma a  $14.5\text{ bits s}^{-1}$  perché la distribuzione semi-alternata dei bit obbliga il sezione 3.4 a cambiare segno quasi ogni due posizioni, **riducendo l'ampiezza media dei run a poco più di due bit**.

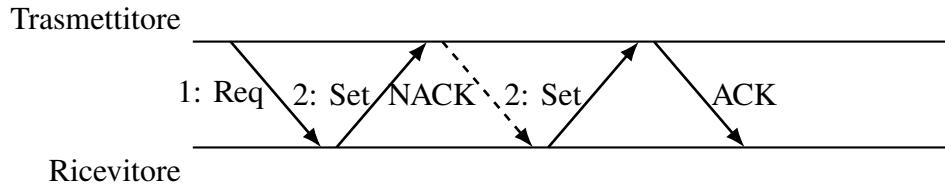
La media complessiva del bitrate utile su tutte le prove si attesta a  $14.4\text{ bits s}^{-1}$ , in linea con quanto affermato nella sezione 3.3.



**Figura 6.1.** Trend del bitrate utile e della percentuale di ritrasmissioni misurati nelle tre casistiche di prova

La bassa percentuale di ritrasmissioni riportata nel grafico deriva da un meccanismo di conferma continua che obbliga il nodo ricevente a replicare **l'acknowledgement** sul canale slave o master entro 100 ms dalla decodifica del frame; lo schema temporale di Figura 6.2 mostra come l'unico NACK (nessun ACK) registrato durante la sequenza di configurazione venga seguito dal reinvio di un singolo pacchetto e dalla successiva conferma positiva, penalizzando il tempo complessivo della macro-operazione.

Anche nei cicli telemetrici più gravosi la logica di ritrasmissione non ha richiesto più di un tentativo aggiuntivo e, grazie alla persistenza di fase nel generatore audio di Capitolo 5, l'aggancio dello spettro è rimasto stabile.



**Figura 6.2.** Sequenza temporale semplificata con l'unica ritrasmissione registrata durante la fase di configurazione

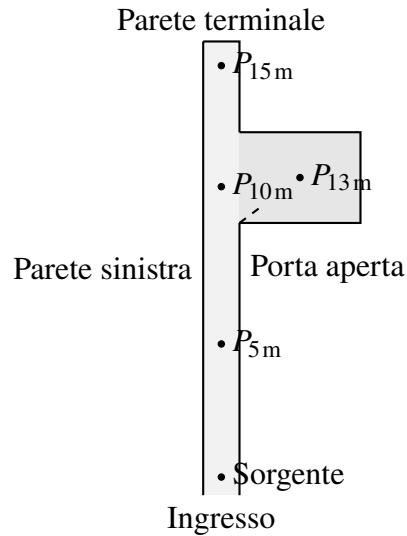
## 7. Sperimentazione della propagazione del suono

Il percorso metodologico ideato dalla U.S. National Institute for Occupational Safety and Health's[Jacksha and Zhou, 2016, sec. 2] nel [Capitolo 2](#) è stato usato come riferimento per la campagna di misurazioni svolta in un corridoio di servizio situato a circa cinque metri sotto terra, largo 1.2 m, lungo 15 m e chiuso da una parete cieca, con una porta sul lato destro posta a 10 m dall'origine che conduce nella camera laterale raffigurata nella Figura 7.1; la deviazione a novanta gradi corrisponde all'accesso a questa stanza, mentre il tunnel principale resta rettilineo fino al muro finale posto a 15 metri dall'origine.

La prova ha impiegato il dispositivo sperimentale sviluppato nella tesi, basato sulla piattaforma ESP32, programmato con il protocollo definito nella parte di implementazione software [Capitolo 5](#).

Il protocollo in modalità "testing" ha riprodotto in sequenza tutte le frequenze previste dal protocollo, da 1000 Hz a 9000 Hz con passi di 400 Hz, mantenendo per ciascuna un livello in emissione di circa 70 dB<sub>SPL</sub> per via dell'amplificatore discusso nella [sottosottosezione 4.1.2.2](#), misurato all'origine con un fonometro.

Le misure di attenuazione sono state raccolte con il **Fonometro Tadeto**, posato accanto alla scheda ricevente equipaggiata con l'implementazione software di decodifica, così da correlare in tempo reale la variazione del livello sonoro in 35 dB<sub>SPL</sub> con la percentuale di pacchetti persi al crescere della distanza; i valori registrati alimentano le curve di Figura 7.2.



