



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACOLTÀ DI INGEGNERIA
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ELETTRONICA

**Studio e modellazione del sistema di
diffusione ad altoparlanti rotanti "Leslie"
Study and modeling of the "Leslie"
rotating speakers diffusion system**

Candidato:
Elia Bonci

Relatore:
Prof. Leonardo Gabrielli

Anno Accademico 2021-2022



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

FACOLTÀ DI INGEGNERIA
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ELETTRONICA

**Studio e modellazione del sistema di
diffusione ad altoparlanti rotanti "Leslie"
Study and modeling of the "Leslie"
rotating speakers diffusion system**

Candidato:
Elia Bonci

Relatore:
Prof. Leonardo Gabrielli

Anno Accademico 2021-2022

UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ELETTRONICA
Via Brezze Bianche – 60131 Ancona (AN), Italy

A mio padre

Ringraziamenti

Desidero ringraziare il Prof. Leonardo Gabrielli, relatore di questa tesi, per avermi seguito con pazienza e disponibilità durante tutto il percorso di stesura dell'elaborato, per gli indispensabili consigli e per le conoscenze trasmesse.

Ringrazio poi il personale della Viscount International, in cui ho svolto il tirocinio, per avermi fornito un Leslie su cui effettuare i miei test ed aiutato nelle misurazioni.

Grazie infine a tutta la mia famiglia per il supporto ricevuto in questi anni di studio; in particolare, ringrazio mia madre per avermi dato la possibilità di intraprendere il percorso universitario e incoraggiato a portarlo a termine.

Cingoli, Settembre 2022

Elia Bonci

Sommario

Questa tesi tratta lo sviluppo di un'emulazione digitale del sistema di diffusione Leslie. Il suono caratteristico del Leslie è dato dai suoi altoparlanti rotanti, che applicano una modulazione in ampiezza e in frequenza al segnale. Inoltre, per via delle riflessioni del suono che si generano sulle pareti del cabinet in legno, all'interno del quale gli altoparlanti sono contenuti, si crea una modulazione della risposta impulsiva.

Il modello del Leslie proposto in questo lavoro è costruito sulla base dei segnali rilevanti dal punto di vista percettivo di una sorgente in movimento, ovvero le variazioni di intensità, le variazioni di timbro e l'effetto Doppler, e sfrutta il teorema delle immagini per simulare le riflessioni del suono. Si vedrà in modo dettagliato come l'effetto Doppler può essere implementato con una delay line a lunghezza variabile; in particolare, è stata svolta un'analisi che mostra le differenze implementative tra l'utilizzo della lettura interpolata e della scrittura interpolata della delay line.

Il Leslie prevede tre modi di funzionamento con velocità di rotazione differenti degli altoparlanti. La loro dinamica durante i cambi di velocità è stata misurata e modellata utilizzando una funzione polinomiale.

Una parte importante del suono del Leslie è data dall'amplificatore valvolare. Sono state testate diverse tecniche per la sua modellazione: waveshaping; modello di Wiener-Hammerstein identificato con una procedura di ottimizzazione; modello di Wiener-Hammerstein identificato con un approccio basato sul deep learning, in cui i blocchi lineari sono costituiti da filtri IIR che vengono implementati come layer ricorrenti e allenati con la backpropagation.

Il modello è stato implementato come plug-in per il software di elaborazione audio VCV Rack e la sua accuratezza è stata valutata tramite test di ascolto e confronti con un vero Leslie.

Indice

1	Introduzione	1
1.1	Descrizione del Leslie	1
1.2	Motivazioni	7
1.3	Panoramica della tesi	7
2	Modello di riferimento	8
2.1	I segnali percettivi rilevanti	8
2.1.1	Pressione sonora	8
2.1.2	Timbro	9
2.1.3	Effetto Doppler	9
2.1.4	Riverberazione	9
2.2	Modello di sorgente in movimento real-time	10
2.2.1	Implementazione delle variazioni di intensità	10
2.2.2	Implementazione delle variazioni del timbro	11
2.2.3	Implementazione dell'effetto Doppler	11
2.2.4	Implementazione della riverberazione	11
2.3	Modello del Leslie	12
3	Modello proposto	17
3.1	Descrizione del modello	17
3.2	Modello di sorgente in movimento	19
3.2.1	Delay line con lettura interpolata	23
3.2.2	Delay line con scrittura interpolata	26
3.3	Crossover	27
3.4	Filtro di assorbimento dell'aria	28
3.5	Filtro di assorbimento del legno	29
3.6	Modulazione in ampiezza	31
3.7	Dinamica dei diffusori rotanti	33
3.8	Amplificatore	37
3.8.1	Waveshaper	37
3.8.2	Modello di Wiener-Hammerstein allenato con backpropagation	38
3.8.3	Modello di Wiener-Hammerstein ricavato da procedura di ottimizzazione	40
4	Implementazione	44
4.1	Cenni su VCV Rack	44

Indice

4.2	Sviluppo del modulo del Leslie	46
4.2.1	Geometria delle sorgenti	47
4.2.2	Implementazione della funzione di elaborazione del modulo	49
5	Risultati	53
5.1	Amplificatore	53
5.1.1	Modello di Wiener-Hammerstein allenato con backpropagation	53
5.1.2	Modello di Wiener-Hammerstein ricavato da procedura di ottimizzazione	54
5.1.3	Osservazioni	57
5.2	Simulatore degli altoparlanti rotanti	58
5.2.1	Simulatore della tromba	58
5.2.2	Simulatore del woofer	58
5.2.3	Osservazioni	61
5.2.4	Filtro di assorbimento dell'aria	65
5.2.5	Confronto tra scrittura interpolata e lettura interpolata	65
6	Conclusioni	67

Elenco delle figure

1.1	Leslie 122 con e senza il pannello posteriore.	2
1.2	Schema del Leslie per i modelli 145, 147 e 122.	3
1.3	Unità degli alti del Leslie.	3
1.4	Diagramma di radiazione della tromba del Leslie senza il cono diffusore.	4
1.5	Diagramma di radiazione della tromba del Leslie con il cono diffusore.	4
1.6	Risposta in ampiezza della tromba del Leslie senza il cono diffusore.	5
1.7	Risposta in ampiezza della tromba del Leslie con il cono diffusore.	5
1.8	Unità dei bassi del Leslie.	6
2.1	Modello per la simulazione di una sorgente sonora in movimento.	10
2.2	Implementazione del modello di Figura 2.1.	10
2.3	Modello del simulatore della tromba del Leslie con 5 sorgenti immagine.	12
2.4	Risultati ottenuti dalla registrazione del vero Leslie.	14
2.5	Risultati ottenuti dalla registrazione dell'uscita del simulatore in cui sono state escluse le sorgenti immagine.	15
2.6	Risultati ottenuti dalla registrazione dell'uscita del simulatore con le sorgenti immagine.	16
3.1	Schema a blocchi del modello del Leslie proposto in questo lavoro.	17
3.2	Schema interno dei blocchi di elaborazione dei segnali della tromba e del wooper.	18
3.3	Delay line con puntatore di scrittura fisso e puntatore di lettura variabile.	19
3.4	Implementazione del modello di sorgente in movimento con delay line con lettura variabile. Le operazioni di filtraggio e AM sono effettuate sul campione in ingresso alla delay line.	20
3.5	Implementazione del modello di sorgente in movimento con delay line con lettura variabile. Le operazioni di filtraggio e AM sono effettuate sul campione in uscita dalla delay line.	21
3.6	Confronto tra il ritardo di propagazione reale (Eq. 3.4) e quello ottenuto dal calcolo approssimato (Eq. 3.5) nel caso di una sorgente che ruota con raggio di 20 cm e frequenza di 7 Hz attorno ad un punto a distanza 1 m da un ascoltatore.	21
3.7	Modello di due sorgenti in movimento con lettura variabile della delay line.	22
3.8	Delay line con puntatore di scrittura variabile e puntatore di lettura fisso.	22

Elenco delle figure

3.9	Modello di due sorgenti in movimento con scrittura variabile della delay line.	23
3.10	Schema dell'implementazione della lettura interpolata di una delay line.	23
3.11	Risposta in ampiezza (in alto) e in fase (in basso) di un interpolatore lineare per undici diversi delay frazionali ($D = 0, 0.1, 0.2, \dots 1.0$). Si noti che ci sono solo sei curve nel grafico in alto perché le risposte in ampiezza sono uguali per le parti frazionali d e $1 - d$. Nel grafico in basso, le linee tratteggiate indicano la risposta in fase ideale.	24
3.12	Schema dell'implementazione della scrittura interpolata (nel caso di interpolazione lineare) di una delay line.	26
3.13	Risposte in ampiezza passa-basso e passa-alto del crossover.	28
3.14	Funzioni di trasferimento dell'aria (linee continue) e funzioni di trasferimento simulate (linee tratteggiate) ottenute dal processo di ottimizzazione per varie distanze tra sorgente e ascoltatore. Le funzioni di trasferimento dell'aria sono state calcolate considerando una temperatura di 20° , pressione atmosferica pari a 1013 hPa e 45% di igrometria. La frequenza di taglio del filtro simulato è stata settata a 10 kHz e il guadagno è stato calcolato usando la Eq. 3.9.	29
3.15	Coefficienti di assorbimento misurati per alcuni tipi di legno.	30
3.16	Risposta in ampiezza del filtro progettato (linea continua) e risposta in ampiezza da approssimare (linea tratteggiata).	31
3.17	Poli e zeri del filtro progettato.	31
3.18	Schema di una sorgente sonora S che ruota intorno ad un punto H ad una certa distanza da un microfono.	32
3.19	Andamento della modulante per differenti valori dei parametri <i>directivityAngle</i> e <i>directivityLevel</i> . Le curve sono riferite ad una sorgente che ruota con una frequenza di 6.5 Hz e un raggio di 0.19 m attorno ad un punto distante 0.8 m da un microfono.	33
3.20	Sono mostrati l'andamento nel tempo del segnale registrato per la transizione <i>fast</i> \rightarrow <i>slow</i> , l'involuppo calcolato e le posizioni dei picchi dell'involuppo.	34
3.21	Risultato del fitting per la transizione <i>brake</i> \rightarrow <i>fast</i>	34
3.22	Risultato del fitting per la transizione <i>brake</i> \rightarrow <i>slow</i>	35
3.23	Risultato del fitting per la transizione <i>slow</i> \rightarrow <i>fast</i>	35
3.24	Risultato del fitting per la transizione <i>slow</i> \rightarrow <i>brake</i>	35
3.25	Risultato del fitting per la transizione <i>fast</i> \rightarrow <i>slow</i>	36
3.26	Risultato del fitting per la transizione <i>fast</i> \rightarrow <i>brake</i>	36
3.27	Schema a blocchi del modello di Wiener-Hammerstein.	37
3.28	Circuito interno al blocco non lineare.	40
3.29	Funzione di mapping per differenti valori dei suoi parametri. Nella legenda i parametri sono ordinati come: k_p, k_n, g_p, g_n . I valori di g_p e g_n sono in dB.	41