**Figura 7.1.** Planimetria del corridoio sotterraneo. La porta sul lato destro è stata mantenuta aperta durante la prova.

All'ingresso del tunnel il segnale era di 70 dB<sub>SPL</sub>, mentre a 5 m dall'origine, il livello medio scendeva a 65 dB<sub>SPL</sub> e a 10 m si attestava su 62 dB<sub>SPL</sub>; il valore minimo registrato vicino alla parete finale collocata a 15 m è stato di 55 dB<sub>SPL</sub>, confermando una perdita complessiva di circa 15 dB.

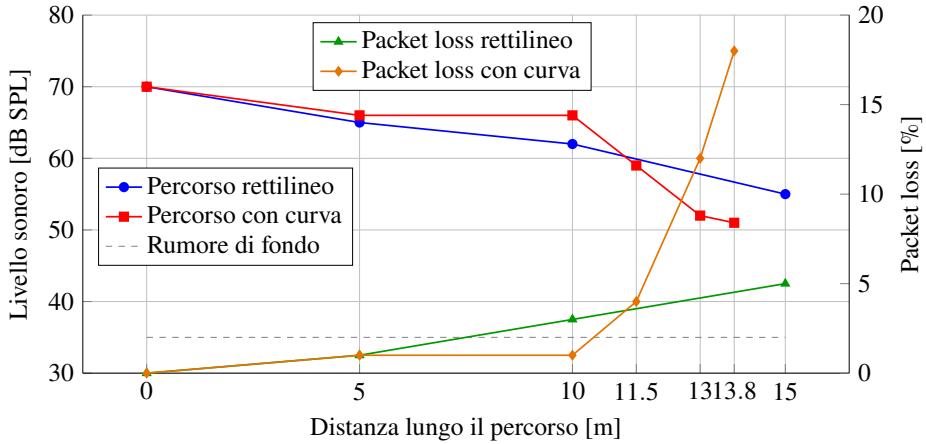
Attraversando la porta a destra, lasciata completamente aperta, è stata seguita la curva verso la camera larga 4.25 m e lunga 3.67 m.

A 1.5 m dall'ingresso la misura rimaneva su 59 dB<sub>SPL</sub>, ma collocandosi a 3 m lungo la direzione principale della stanza, verso destra quindi, si è osservato un crollo a 52 dB<sub>SPL</sub>, seguito da un ulteriore assestamento a 51 dB<sub>SPL</sub> vicino alla parete opposta.

Le componenti frequenziali tra 5 kHz e 7.4 kHz hanno mostrato attenuazioni leggermente più pronunciate rispetto alle frequenze basse, ma sempre inferiori a 2 dB di differenza rispetto al valore medio, tuttavia le bande portanti 1000Hz, 8600Hz ed 9000Hz sono state preservate.

Il rumore di fondo dell'ambiente sotterraneo è rimasto stabile su 35 dB<sub>SPL</sub>, avendo, quindi, un margine di 16 dB anche nel punto più critico dopo la curva e permettendo al ricevente di distinguere con chiarezza i pacchetti trasmessi.

L'ulteriore analisi, effettuata sulla **perdita di pacchetti**, confermano che il percorso rettilineo ha mantenuto un tasso di perdita inferiore al 5 % fino alla parete finale, mentre nella camera laterale l'incremento di attenuazione dopo i 3 m dalla curva ha fatto salire la packet loss fino a circa 18 %; le due serie sono riportate dalle linee verdi e arancioni di Figura 7.2.



**Figura 7.2.** Andamento medio dei livelli sonori misurati lungo il corridoio e corrispondente packet loss. La linea blu mostra la perdita quasi lineare nel tratto rettilineo, la linea rossa il crollo dopo la curva a novanta gradi, la linea tratteggiata il rumore di fondo e le curve verdi e arancioni riportano il tasso di pacchetti persi rilevato con il fonometro accanto alla scheda ricevente.

In una produzione reale a scopi industriali, come evidenziato nell’analisi del Capitolo 2, l’alimentazione a 3.3 V verrebbe sostituita da una intorno ai 220 V, con un incremento di livello pari a

$$20 \log_{10} \left( \frac{220}{3.3} \right) \approx 36.5 \text{ dB}. \quad (7.1)$$

L’emissione iniziale arriverebbe quindi a circa 106.5 dB<sub>SPL</sub> e, mantenendo invariata l’attenuazione osservata, porterebbe i livelli lungo il corridoio a 101.5 dB<sub>SPL</sub> dopo 5 m, 98.5 dB<sub>SPL</sub> in prossimità della porta e 91.5 dB<sub>SPL</sub> contro il muro; nel ramo laterale si stimerebbero 95 dB<sub>SPL</sub> tre metri oltre la curva, valori che resterebbero nettamente separati dal rumore di fondo.

# 8. Sviluppi futuri

Il progetto trattato potrà trovare nuovi sbocchi partendo dalle conclusioni raggiunte in Capitolo 1, dove si è visto come il suono riesca a infilarsi nei cunicoli in cui le onde radio fanno fatica a passare.

Pensando a miniere e gallerie, ma anche a scenari di emergenza, lo stesso impianto hardware e software può diventare un supporto per squadre di sicurezza che devono tenere sotto controllo vibrazioni anomale, presenza di gas o segnali vitali di operatori isolati in caso di disastri.

La rete acustica può guidare piccoli robot mobili controllati da remoto, lasciando ai nodi fissi il compito di fare da ripetitori tra un corridoio e l’altro e di aggiornare la sala operativa sulla posizione del robot, anche quando il tracciato cambia dopo un crollo.

## 8.1 Evoluzione degli scenari applicativi

Gli sviluppi immaginati per gli ambienti minerari sono il primo esempio concreto: schede come quella descritta in questa tesi alimentate a batteria e installate lungo le gallerie possono offrire un canale acustico di comunicazione per raggiungere nodi finali addibiti alla raccolta di dati ambientali (Sensori PIR o Sensori Radar Doppler) o al controllo di robot mobili.

Il paragrafo sezione 8.4 torna su questo punto perché la velocità diventa cruciale nelle operazioni di ricerca di personale disperso; in parallelo, il nodo master può comandare una squadra di robot grazie al protocollo master–slave già enunciato nel Capitolo 3, chiedendo loro di esplorare diramazioni pericolose e di tornare indietro in caso di crolli improvvisi.

L’idea si espande anche ad ambienti marini, adattando il hardware e dotandolo di trasduttori infrasonici ad alta sensibilità in grado gestire frequenze compresse tra 0 e 20 Hz, come per la versione "terrestre" ispirati alla natura, in particolare ai pappagalli, la versione sottomarina potrebbe sfruttare le onde ELF (Extremely Low Frequency, 3–30 Hz) per comunicare a grandi profondità, ispirandosi alle balenottere azzurre che usano queste frequenze per coprire distanze oceaniche [Dziak et al., 2017, sec. 2], questo permetterebbe la comunicazione con sommergibili o con sensori di profondità.

Studi specifici sulla propagazione in banda ELF confermano che segnali così lenti, se filtrati correttamente, riescono a raggiungere sommergibili che stazionano a grande profondità senza dover riemergere.

In maniera quasi speculare, in una situazione estrema, il protocollo potrebbe tornare utile come backup terrestre: nel caso di una tempesta solare capace di mandare fuori gioco i ponti radio, una dorsale acustica basata su condensatori (microfoni digitali a condensatori) e attuatori elettrodinamici, come quelli usati nei prototipi, potrebbe offrire un canale di comunicazione lento ma affidabile.

## 8.2 Revisione del formato di pacchetto

Nell'ipotesi di questi sviluppi futuri, allora si rende necessaria una modifica del protocollo descritto in particolare nella [sezione 3.4](#).

L'assetto attuale, che codifica l'intestazione come stringa

**ID:id\_destinatarioOPERAZIONEkid\_mittente**, ha delle criticità che rendono il protocollo più lento di quello che potrebbe diventare.

La riscrittura prevede di riservare un solo carattere ASCII di controllo, per esempio il simbolo "?", così da separare i campi del pacchetto senza l'utilizzo di parentesi graffe o scritte complesse come "ID:". In questo modo il ricevitore non deve più cercare pattern letterali ma può fermarsi alla prima occorrenza del delimitatore, impacchettando ogni blocco in una specifica categoria.

La scelta del carattere però deve essere ben ponderata, questo deve essere un carattere **con un alto tasso di compressione** utilizzando l'algoritmo di compressione run-length descritto nel [Capitolo 5](#), in modo da rappresentare un separatore con il minor numero di bit possibile.

Le modifiche sopra verebbero implementate nel "frame" attuale, mentre un ulteriore modifica al di fuori di quest'ultimo sarebbe la compressione **Huffman** del pacchetto binario così da ridurre ulteriormente la dimensione del messaggio.

La compressione di Huffman è un algoritmo di compressione senza perdita che assegna codici di lunghezza variabile ai simboli in base alla loro frequenza di occorrenza.