Elenco delle figure

3.30	Diagramma a blocchi per il calcolo della funzione costo $C(\mathbf{p})$ e della metrica PEAS. Gli step aggiuntivi per il calcolo della PEAS sono mostrati in grigio.	43
4.1	Interfaccia grafica di VCV Rack.	44
4.2	Interfaccia grafica del modulo sviluppato.	47
4.3	Geometria del simulatore della tromba (vista dall'alto del Leslie). . .	48
4.4	Geometria del simulatore della tromba (vista frontale del Leslie). . .	48
4.5	Geometria della sorgente principale.	49
4.6	Schema a blocchi della funzione <code>process</code> . Viene mostrato come i parametri del modulo (in arancione) controllano le varie operazioni.	50
4.7	Implementazione dei blocchi <code>HORN PROCESSING</code> e <code>WOOFER PROCESSING</code> nel caso di utilizzo di delay line con lettura interpolata. Si considerano la sorgente di suono diretto e due sorgenti immagine.	51
4.8	Implementazione dei blocchi <code>HORN PROCESSING</code> e <code>WOOFER PROCESSING</code> nel caso di utilizzo di delay line con scrittura interpolata. Si considerano la sorgente di suono diretto e due sorgenti immagine.	52
5.1	Setup utilizzato per la misurazione dell'amplificatore del Leslie. . . .	53
5.2	Risultato dell'adattamento dei filtri IIR che formano il blocco lineare d'ingresso del modello.	55
5.3	Risultato dell'adattamento dei filtri IIR che formano il blocco lineare d'uscita del modello.	55
5.4	Sono mostrati: lo spettrogramma del segnale di riferimento; lo spettrogramma del modello <code>WH_SVF66_MLP2</code> ; lo spettrogramma del modello ottenuto dalla procedura di ottimizzazione.	57
5.5	Forme d'onda del segnale in uscita dal simulatore e del segnale misurato. Il numero di sorgenti virtuali considerate viene progressivamente aumentato: nel primo grafico viene considerata solamente la sorgente di suono diretto; nel secondo grafico vengono aggiunte le immagini corrispondenti alle pareti laterali; nel terzo grafico viene aggiunta l'immagine corrispondente alla parete frontale; nel quarto grafico si considerano tutte le riflessioni del primo ordine.	59
5.6	Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando solamente la sorgente di suono diretto, e lo spettrogramma del segnale misurato.	59
5.7	Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali, e lo spettrogramma del segnale misurato.	60
5.8	Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale, e lo spettrogramma del segnale misurato.	60

Elenco delle figure

5.9	Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando tutte le immagini del primo ordine, e lo spettrogramma del segnale misurato.	61
5.10	Forme d'onda del segnale in uscita dal simulatore e del segnale misurato. Nel primo grafico vengono considerate la sorgente principale e le immagini corrispondenti alla pareti laterali; nel secondo e terzo grafico viene aggiunta l'immagine corrispondente alla parete frontale; nel quarto grafico si considerano tutte le riflessioni del primo ordine.	62
5.11	Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali e settando <i>directivityLevel</i> = 5, e lo spettrogramma del segnale misurato.	62
5.12	Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale e settando <i>directivityLevel</i> = 5, e lo spettrogramma del segnale misurato.	63
5.13	Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale e settando <i>directivityLevel</i> = 1, e lo spettrogramma del segnale misurato.	63
5.14	Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando tutte le immagini del primo ordine e settando <i>directivityLevel</i> = 5, e lo spettrogramma del segnale misurato.	64
5.15	Applicazione del teorema delle immagini. S è la sorgente principale e S^1 è la sorgente immagine del primo ordine corrispondente alla parete destra. In questa situazione S^1 è una sorgente immagine valida. La linea blu indica il percorso dal suono corrispondente alla riflessione del primo ordine sulla parete destra.	64
5.16	Applicazione del teorema delle immagini. S è la sorgente principale, e S^1 , S^2 , S^3 sono, rispettivamente, immagini del primo, secondo e terzo ordine. In questa situazione S^1 non è una sorgente immagine valida. Considerare il percorso arancione del suono è un errore perché la linea non intercetta la parete destra del cabinet. Comunque, nello stesso istante, dovrebbero essere considerate delle riflessioni di ordine superiore. Un esempio è la riflessione del terzo ordine data dal percorso blu.	65
5.17	Forme d'onda ottenute in uscita del simulatore della tromba per i due possibili schemi implementativi. L'ingresso del modello è un segnale sinusoidale con una frequenza di 2 kHz.	66
5.18	Spettrogrammi calcolati sui segnali di Figura 5.17.	66

Elenco delle tabelle

4.1	Coordinate dei centri di rotazione.	47
4.2	Coordinate delle sorgenti.	49
5.1	Elenco dei modelli su cui è stato effettuato il training.	54
5.2	Risultati del training e del testing delle reti.	54
5.3	Risultato dell'identificazione dell'amplificatore del Leslie. Sono riportati i valori calcolati delle metriche.	56
5.4	Risultato dell'identificazione del modello digitale di una valvola. Sono riportati i valori calcolati delle metriche.	56
5.5	Valori delle metriche per il modello WH_SVF66_MLP2. Il calcolo è effettuato sui segnali di Figura 5.4.	58
5.6	Valori delle metriche per il modello ricavato dalla procedura di ottimizzazione. Il calcolo è effettuato sui segnali di Figura 5.4.	58

Capitolo 1

Introduzione

1.1 Descrizione del Leslie

"Leslie" è un marchio di fabbrica usato universalmente per indicare un sistema di altoparlanti rotanti inventato da Don Leslie alla fine degli anni quaranta. Il Leslie non è un altoparlante "Hi-Fi", infatti venne ideato come complemento per i primi organi Hammond, per modificarne il suono. Tuttora, è prevalentemente utilizzato dagli organisti in questo modo, ma è stato impiegato anche come effetto audio per altri strumenti, come la chitarra elettrica, e per la voce umana.

Il Leslie opera secondo un semplice principio: una sorgente sonora viene fatta ruotare attorno ad un fulcro. In un punto di ascolto, ad una certa distanza dalla sorgente, accadono tre cose:

1. Per via del cambiamento della distanza tra sorgente e ascoltatore e della direttività della sorgente, il livello sonoro percepito cambia nel tempo. Si ha quindi una modulazione in ampiezza del suono, chiamata effetto "tremolo".
2. L'effetto Doppler causa un aumento del pitch percepito dall'ascoltatore quando la sorgente si avvicina e una diminuzione del pitch quando la sorgente si allontana. Si ha quindi una modulazione in frequenza del suono, chiamata effetto "vibrato".
3. L'accoppiamento della sorgente con l'ambiente circostante varia nel tempo causando una modulazione della risposta impulsiva.

Nel corso degli anni, il produttore di Leslie ha proposto un'ampia varietà di modelli con caratteristiche differenti. Tuttavia, secondo gli appassionati, solamente i modelli 145, 147 e 122 sono quelli con il vero "suono Leslie". Tutti e tre i modelli condividono le stesse caratteristiche di base: un amplificatore a valvole monofonico da 40 Watt, un crossover passivo da 800 Hz a 16 Ω , una tromba rotante per gli alti e un altoparlante rotante per i bassi. Entrambi gli altoparlanti funzionano secondo lo stesso principio: un driver fisso e un "proiettore" acustico rotante. Inoltre, questi modelli sono simili in quanto hanno tutti i componenti montati in un cabinet in legno a tre scomparti (Figura 1.1). Il vano superiore ospita la tromba rotante ad alta frequenza, lo scomparto centrale i driver delle alte e basse frequenze e il crossover,

il vano inferiore il deflettore dei bassi e l'amplificatore. Le feritoie situate negli scomparti superiore e inferiore emettono, rispettivamente, il suono degli alti e dei bassi.



Figura 1.1: Leslie 122 con e senza il pannello posteriore.

Ogni rotore è pilotato da due motori separati (uno lento e uno veloce) montati in un singolo contenitore. Questi modelli prevedono, infatti, tre differenti modi di funzionamento:

- *fast* (anche detto "tremolo"), in cui i diffusori ruotano ad una velocità di circa 400 rpm;
- *slow* (anche detto "chorale"), in cui i diffusori ruotano ad una velocità di circa 40 rpm;
- *brake* (anche detto "stop"), in cui i diffusori sono fermi.

Un'altra caratteristica è che i versi di rotazione dei due rotori sono opposti.

Uno schema del Leslie completo è mostrato in Figura 1.2. L'ingresso del sistema è quello dell'amplificatore, che consiste in una presa a 6 pin a cui vengono inviati il segnale d'ingresso, i controlli dei quattro motori e la tensione di linea alternata. L'uscita dell'amplificatore è inviata al crossover, che pilota i due driver a 16 Ω nelle appropriate bande di funzionamento.

Si riportano ora alcune informazioni tecniche prese da [1]. L'unità degli alti consiste in un driver a connessione che sfoga all'interno di una coppia di trombe nere in

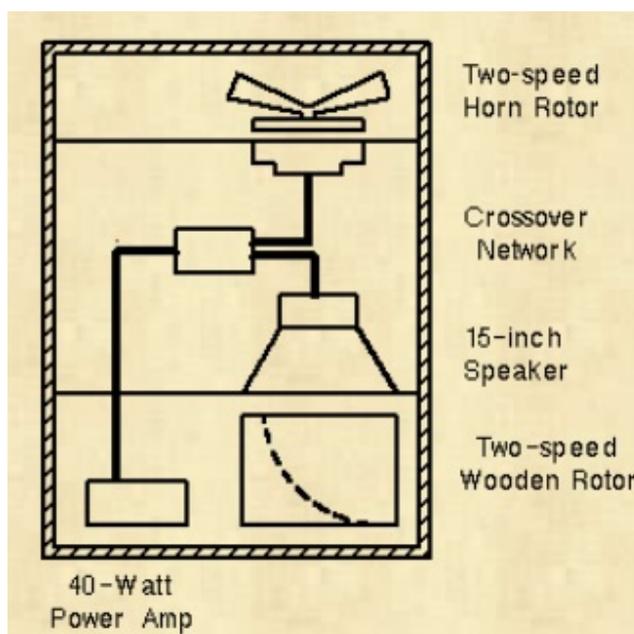


Figura 1.2: Schema del Leslie per i modelli 145, 147 e 122.

bachelite. Come si vede in Figura 1.3, solo una delle due trombe è attiva (emette suono), mentre l'altra serve solo per l'equilibratura dinamica della rotazione, fornendo una resistenza dell'aria simmetrica ad alta velocità. La struttura risultante ruota senza intoppi e senza forze eccentriche di oscillazione.

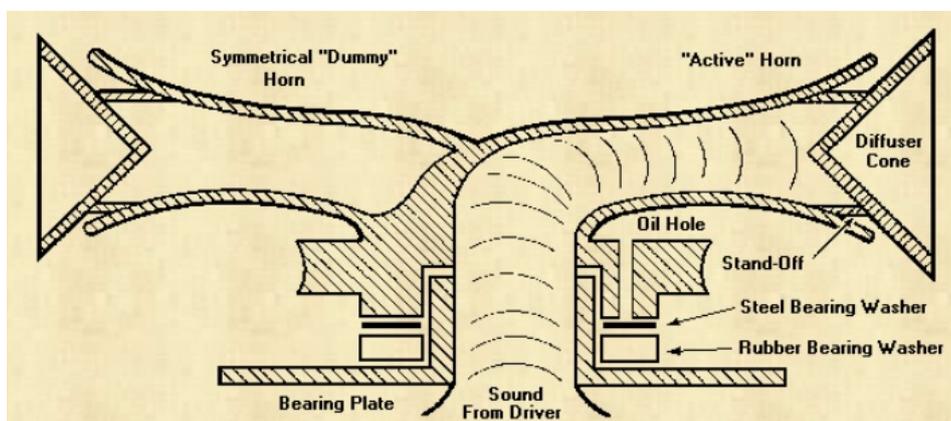


Figura 1.3: Unità degli alti del Leslie.

Sulla bocca della tromba è posizionato un cono diffusore che colora e diffonde ulteriormente il suono. Il diagramma di radiazione della tromba, con e senza il diffusore, è mostrato in Figura 1.4 e in Figura 1.5. Si nota che senza il diffusore il suono è molto più concentrato sull'asse della tromba, mentre con il diffusore il pattern di radiazione è quasi omnidirezionale. In Figura 1.6 e in Figura 1.7 è mostrata la risposta in ampiezza della tromba con e senza il cono diffusore. Si nota una caratteristica passa-banda centrata sui 2 kHz. Inoltre, con il diffusore

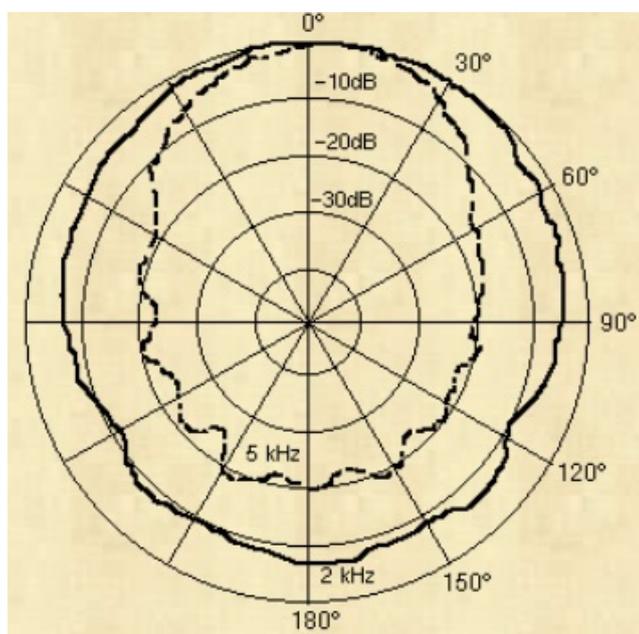


Figura 1.4: Diagramma di radiazione della tromba del Leslie senza il cono diffusore.

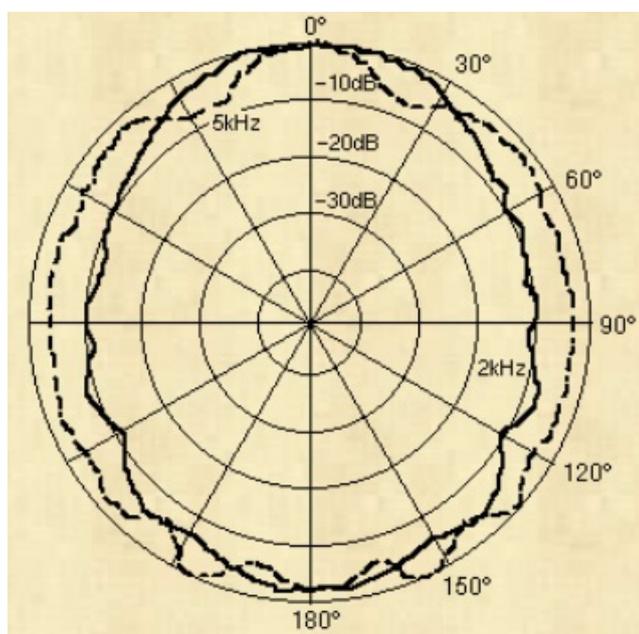


Figura 1.5: Diagramma di radiazione della tromba del Leslie con il cono diffusore.

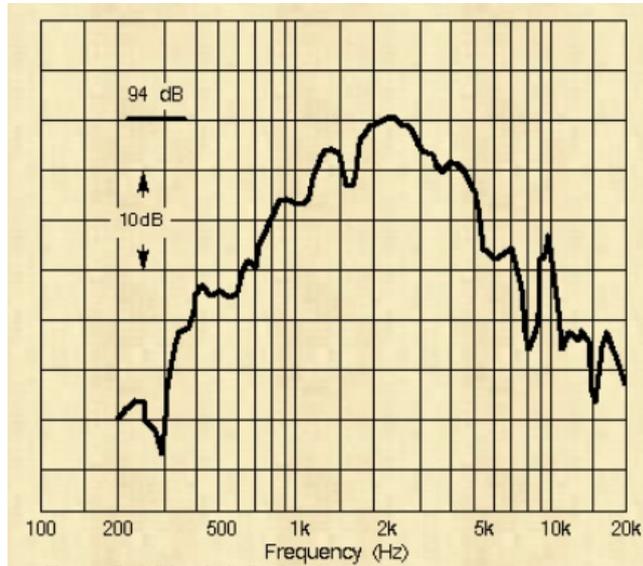


Figura 1.6: Risposta in ampiezza della tromba del Leslie senza il cono diffusore.

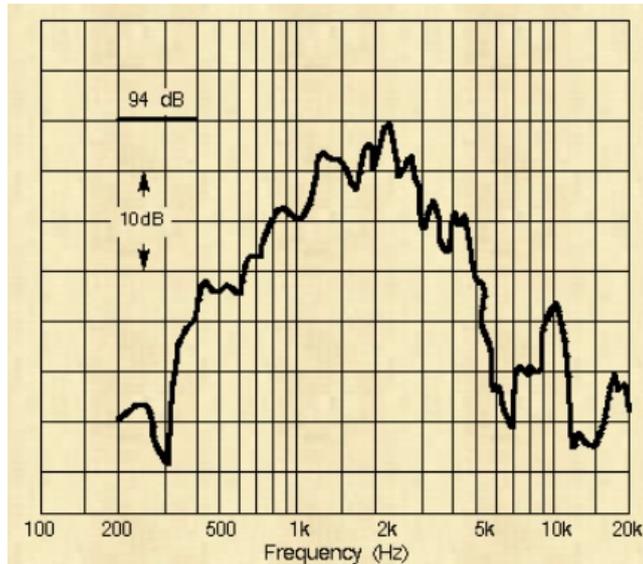


Figura 1.7: Risposta in ampiezza della tromba del Leslie con il cono diffusore.

si ha una sensibilità maggiore al di sopra dei 1.5 kHz. L'intero discorso è basato su misurazioni effettuate in camera anecoica. Nella pratica, la tromba è montata all'interno del cabinet, quindi la figura sonora cambierà per via delle riflessioni del suono sulle pareti. Il cono diffusore ha anche un altro effetto: sposta la posizione apparente della sorgente sonora per via del cambiamento d'impedenza visto dall'onda sonora all'uscita della tromba. Con il diffusore sembra che l'intero range del suono venga emesso dalla bocca della tromba; senza il diffusore, mentre le basse frequenze sembrano essere emesse dalla bocca della tromba, le frequenze più alte sembrano provenire progressivamente da un punto più in basso nella gola della tromba. Quindi, è come se le alte frequenze ruotassero con un raggio minore, il che risulta in un effetto di modulazione in frequenza minore. Per questo, con il cono diffusore si ha la massima modulazione in frequenza e una modulazione in ampiezza relativamente bassa (per via della minore direttività della sorgente). Viceversa, senza il diffusore si ha un minore modulazione in frequenza e una maggiore modulazione in ampiezza.