I simboli più frequenti ricevono codici più corti, mentre quelli meno frequenti ricevono codici più lunghi.

Questo permetterebbe di assegnare ai comandi più frequenti quali: delimitatori, ID di rete ed operazioni più frequenti o più importanti per l'azienda committitrice codici più corti, riducendo la dimensione complessiva del pacchetto.

Un esempio di questa evoluzione è mostrato in Figura 8.1.



**Figura 8.1.** Roadmap di evoluzione del frame, dal formato attuale verso blocchi compatti e pronti alla compressione run-length attuale

La roadmap di Figura 8.1 riassume gli step principali: il lavoro sul delimitatore prepara il terreno a un lessico binario che potrà essere precompresso con Huffman prima di affrontare la compressione a livello link.

### 8.3 Sensoristica avanzata

Come già accennato in sezione 8.1, l'adozione di sensori più sofisticati permetterebbe di estendere il protocollo a scenari più complessi, come le comunicazioni sottomarine o il monitoraggio ambientale in miniere profonde.

Per la ricezione del segnale, si rende indispensabile sostituire i microfoni MEMS standard con microbarometri capacitivi.

In parallelo, un convertitore sigma-delta, attualmente limitato a 16 bit, potrebbe essere cambiato con un modello a 24 bit, in grado di campionare anche frequenze ultrabasse, rendendo il sistema a livello hardware in grado di operare in banda ELF, quindi implementarlo per usi sottomarini.

Dal lato della trasmissione, invece, sarebbe meglio adottare amplificatori di classe-D come quello attualmente utilizzato, ma con una potenza di uscita più elevata, in modo da poter pilotare trasduttori elettrodinamici più grandi e in grado di spostare più aria o acqua.

## 8.4 Riduzione dei tempi di segnalazione

Attualmente ogni signal code occupa 152 ms, somma dell'emissione ripetuta tre volte e della finestra di silenzio da 80 ms come mostrato nella sequenza di Figura 6.2.

Utilizzando l'elettronica discussa nella sezione precedente sezione 8.3 potrebbe essere possibile, allora, la riduzione del tempo di emissione per ogni signal code, presumibilmente, nel caso migliore emettendolo solo per 10.7 ms (tempo di campionamento discusso nella sezione 3.2) susseguito da un silenzio di 10.7 ms per un totale di 21.4 ms.

Questo porterebbe diversi vantaggi, oltre ad un throughput migliore, che si attesterebbe intorno a 47.6 signal codes per secondo, che con una compressione run-length di circa 4:1, come discusso nella sezione 3.4, porterebbe a circa 190 bits per secondo.

## 9. Conclusioni

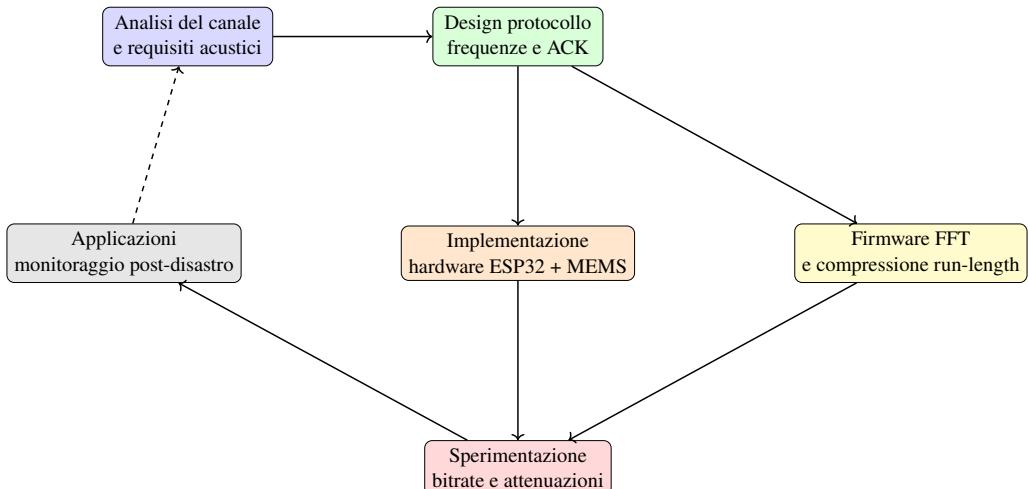
Il percorso intrapreso ha permesso di verificare che la scelta di un canale acustico per collegare nodi master–slave in ambienti sotterranei è realistica quando le onde elettromagnetiche risultano inaffidabili.

La progettazione è partita dall’analisi delle condizioni di propagazione e dei vincoli strutturali descritti nei capitoli iniziali [Capitolo 1](#), [Capitolo 2](#), traducendosi in un protocollo capace di sfruttare coppie di frequenze dedicate e schemi di conferma rapidi gestiti da handles ad ACK come descritto nella [sezione 3.4](#).

Le scelte progettuali sono state costantemente verificate con prove sperimentali diverse, sono state effettuate prove di emissione e verifica della stabilità del suono al fine di evitare fruscii come evidenziato nella [sottosezione 3.2.2](#), poi si è passati alla verifica del riconoscimento delle frequenze, che ha richiesto 4 mesi di lavoro, con l’ausilio di un trasduttore piezoelettrico collegato ad un amplificatore Classe-D da 230 V a sua volta collegato all’uscita AUX di un computer; si è poi proceduto alla verifica del protocollo completo in ambienti controllati e infine in un corridoio sotterraneo di 15 metri con curve a 90 gradi con l’ausilio di un fonometro.

La parte hardware e software si è evoluta insieme, con il microcontrollore ESP32 che ha offerto abbastanza margine per campionare a 48 kHz, eseguire l’FFT a 512 campioni e mantenere in tempo reale il riconoscimento delle frequenze utili, più la derivante elaborazione del pacchetto dai.

L’implementazione hardware discussa nel [Capitolo 4](#) ha permesso di realizzare un nodo economico, con un costo totale inferiore a 20 euro per unità, che integra un microfono MEMS e un DAC I2S con amplificatore di potenza in Classe-D, il tutto alimentato a 3.3 V con un consumo di picco inferiore a 200 mA durante la trasmissione, risultando perfettamente in linea con quanto previsto nei requisiti iniziali ([sezione 1.5](#)).



**Figura 9.1.** Flusso progettuale e sperimentale della tesi.

Le prove in ambiente controllato e nel corridoio sotterraneo hanno confermato che, con un'emissione iniziale di circa 70 dB<sub>SPL</sub>, la perdita lungo quindici metri resta nell'ordine di 15 dB e la differenza tra percorso rettilineo e curva a novanta gradi si traduce in un aumento della packet loss dall'1% al 18% dopo i 3 m dall'imbocco laterale come è mostrato nella Figura 7.2.

Il protocollo ha mantenuto un bitrate utile medio di  $14.4 \text{ bits}^{-1}$  con una percentuale di ritrasmissioni sotto l'1%, valori che per un sistema ancora in stato embrionale, creato con pochissimi fondi e materiale di recupero, alimentato a 3.3 V risultano incoraggianti.

Le registrazioni di sequenze REQ/SET/OK e le trasmissioni telemetriche hanno mostrato come l'auto-taratura della soglia spettrale riesca a seguire i cambiamenti del rumore, evitando falsi positivi anche quando l'ambiente risultava particolarmente rumoroso, con voci di sottofondo o rumore di dispositivi elettronici accesi.

Una dimostrazione pratica del funzionamento del protocollo è disponibile nel video presente nell' Appendice A.

Il sistema ha ancora alcuni limiti, soprattutto legati alla velocità di trasmissione e al fatto che il segnale si indebolisce dopo curve strette, dove l'attenuazione diventa più marcata come osservato nel Capitolo 7.

Nonostante questo, il progetto non è affatto chiuso: anzi, ci sono margini di miglioramento ben definiti citati nel Capitolo 8, che potrebbero portare ad incrementi significativi delle prestazioni, in entrambi gli aspetti sopra.

**In conclusione** Il lavoro ha dimostrato che un protocollo **acustico low cost** può funzionare in contesti critici, raggiungendo l’obiettivo principale, superando in alcuni contesti pure certi standard di industria osservati nel [Capitolo 2](#).

Nonostante i limiti attuali, i risultati ottenuti rappresentano una base per sviluppare versioni più robuste, rapide ed adatte ad ambienti diversi da quelli trattati in questa tesi, questa era volta più ad aprire una strada verso queste applicazioni che a fornire una soluzione definitiva.

# Ringraziamenti

Alla fine di questo elaborato ho il piacere di ringraziare tutti quelli che hanno contribuito in questi anni di carriera universitaria.