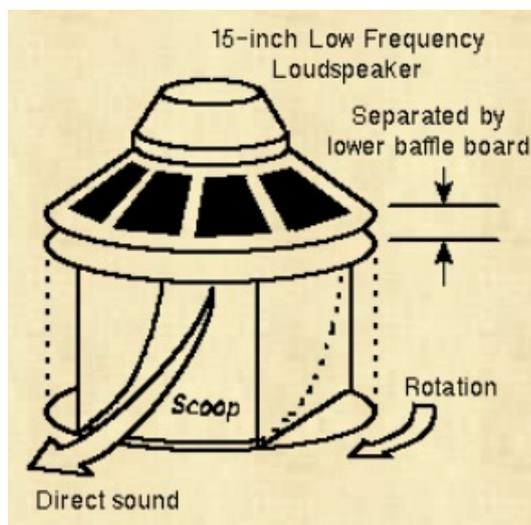


Figura 1.8: Unità dei bassi del Leslie.

L'unità dei bassi (Figura 1.8) consiste in un altoparlante a cono posto con il magnete in alto e la membrana orientata verso il basso. Sotto l'altoparlante c'è un deflettore acustico rotante in legno ricoperto da una tela nera, che fornisce una minore resistenza aerodinamica ad alta velocità. Il suono è emesso da una apertura a forma di parabola concava. Il deflettore ruota approssimativamente alla stessa velocità della tromba. L'unica differenza è che la sua inerzia lo fa avvicinare alla velocità finale in un periodo di tempo molto più lungo. L'effetto principale del woofer è la modulazione in ampiezza (questa probabilmente non influisce sulle frequenze inferiori a 200 Hz), anche se c'è una certa modulazione in frequenza vicino alla frequenza di crossover (800 Hz).

1.2 Motivazioni

Questo lavoro tratta lo sviluppo di un modello digitale per la simulazione del Leslie. La necessità di un'implementazione digitale è dovuta a vari motivi: la grandezza e il peso del Leslie ne rendono difficile il trasporto; alcuni componenti hanno una durata limitata nel tempo (ad esempio le valvole dell'amplificatore) e potrebbe essere difficile trovare dei ricambi; l'elevato consumo di energia; il prezzo elevato; la scarsità di produzione di nuovi prodotti ispirati ai modelli vintage. Un ulteriore vantaggio di un'emulazione digitale è che un unico hardware può essere utilizzato per simulare molteplici modelli di Leslie, semplicemente caricando nuovi parametri nel dispositivo. Inoltre, il modello può essere implementato come plug-in software ed essere utilizzato connettendo il proprio strumento alla scheda audio del computer (particolarmente utile in un home studio). Tuttavia, il vero Leslie continua ad essere insostituibile per molti musicisti per via della difficoltà di modellare adeguatamente i fenomeni complessi che avvengono quando questo viene suonato. Questi sono, ad esempio, la direzionalità dipendente dalla frequenza delle sorgenti, l'accoppiamento del suono con il cabinet e la stanza, la timbrica degli altoparlanti e la propagazione del suono attraverso le feritoie. Un modello accurato sotto tutti gli aspetti comporterebbe un elevato costo computazionale, che ne renderebbe impossibile l'implementazione su piattaforme DSP, in cui spesso le risorse sono limitate. L'obiettivo di questo lavoro è quello di proporre un modello che sia computazionalmente efficiente ma che riesca a simulare gli effetti percettivi fondamentali che contraddistinguono l'"effetto Leslie".

1.3 Panoramica della tesi

La tesi è organizzata come segue. Nel capitolo 2 si descriverà l'articolo di Kronland et al. [2], che propongono un modello digitale per la simulazione di sorgenti in movimento, prendendo il Leslie come caso di studio. Questo lavoro è stato utilizzato come base per lo sviluppo del modello del Leslie oggetto di questa tesi, che sarà presentato nel capitolo 3. Nel capitolo 4 si descriverà l'implementazione del modello come plug-in per il software VCV Rack. Il plug-in è stato utilizzato per parametrizzare il modello e valutarne l'accuratezza tramite test di ascolto e confronti con un vero Leslie. I risultati sono riportati nel capitolo 5.

Capitolo 2

Modello di riferimento

In questo capitolo si descriverà il lavoro di Kronland et al. che, in [2], hanno proposto un modello real-time per la simulazione di una sorgente sonora in movimento che può essere applicato a qualsiasi segnale audio. Questo modello, che è basato sugli studi di Chowning, è stato calibrato usando la conoscenza fisica che riguarda la propagazione del suono, includendo il fenomeno di assorbimento dell'aria, il processo di riverberazione e l'effetto Doppler. Come applicazione audio viene presentato un modello del Leslie.

2.1 I segnali percettivi rilevanti

Secondo gli autori, sono quattro i segnali percettivi rilevanti che possono essere utilizzati per costruire un generico modello per una sorgente sonora in movimento:

- pressione sonora;
- timbro;
- effetto Doppler;
- riverberazione.

Molti di questi non dipendono dal processo di spazializzazione coinvolto, ma sono comunque molto influenti nella percezione dei suoni, compresi quelli emessi da sorgenti fisse.

2.1.1 Pressione sonora

Da un punto di vista fisico, la pressione sonora è correlata all'intensità sonora e, in maniera più complessa, alla loudness. La pressione sonora varia con l'inverso della distanza tra sorgente e ascoltatore. Questa regola è di grande importanza da un punto di vista percettivo ed è forse decisiva nel caso di sorgenti che si muovono lentamente. Vale la pena notare che solo i cambiamenti relativi della pressione sonora dovrebbero essere presi in considerazione, visto che la pressione assoluta ha un effetto minimo sul risultante effetto percettivo.

2.1.2 Timbro

Il timbro è l'attributo percettivo che rende possibile discriminare tra suoni differenti aventi stesso pitch, loudness e durata. Dal un punto di vista dell'elaborazione del segnale, le variazioni del timbro si riflettono in cambiamenti sia dell'evoluzione temporale che della distribuzione spettrale dell'energia sonora. I cambiamenti del timbro dovuti al movimento di una sorgente possono essere tenuti in conto, fisicamente, in termini di assorbimento dell'aria. L'effetto percettivo principale dell'assorbimento dell'aria sul suono è dato da un processo di filtraggio passa-basso, il risultato del quale dipende dalla distanza tra sorgente e ascoltatore. In condizioni abituali, nella banda che va da 0 a 3 kHz, il timbro varia molto poco anche per grandi distanze tra sorgente e ascoltatore. Per simulare sorgenti sonore in movimento che coprono grandi distanze l'effetto dell'assorbimento dell'aria deve essere tenuto in conto.

2.1.3 Effetto Doppler

Da un punto di vista fisico, sorgenti sonore in movimento provocano uno shift di frequenza conosciuto come effetto Doppler. A seconda della velocità relativa della sorgente rispetto all'ascoltatore, la frequenza ω_1 misurata nella posizione dell'ascoltatore è data da:

$$\omega_1 = \omega_s \left(\frac{1 + v_{ls}/c}{1 - v_{sl}/c} \right) \quad (2.1)$$

dove ω_s è la frequenza emessa dalla sorgente, v_{ls} e v_{sl} denotano, rispettivamente, la velocità relativa dell'ascoltatore in direzione della sorgente e la velocità relativa della sorgente in direzione dell'ascoltatore, e c è la velocità del suono. Durante una traiettoria della sorgente sonora la frequenza percepita è quindi dipendente dal tempo e il suo schema specifico è un segnale percettivo altamente rilevante che consente all'ascoltatore di costruire una rappresentazione mentale della traiettoria.

2.1.4 Riverberazione

La qualità del suono è fortemente influenzata dall'ambiente circostante. L'effetto della riverberazione può essere spiegato dalle leggi fisiche della propagazione del suono, le quali impongono che sorgenti sonore distanti portano a segnali più altamente riverberati rispetto a sorgenti vicine perché, con sorgenti sonore distanti, il percorso diretto e le riflessioni del suono sono dello stesso ordine di ampiezza, mentre, con sorgenti vicine, il suono diretto ha un'ampiezza maggiore rispetto al suono riflesso. Sorgenti sonore in movimento implicano, quindi, un rapporto tra suono diretto e riflesso dipendente dal tempo, il cui valore dipende dalla distanza tra sorgente e ascoltatore.

2.2 Modello di sorgente in movimento real-time

In linea con le considerazioni di cui sopra, viene proposto un modello per la simulazione di una sorgente in movimento composto dalla combinazione delle quattro componenti percettive principali (Figura 2.1). La velocità relativa della sorgente mobile e la sua distanza dall'ascoltatore controllano i parametri del modello. Il suono risultante è destinato ad un ascolto monofonico, ma potrebbe essere legato ad un sistema di spazializzazione, esaltandone il realismo. Lo schema implementativo del modello è mostrato in Figura 2.2.

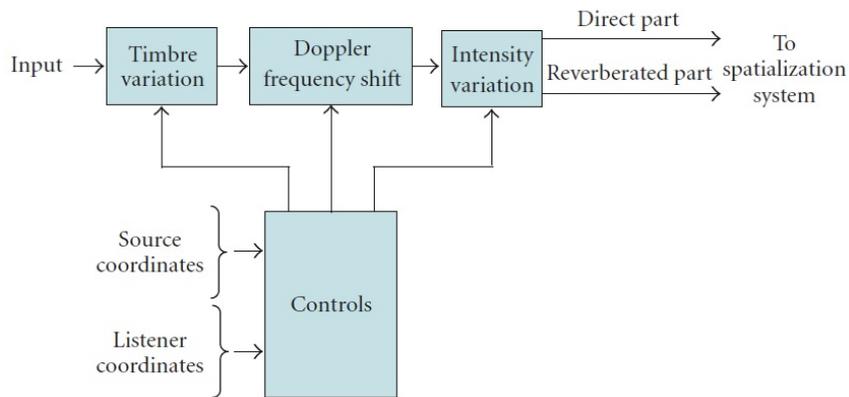


Figura 2.1: Modello per la simulazione di una sorgente sonora in movimento.