In primis, ringrazio la mia famiglia: mia madre, mio padre, mia sorella e i miei nonni; voi mi avete incoraggiato e avete creduto in me, anche quando io stesso avevo dei dubbi. In ogni momento difficile, in cui qualsiasi altra persona mi avrebbe abbandonato, voi ci siete stati, lì pronti per me. Quando dopo le superiori volevo mollare e ho fatto sbagli uno dietro l'altro, quando durante l'università volevo arrendermi, voi mi avete sempre sostenuto, in modo discreto, con quella classica “spinta” che non ti accorgi di avere ma che ti fa andare avanti.

Mi sono ripromesso che non farò più errori e che andrò avanti fino alla fine, per voi. Grazie.

Un ulteriore ringraziamento va ai miei cugini che mi hanno sempre appoggiato, soprattutto in questi ultimi anni: le chiacchierate pomeridiane parlando dei nostri esami e le giornate di studio intenso insieme rimarranno sempre scolpite nella mente.

特別なありがとうはしーちゃんに。知り合ってからそんなに時間は経っていないし、一緒に過ごした時間も多くなかったけど、Campus Hubや図書館、うちに一緒に勉強した午後はずっと心に残ると思う。

君と出会ったこと、そして一緒に未来を想像することが、前に進む理由をくれた。本当にありがとう。

Ai miei amici di Catania, che ci sono sempre stati nello sconforto e nella gioia, che mi hanno sempre appoggiato e incoraggiato, grazie.

Negli anni gli amici si incontrano e si perdono, a noi è andata così.

Ma alla fine l'importante è chi rimane, e voi siete rimasti. Grazie per le belle serate, le francescane bevute insieme dopo ogni esame e le risate.

Ai miei amici di Malta, che mi hanno fatto passare un anno indimenticabile: siamo stati una vera famiglia, in un'isola che ci ha lasciato ricordi indelebili.

Ci siamo divertiti e abbiamo studiato pomeriggi interi insieme, ci siamo aiutati a vicenda e siete stati una seconda famiglia quando ero lontano da casa, grazie.

Ulteriori ringraziamenti sono rivolti a quei docenti che in questi anni hanno saputo trasmettere con passione, dedizione e preparazione l'amore che essi stessi nutrono per le loro materie. Grazie al loro impegno e alla loro capacità di coinvolgere hanno creato ambienti stimolanti, capaci di lasciare a tutti gli studenti, me compreso, amore e passione per gli insegnamenti svolti.

Infine, un ringraziamento va al mio relatore, che mi ha sostenuto in ogni fase del lavoro, lasciandomi la libertà di seguire le mie idee e incoraggiandomi a svilupparle senza paura. La sua guida non è mai stata un vincolo, ma un invito a crescere, a migliorarmi, a credere nelle mie capacità.

E quando, alla fine, mi ha proposto la pubblicazione e la partecipazione a una conferenza, che ha dato ancora più valore a questo percorso.

# Elenco delle figure

1.1	Corner Loss ~30 dB in un condotto con angolo di 90° per radio frequenze HF/SHF [Jacksha and Zhou, 2016].....	3
2.1	Mappa Tunnel U.S. National Institute for Occupational Safety and Health's[Jacksha and Zhou, 2016, sec. 3]. .....	8
3.1	Livelli del Protocollo.....	11
3.2	Verifica della cattura del tono a 9kHz con campionamento a 48kHz nel periodo 0-0.010 secondi, utilizzando finestra di Hann su 512 elementi .....	12
3.3	Grafico contenente lo spettro di frequenza calcolato tramite FFT .....	13
3.4	Grafico contenente lo spettro di frequenza di tre frequenze (bassa, media, alta), utilizzato per identificare l'attenuazione in frequenza del microfono...	16
3.5	Picco Locale nell'intorno dei sei Bins .....	16
3.6	Interpolazione Parabolica [Smith and Serra, 2020] .....	17
3.7	Finestra d'ascolto della FFT sul segnale emesso .....	18
3.8	Schema del flusso dei pacchetti tra FIFO Input, Livello Link e FIFO Output	23
3.9	Schema del flusso dei Token tra hotspot e nodi non-hotspot .....	24
3.10	Utilizzo del Livello Applicazione tramite l'interfaccia Blynk.io .....	27
4.1	Componenti hardware principali di un microcontrollore ESP32.....	29
4.2	Microfono MEMS I2S digitale Adafruit SPH0645LM4H-B.....	30
4.3	Sezione verticale del microfono Adafruit MEMS. .....	32
4.4	Amplificatore digitale Classe-D Adafruit MAX98357A. ....	33
4.5	Funzionamento interno del MAX98357A.....	34
4.6	Collegamento fisico tra ESP32 e microfono I <sup>2</sup> S (Adafruit SPH0645LM4H) su breadboard. ....	35
4.7	Collegamento fisico tra ESP32 e amplificatore I <sup>2</sup> S (Adafruit MAX98357A) su breadboard. ....	37
4.8	Collegamento fisico tra ESP32 e LED di segnalazione su breadboard. ....	38
4.9	Schema di collegamento del bottone e dei LED di stato per la modalità hotspot. ....	40

4.10	Schema del circuito completo con ESP32, microfono I <sup>2</sup> S, amplificatore MAX98357A, LED di stato e pulsante di hotspot. ....	41
4.11	Layout del circuito stampato (PCB), con ESP32, LED di stato, pulsante di hotspot e connettori di alimentazione. ....	42
5.1	Swapping degli array di campioni .....	47
5.2	Compressione ed impacchettamento mediante Algoritmo Run-Length S= Silenzio .....	55
6.1	Trend del bitrate utile e della percentuale di ritrasmissioni misurati nelle tre casistiche di prova .....	58
6.2	Sequenza temporale semplificata con l'unica ritrasmissione registrata durante la fase di configurazione .....	59
7.1	Planimetria del corridoio sotterraneo. La porta sul lato destro è stata mantenuta aperta durante la prova. ....	61
7.2	Andamento medio dei livelli sonori misurati lungo il corridoio e corrispondente packet loss. La linea blu mostra la perdita quasi lineare nel tratto rettilineo, la linea rossa il crollo dopo la curva a novanta gradi, la linea tratteggiata il rumore di fondo e le curve verdi e arancioni riportano il tasso di pacchetti persi rilevato con il fonometro accanto alla scheda ricevente. ....	62
8.1	Roadmap di evoluzione del frame, dal formato attuale verso blocchi compatti e pronti alla compressione run-length attuale .....	65
9.1	Flusso progettuale e sperimentale della tesi.....	68

# Elenco delle tabelle

3.1	Frequenze e tipi di segnale.....	15
3.2	Struttura di comunicazione tra Livello Bit e Livello Fisico .....	19
3.3	Mappatura frequenze, codici e portanti. ....	20
3.4	Indirizzi di rete del Livello Link .....	25
3.5	Mappatura comandi Blynk, operazioni interne e azioni eseguite dai nodi ...	26
4.1	Mappatura hardware dei pin I <sup>2</sup> S tra ESP32 e microfono SPH0645LM4H. ...	36
4.2	Mappatura hardware dei pin I <sup>2</sup> S tra ESP32 e amplificatore MAX98357A... ...	37
4.3	Mappatura hardware dei pin di uscita dell'ESP32 verso i LED di stato. ....	39
5.1	Mappatura frequenze con divisione dei Signal Code in 0 e 1 .....	51
6.1	Sintesi numerica delle tre campagne sperimentali condotte sul protocollo audio .....	57

## **A. Repository del codice e risorse multimediali**

Il codice sorgente sviluppato nell’ambito di questa tesi è reso disponibile nella repository GitHub [github.com/G103L3/Project01Giunta](https://github.com/G103L3/Project01Giunta). Tale archivio contiene tutte le componenti software realizzate, organizzate in modo da consentire la riproducibilità dei risultati e favorire eventuali estensioni future del progetto.

Durante le fasi di sviluppo è stata adottata una metodologia di lavoro basata su sprint iterativi, sfruttando la metodologia Agile.

A supporto della documentazione tecnica è stato inoltre realizzato un video dimostrativo del funzionamento del sistema, consultabile al seguente indirizzo: [youtu.be/F4CkQpI0kls](https://youtu.be/F4CkQpI0kls). Il filmato illustra le principali funzionalità offerte.

# Bibliografia

Sph0645lm4h mems microphone datasheet, 2015. URL <https://cdn-shop.adafruit.com/product-files/3421/i2S+Datasheet.PDF>. Adafruit product files.

Q&a: Using sph0645lm4h with nordic socs (i2s timing and osr), 2019. URL <https://devzone.nordicsemi.com/f/nordic-q-a/65863/i2s-clocks-and-softdevice>.

Esp32 technical reference misc, 2020. URL [https://www.espressif.com/sites/default/files/documentation/esp32\\_datasheet\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf).

Ian F. Akyildiz and Erich P. Stuntebeck. Wireless underground sensor networks: Research challenges. *Ad Hoc Networks*, 4(6):669–686, 2006. doi: 10.1016/j.adhoc.2006.03.004. URL <https://doi.org/10.1016/j.adhoc.2006.03.004>. Paragrafo: Sec. 2 (Underground Channel Modeling).