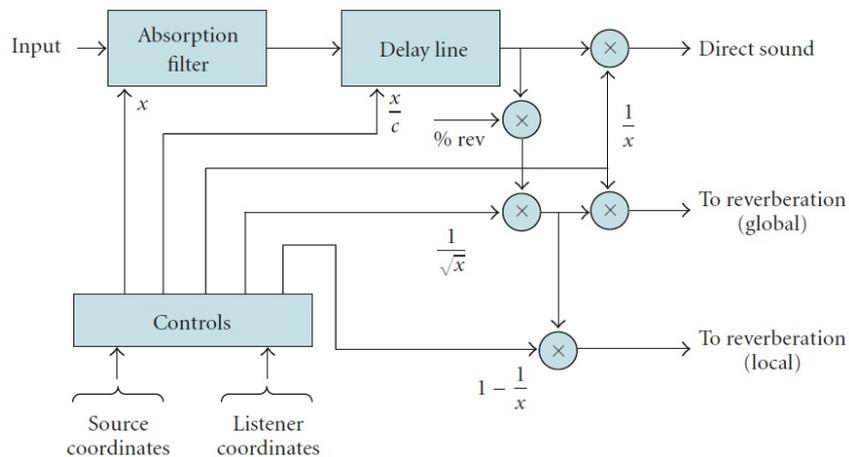


Figura 2.2: Implementazione del modello di Figura 2.1.

2.2.1 Implementazione delle variazioni di intensità

Le variazioni di intensità sono controllate direttamente dal livello sonoro. Assumendo che la propagazione sonora coinvolge onde sferiche, l'ampiezza del segnale varia rispetto a $1/x$, dove x è la distanza tra sorgente e ascoltatore. Da un punto di

vista pratico, bisogna prestare attenzione per evitare problemi di divergenza quando $x = 0$.

2.2.2 Implementazione delle variazioni del timbro

Le variazioni del timbro causate dall'assorbimento dell'aria vengono simulate con un filtro, la cui risposta in frequenza varia in funzione della distanza tra sorgente e ascoltatore. Visto che questo fattore è probabilmente di minore importanza percettiva rispetto ad altri, è possibile semplificare il suo trattamento nel processo di implementazione. In [3] è stato stabilito che la risposta in ampiezza del filtro può essere modellata utilizzando filtri IIR.

2.2.3 Implementazione dell'effetto Doppler

L'effetto Doppler è dovuto al cambiamento della lunghezza del percorso tra sorgente e ascoltatore; quindi, al cambiamento del tempo di propagazione $\tau(t)$. Lo shift in frequenza (Eq. 2.1) può essere quindi controllato con una delay line variabile. Per una sorgente sonora che emette un segnale monocromatico e si muove rispetto ad un ascoltatore fisso, Smith et al. [4] hanno ottenuto la seguente espressione:

$$\frac{d\tau(t)}{dt} = -\frac{v_{sl}}{c} \quad (2.2)$$

La proiezione della velocità della sorgente sulla linea che congiunge sorgente e ascoltatore v_{sl} può essere quindi calcolata, per ogni campione del segnale, ed utilizzata per ottenere l'incremento da applicare al delay. Il metodo utilizzato in [2] per calcolare il valore corrente del delay ha un costo computazionale minore ed è quello proposto da Tsingos [5], che dà la seguente espressione per $\tau(t)$:

$$\tau(t) = \frac{1}{c} \|\mathbf{L}(t) - \mathbf{S}(t - \tau(t))\| \quad (2.3)$$

dove $\mathbf{L}(t)$ e $\mathbf{S}(t)$ sono, rispettivamente, i vettori posizione dell'ascoltatore e della sorgente all'istante t , e $\|\cdot\|$ denota la distanza Euclidea. Questa espressione è stata semplificata, visto che si ottengono effetti percettivi simili anche a velocità elevate delle sorgenti:

$$\tau(t) = \frac{1}{c} \|\mathbf{L}(t) - \mathbf{S}(t)\| = \frac{x(t)}{c} \quad (2.4)$$

Si noti che la delay line deve trattare con valori frazionari di $\tau(t)$.

2.2.4 Implementazione della riverberazione

La riverberazione dipende dall'ambiente circostante e il trattamento è solitamente lasciato all'utente. Comunque, in [2] viene spiegato, sulla base di [6], in che modo settare i livelli dei segnali inviati alla riverberazione. Si assume che in uno spazio piccolo l'ampiezza del segnale riverberato prodotto da una sorgente sonora con

intensità costante a varie distanze dall'ascoltatore vari poco, ma che in spazi grandi vari molto. Per questo, l'ampiezza del segnale riverberato (che è dato da una percentuale del segnale diretto) è resa proporzionale a $1/\sqrt{x}$. In seguito, il segnale riverberato viene separato in due componenti: la riverberazione globale, che ha origine dall'intero spazio, e la riverberazione locale, che ha origine dalla direzione della sorgente. La prima è resa proporzionale a $1/x$, mentre la seconda a $(1 - \frac{1}{x})$. In questo modo, all'aumentare della distanza tra sorgente e ascoltatore la riverberazione diventa più localizzata, compensando la perdita di energia del segnale diretto. Questa è una giusta approssimazione di una situazione acustica reale in cui all'aumentare della distanza della sorgente la sua distanza dalle pareti riflettenti diminuisce. Nel caso di un sistema di spazializzazione con altoparlanti multipli, la riverberazione globale va distribuita in egual modo su tutti i canali, mentre la riverberazione locale segue la sorgente in movimento.

2.3 Modello del Leslie

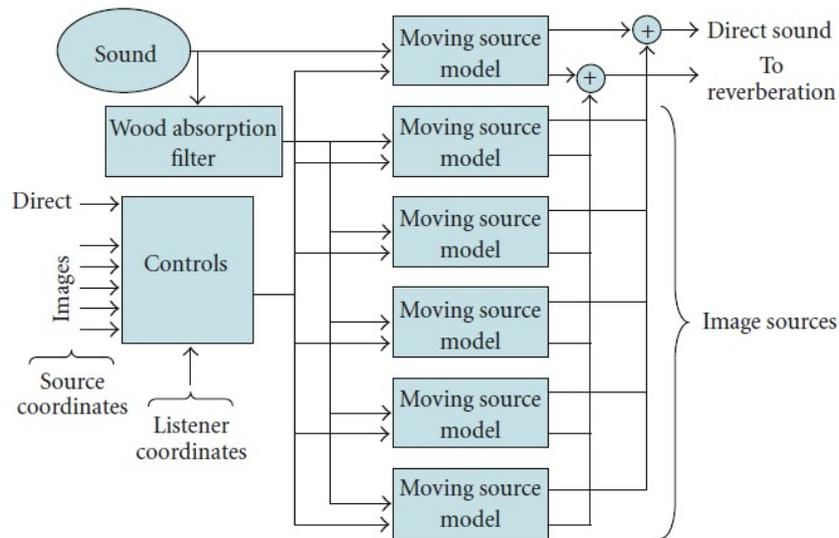


Figura 2.3: Modello del simulatore della tromba del Leslie con 5 sorgenti immagine.

Lo schema del simulatore della tromba del Leslie, proposto in [2], è mostrato in Figura 2.3. Oltre alla sorgente principale, che genera il suono diretto, sono state considerate delle sorgenti virtuali aggiuntive, che simulano le riflessioni sulle pareti del cabinet. Le coordinate di queste sorgenti possono essere facilmente dedotte dalla geometria del cabinet utilizzando il teorema delle immagini [7]. La complessità computazionale del metodo delle immagini aumenta linearmente con il numero di riflessione prese in considerazione. Nell'articolo si afferma che considerare una sorgente immagine per ogni piano riflettente (primo ordine) è sufficiente per ottenere dei risultati soddisfacenti dal punto di vista percettivo. Ogni sorgente è processata

utilizzando il modello di sorgente in movimento visto in precedenza. Inoltre, i segnali relativi alle sorgenti immagine sono filtrati per tener conto dell'assorbimento, dipendente in frequenza, delle pareti in legno (il segnale diretto non viene filtrato perché si assume che l'ascoltatore sia posizionato sul lato del cabinet in cui il pannello è stato rimosso). Per il simulatore del woofer si utilizza lo stesso modello. Come nel Leslie reale, gli ingressi dei simulatori della tromba e del woofer sono generati da un crossover.

Per valutare l'accuratezza del modello, le caratteristiche principali del segnale simulato sono state confrontate con quelle di una registrazione del vero Leslie. A tale scopo, le uscite del simulatore e di un Leslie 122A sono state registrate mandando in ingresso un segnale sinusoidale con una frequenza di 800 Hz (frequenza d'incrocio del crossover), in modo da includere sia l'effetto della tromba che del woofer. Per la misura, il vero Leslie è stato posizionato in una stanza anecoica e la pressione acustica è stata misurata usando un microfono posizionato a 1.2 m di distanza dal cabinet, alla stessa altezza del piano di rotazione della tromba. Dai segnali registrati $s(t)$, i segnali analitici, dati da $Z(t) = s(t) + iH[s](t) = A(t)e^{i\phi(t)}$ (dove H denota la trasformata di Hilbert), sono stati calcolati per dedurre le leggi di modulazione dell'ampiezza $A(t)$ e della frequenza istantanea $d\phi/dt$. I risultati sono mostrati in Figura 2.4-Figura 2.6.