Md Abir Al Moshi et al. Wireless underground sensor communication using acoustic technology: A review. *Sensors*, 2024. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC11125261/>. Open access review on acoustic WUSNs.

Simone Bianchi. Il suono in uno spazio chiuso, 2023. URL [https://www.simonebianchi.net/wp-content/uploads/2023/01/acustica\\_ambient.pdf](https://www.simonebianchi.net/wp-content/uploads/2023/01/acustica_ambient.pdf). Paragrafo “Teoria modale”, p. 2; adattamento della tesi di laurea di T. Giunti.

James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19(90):297–301, 1965. doi: 10.1090/S0025-5718-1965-0178586-1. URL <https://www.ams.org/mcom/1965-19-090/S0025-5718-1965-0178586-1/S0025-5718-1965-0178586-1.pdf>.

Robert P. Dziak, Julia H. Haxel, T.-K. Lau, S. Heimlich, Jacqueline Caplan-Auerbach, D. K. Mellinger, H. Matsumoto, and Bruce Mate. A pulsed-air model of blue whale b call vocalizations. *Scientific Reports*, 7(1):9122, 2017. doi: 10.1038/s41598-017-09423-7. URL <https://www.nature.com/articles/s41598-017-09423-7>.

O. O. Farai, J. M. Muggleton, E. Rustighi, and M. J. Brennan. Analysis of acoustic signal propagation for reliable digital communication along mdpe pipes. *Applied Sciences*, 13(15):8886, 2023. doi: 10.3390/app13158886. URL <https://doi.org/10.3390/app13158886>.

Marenglen Fishta, Giorgia Franceschinis, Andrea Braglia, Alj, S. Bagheri, and Cristina Seceleanu. From radio to in-pipe acoustic communication for smart water networks. *Information*, 14(8):435, 2023. doi: 10.3390/info14080435. URL <https://doi.org/10.3390/info14080435>.

Alexander Heifetz and et al. Transmission of information by acoustic communication along metal pathways. Technical Report ANL-17/13, Argonne National Laboratory, 2017. URL <https://doi.org/10.2172/1353103>.

Adafruit Industries. Adafruit max98357 i2s class-d mono amp guide, 2016a. URL <https://www.analog.com/media/en/technical-documentation/data-sheets/max98357a-max98357b.pdf>. Guide / Datasheet PDF.

Adafruit Industries. Adafruit max98357 i2s class-d mono amp guide, 2016b. URL [https://www.ije.ir/article\\_71856\\_e8a71864557e41de20df5091667.pdf](https://www.ije.ir/article_71856_e8a71864557e41de20df5091667.pdf). Guide / Datasheet PDF.

R. Jacksha and C. Zhou. Measurement of rf propagation around corners in underground mines and tunnels. *Transactions of the Society for Mining, Metallurgy, and Exploration*, 340(1):30–37, 2016. doi: 10.19150/trans.7324. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC5472338/>.

Jeho Lee. Radio wave propagation in tunnels. Technical Report ADA379290, Polytechnic University, Brooklyn, NY, Brooklyn, NY, 2000. URL <https://apps.dtic.mil/sti/pdfs/ADA379290.pdf>.

Davide Patti. Introduction to microcontrollers – communication interfaces, 2024.

A. Salam et al. Taking wireless underground: A comprehensive summary. *ACM Computing Surveys*, 2023. doi: 10.1145/3587934. URL <https://doi.org/10.1145/3587934>.

Claude E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949. doi: 10.1109/JRPROC.1949.232969.

J.O. Smith and X. Serra. Quadratic interpolation of spectral peaks. misc, CCRMA, Stanford University, 2020. URL [https://ccrma.stanford.edu/~jos/sasp/Quadratic\\_Interpolation\\_Spectral\\_Peaks.html](https://ccrma.stanford.edu/~jos/sasp/Quadratic_Interpolation_Spectral_Peaks.html).

S. Yang, A. Kalantari, S. Markham, J. Saniie, and F. Kong. Development of an underground through-soil wireless acoustic communication system. Technical Report, University of Illinois Chicago, 2020. URL <https://indico.cern.ch/event/923808/contributions/3896723/>.

Eberhard Zwicker and Hugo Fastl. *Psychoacoustics: Facts and Models*. Springer, 2nd edition, 1999. doi: 10.1007/978-3-662-09562-1. URL <https://link.springer.com/book/10.1007/978-3-540-68888-4>.