Quando la parte delle sorgenti immagine non è attiva il segnale di uscita del modello mostra delle modulazioni di ampiezza e frequenza periodiche, comparabili con le informazioni fornite in [4]. In questo caso, l'effetto udibile risultante è una combinazione degli effetti vibrato e tremolo, che però non corrisponde completamente al tipico "effetto Leslie". Quando le sorgenti immagine sono attive le caratteristiche del segnale d'uscita del modello diventano più complesse e simili a quelle del segnale registrato dal vero Leslie. In particolare, si nota che la frequenza istantanea mostra uno schema tipico, dove i picchi alti positivi e i picchi negativi si verificano in corrispondenza di ampiezze del segnale quasi nulle. Questo comportamento aperiodico delle leggi di modulazione, ritenuto responsabile del particolare "effetto Leslie", non è dovuto all'interferenza tra il suono generato dal woofer e quello della tromba, ma dalle riflessioni sulle pareti del cabinet. Sono state infatti effettuate ulteriori misurazioni pilotando la sola tromba e il comportamento è stato ancora osservato. Questo dimostra l'importanza di includere le sorgenti immagine nel modello.

Sono stati effettuati test di ascolto utilizzando dei segnali musicali e il suono ottenuto dal simulatore è stato descritto da musicisti professionisti di una qualità simile a quella di un vero Leslie.

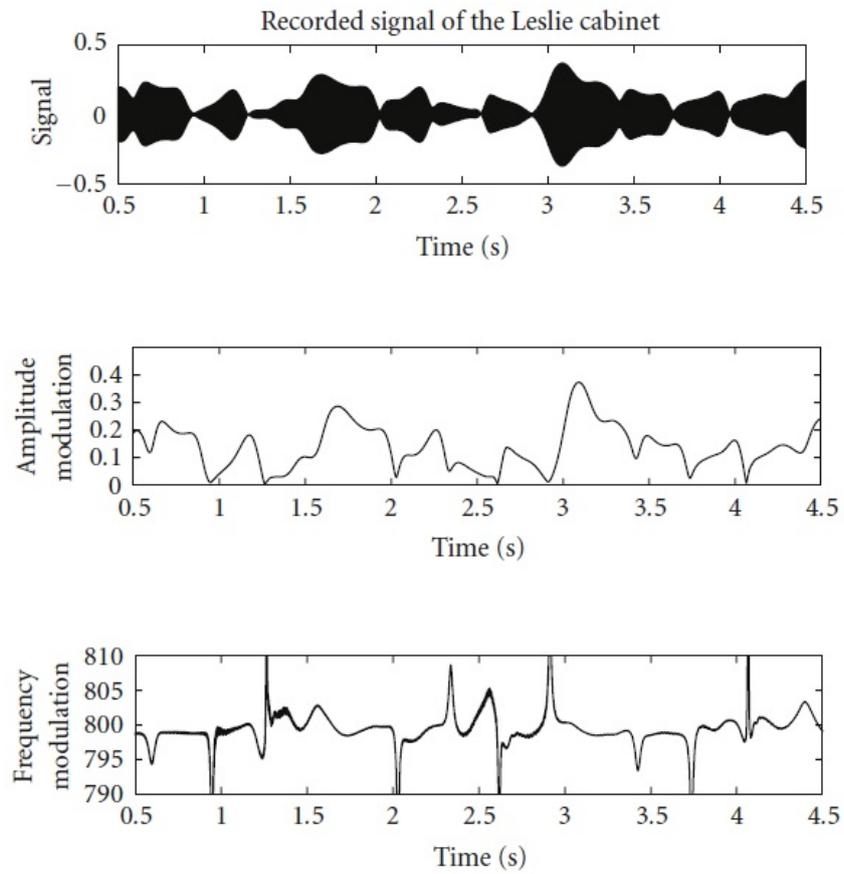


Figura 2.4: Risultati ottenuti dalla registrazione del vero Leslie.

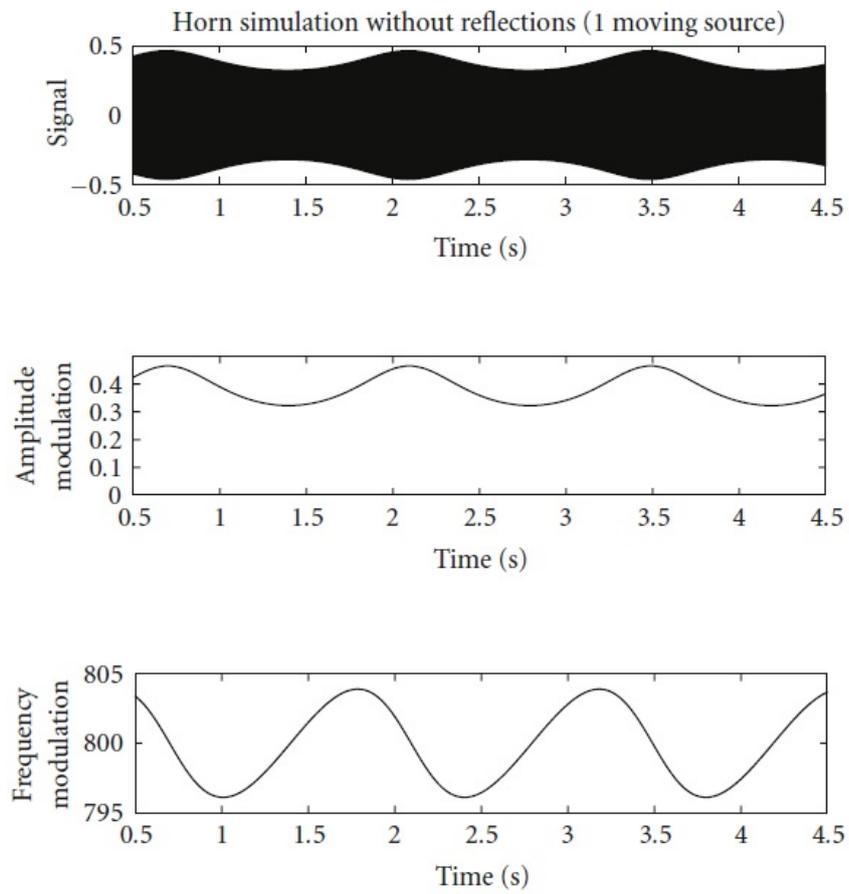


Figura 2.5: Risultati ottenuti dalla registrazione dell'uscita del simulatore in cui sono state escluse le sorgenti immagine.

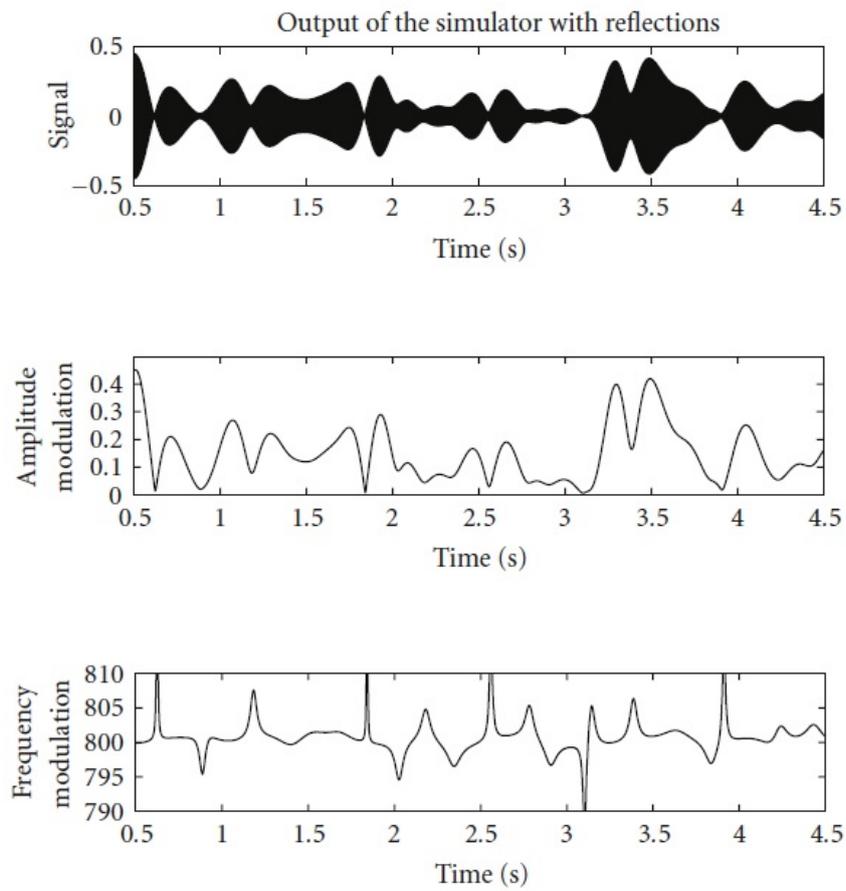


Figura 2.6: Risultati ottenuti dalla registrazione dell'uscita del simulatore con le sorgenti immagine.

Capitolo 3

Modello proposto

Il modello proposto in questo lavoro è basato su quello descritto nel capitolo 2. La modulazione in ampiezza è stata modificata per tener conto della direzionalità delle sorgenti sonore. Le dinamiche della tromba e del deflettore, durante i cambi di velocità, sono state modellate accuratamente. Infine, sono state testate diverse strategie per la modellazione dell'amplificatore.

3.1 Descrizione del modello

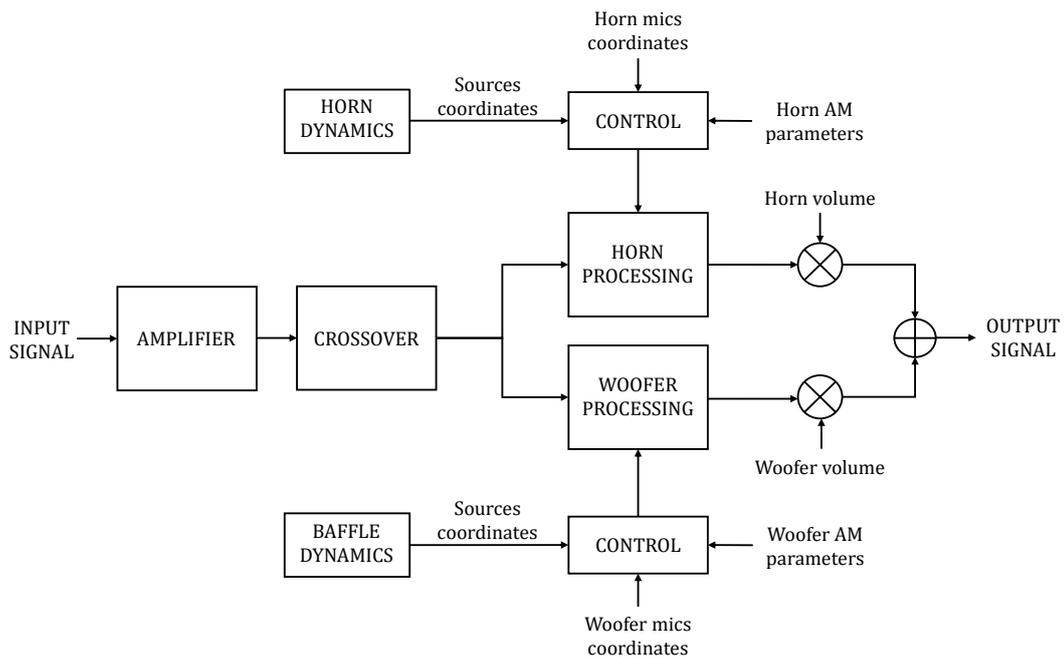


Figura 3.1: Schema a blocchi del modello del Leslie proposto in questo lavoro.

In Figura 3.1 è mostrato lo schema a blocchi del modello proposto in questo lavoro. Il segnale d'ingresso viene processato in successione dall'amplificatore e dal crossover. L'uscita passa-alto del crossover è inviata al blocco di elaborazione del segnale della tromba, mentre l'uscita passa-basso è inviata al blocco di elaborazione del segnale del woofer. Questi blocchi processano il segnale secondo lo schema proposto in [2],

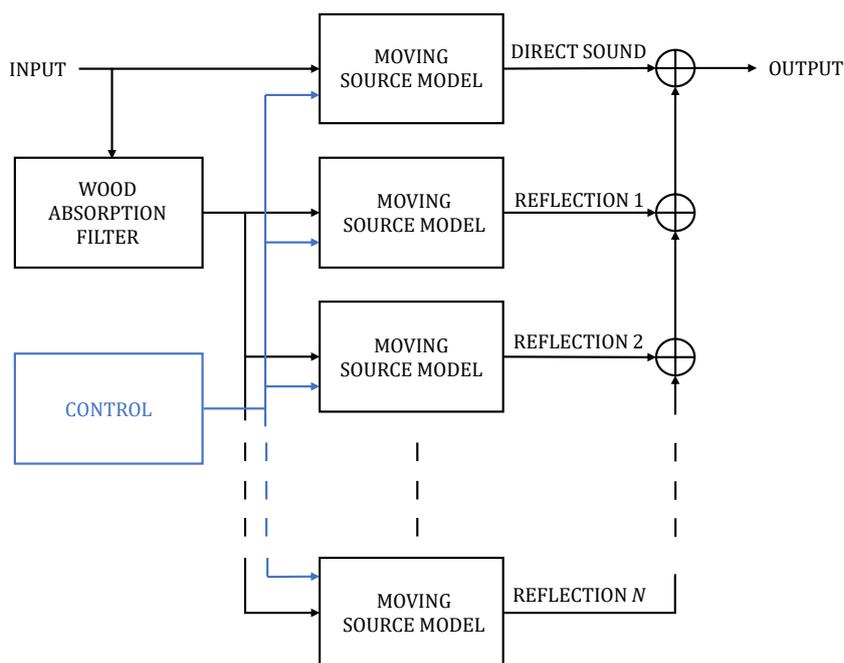


Figura 3.2: Schema interno dei blocchi di elaborazione dei segnali della tromba e del woofer.

descritto nel capitolo 2. In questo caso, lo schema è mostrato in Figura 3.2, in cui le uscite da inviare alla riverberazione non sono state considerate.

I blocchi relativi alla dinamica dei diffusori calcolano le frequenze di rotazione istantanee della tromba e del deflettore, che vengono usate per incrementare gli angoli posizionali delle sorgenti, da cui si ricavano le loro coordinate.

Il sistema di ascolto è pensato per essere stereofonico: si considerano due coppie di microfoni differenti per la ripresa del suono della tromba e del woofer, in modo da poter regolare separatamente l'immagine stereo delle due sorgenti. Non è stato considerato il crosstalk tra il suono emesso dalla tromba e i microfoni del woofer, e viceversa.

Per il woofer, oltre alle sorgenti immagine dovute alle riflessioni sulle pareti del cabinet, è stata considerata una sorgente virtuale aggiuntiva per simulare il suono diretto, non modulato, che viene emesso dal retro del deflettore. Questa sorgente viene processata come le altre con la differenza che è fissa, quindi i coefficienti del filtro di assorbimento dell'aria, così come la delay line e il guadagno per la regolazione del livello sonoro, non sono modulati. Come posizione per questa sorgente si considera il centro di rotazione del woofer.

Nelle sezioni seguenti si descriveranno in dettaglio i blocchi costitutivi del modello.

3.2 Modello di sorgente in movimento

In questa sezione si vedranno diverse alternative per un'implementazione campione-per-campione del modello di sorgente in movimento. Nel modello ogni coppia sorgente-microfono è caratterizzata dalla seguente relazione:

$$m(t) = [s(t - \tau(t)) \cdot g_{AM}(x(t - \tau(t)))] * f_{air}(x(t - \tau(t))) \quad (3.1)$$

dove m è il segnale ripreso dal microfono, s è il segnale generato dalla sorgente, x è la distanza sorgente-microfono, τ è il tempo di propagazione del suono, g_{AM} è il guadagno della modulazione in ampiezza, f_{air} è la risposta impulsiva del filtro dell'aria e $(*)$ è l'operatore di convoluzione. Nel caso di una sorgente in movimento rispetto ad un microfono fisso, il tempo di propagazione è espresso come segue:

$$\tau(t) = \frac{x(t - \tau(t))}{c} = \frac{\|\mathbf{M} - \mathbf{S}(t - \tau(t))\|}{c} \quad (3.2)$$

dove \mathbf{M} e \mathbf{S} denotano, rispettivamente, i vettori posizione del microfono e della sorgente, e c è la velocità del suono.

Il ritardo dovuto al tempo di propagazione del suono può essere simulato con una delay line variabile. Un'implementazione possibile per la delay line consiste nell'utilizzare un buffer con un puntatore di scrittura fisso e un puntatore di lettura variabile (Figura 3.3). Di seguito è riportato uno pseudocodice come esempio.

```
w = 1;
for n = 1:length(in)
    D = calcola_delay();
    buffer(w) = in(n);
    r = w + D;
    out = buffer(r);
    buffer = buffer(0 buffer(end-1));
end
```

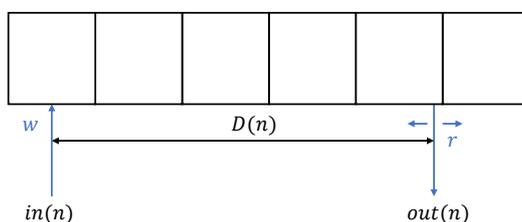


Figura 3.3: Delay line con puntatore di scrittura fisso e puntatore di lettura variabile.

In ogni istante di campionamento si effettuano le seguenti operazioni: il campione d'ingresso corrente viene scritto nella prima posizione del buffer; l'uscita viene letta nella posizione indicata dal puntatore di lettura, ad una distanza pari a D campioni dal puntatore di scrittura; infine, tutti gli elementi del buffer vengono spostati di

Capitolo 3 Modello proposto

una posizione verso destra. Si noti che, in generale, D è un numero frazionale e l'operazione di lettura deve prevedere un'interpolazione tra diverse posizioni del buffer. L'uscita della delay line può quindi essere scritta, in forma matematica, come:

$$out(n) = in(n - D(n)) \quad (3.3)$$

dove n denota l'indice di tempo discreto. Per simulare correttamente lo spostamento della sorgente, il valore del delay in campioni deve essere calcolato come:

$$D(n) = \frac{\tau(n \cdot T_s)}{T_s} = \frac{\|\mathbf{M} - \mathbf{S}(n \cdot T_s - \tau(n \cdot T_s))\|}{c \cdot T_s} \quad (3.4)$$

dove T_s è il tempo di campionamento. Il valore del delay relativo alla posizione della sorgente nell'istante $t - \tau(t)$ può essere calcolato utilizzando l'algoritmo iterativo proposto in [5], riportato di seguito.

```

 $\tau = 1/c \cdot \|\mathbf{M} - \mathbf{S}(t)\|$ 
 $\tau_{old} = -1$ 
 $iter = 0$ 
while  $|\tau_{old} - \tau| > \epsilon$  &  $iter < iter_{max}$  do
     $\tau_{old} = \tau$ 
     $\tau = 1/c \cdot \|\mathbf{M} - \mathbf{S}(t - \tau)\|$ 
     $iter = iter + 1$ 
end while

```

Si noti che, per poter applicare l'algoritmo, è necessario memorizzare la traiettoria della sorgente almeno per alcuni istanti di tempo passati.

In Figura 3.4 e Figura 3.5 sono mostrati gli schemi a blocchi delle due possibili implementazioni del modello di sorgente in movimento che utilizzano una delay line con puntatore di lettura variabile. La differenza sta nel fatto che, se il filtraggio e la modulazione di ampiezza vengono effettuati prima dell'inserimento del campione nella delay line, i coefficienti del filtro e il guadagno AM devono essere ricavati dalla distanza sorgente-ascoltatore corrente. Viceversa, se queste operazioni sono effettuate sul campione in uscita dalla delay line, il calcolo dei coefficienti e del guadagno va fatto utilizzando la distanza sorgente-ascoltatore all'istante $t - \tau(t)$.

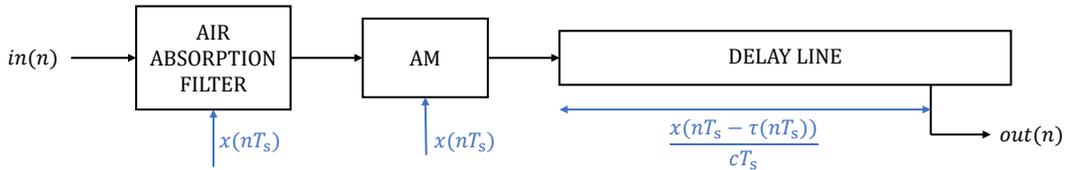


Figura 3.4: Implementazione del modello di sorgente in movimento con delay line con lettura variabile. Le operazioni di filtraggio e AM sono effettuate sul campione in ingresso alla delay line.

L'utilizzo di un metodo iterativo per il calcolo del delay non è desiderabile in

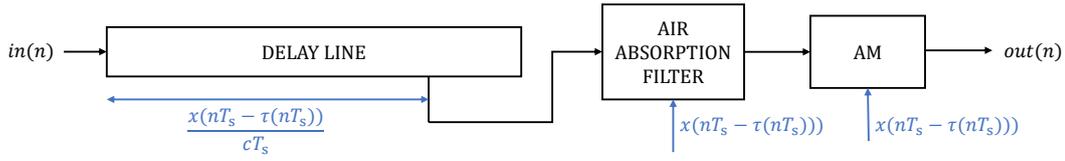


Figura 3.5: Implementazione del modello di sorgente in movimento con delay line con lettura variabile. Le operazioni di filtraggio e AM sono effettuate sul campione in uscita dalla delay line.

un'implementazione real-time. Un'alternativa è quella di approssimare l'espressione del tempo di propagazione come:

$$\tau(t) \approx \frac{x(t)}{c} = \frac{\|\mathbf{M} - \mathbf{S}(t)\|}{c} \quad (3.5)$$

Questa approssimazione è valida in quanto le velocità delle sorgenti del Leslie

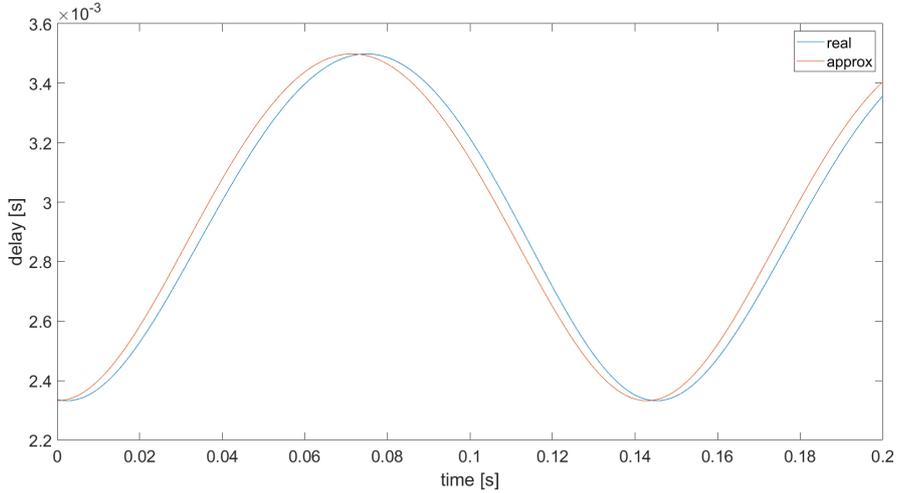


Figura 3.6: Confronto tra il ritardo di propagazione reale (Eq. 3.4) e quello ottenuto dal calcolo approssimato (Eq. 3.5) nel caso di una sorgente che ruota con raggio di 20 cm e frequenza di 7 Hz attorno ad un punto a distanza 1 m da un ascoltatore.

sono molto inferiori alla velocità del suono. Di conseguenza, $S(t - \tau(t)) \simeq S(t)$. In Figura 3.6 è mostrato l'errore che si commette approssimando il tempo di propagazione per una sorgente rotativa dello stesso tipo della tromba del Leslie alla massima frequenza di rotazione. Lo schema dell'implementazione può essere quindi modificato in modo che tutte le operazioni (delay line, filtraggio e AM) siano controllate dalla distanza corrente tra sorgente e microfono. In questo caso, se si utilizza lo schema di Figura 3.4 (filtraggio e AM prima della delay line) si commette un errore solamente sullo shift in frequenza applicato al segnale. Con lo schema di Figura 3.5 (filtraggio e AM dopo della delay line) si commette invece un errore su tutte e tre le operazioni. La seconda implementazione è comunque preferibile nel caso di sorgenti multiple perché

consente di utilizzare un solo buffer con molteplici puntatori di lettura (Figura 3.7), risparmiando così memoria. Infine, bisogna notare che l'approssimazione del tempo di propagazione comporta un errore complessivo maggiore nel caso di più sorgenti. Questo perché ogni sorgente introduce un errore differente, che dipende dalla sua posizione rispetto al microfono.

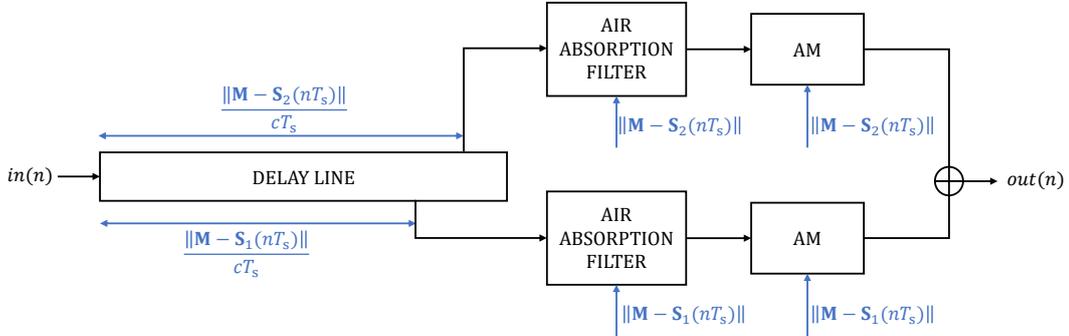


Figura 3.7: Modello di due sorgenti in movimento con lettura variabile della delay line.

Un approccio differente, che permette di implementare un modello corretto (dal punto di vista fisico) senza la necessità di calcolare il tempo di propagazione con la Eq. 3.4, è quello proposto in [4], in cui la delay line è implementata utilizzando un puntatore di scrittura variabile e un puntatore di lettura fisso (Figura 3.8). In questo caso, l'operazione può essere descritta con il seguente pseudocodice:

```

r = length(buffer);
for n = 1:length(in)
    D = calcola_delay();
    w = r - D;
    buffer(w) = buffer(w) + in(n);
    out = buffer(r);
    buffer = buffer(0 buffer(end-1));
end

```

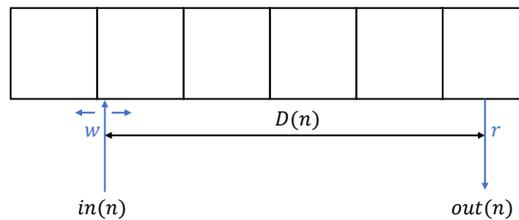


Figura 3.8: Delay line con puntatore di scrittura variabile e puntatore di lettura fisso.

L'uscita della delay line si esprime quindi come:

$$out(n + D(n)) = x(n) \quad (3.6)$$

ovvero, il campione d'ingresso corrente viene letto in uscita dopo un tempo pari a $D(n) \cdot T_s$ secondi. Di conseguenza, per simulare correttamente il tempo di propagazione è sufficiente calcolare il delay come:

$$D(n) = \frac{\|\mathbf{M} - \mathbf{S}(n \cdot T_s)\|}{c \cdot T_s} \quad (3.7)$$

Quando si implementa la delay line in questo modo, lo schema del modello di sorgente in movimento diventa quello di Figura 3.9, in cui le operazioni di filtraggio e AM devono essere eseguite prima dell'inserimento del campione nella delay line. Inoltre, è possibile usare un singolo buffer con molteplici puntatori di scrittura per simulare il movimento di più sorgenti, con differenti percorsi di propagazione. Di seguito, si vedrà però che l'implementazione di una delay line variabile con scrittura interpolata è più complessa rispetto a quella che utilizza la lettura interpolata.

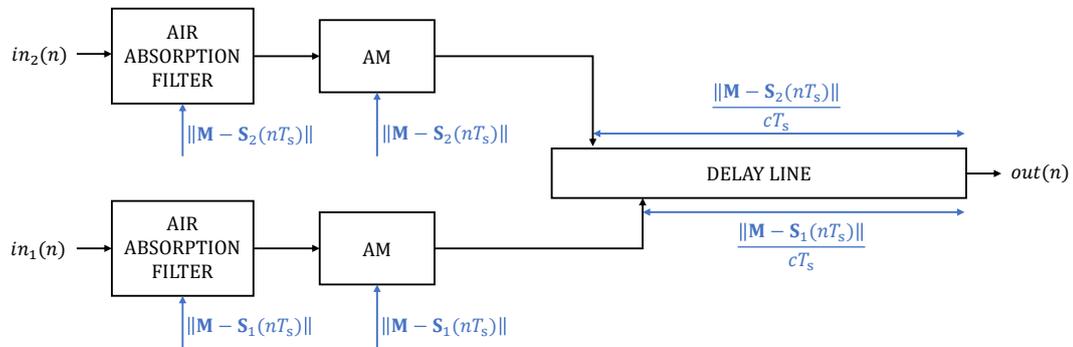


Figura 3.9: Modello di due sorgenti in movimento con scrittura variabile della delay line.

3.2.1 Delay line con lettura interpolata

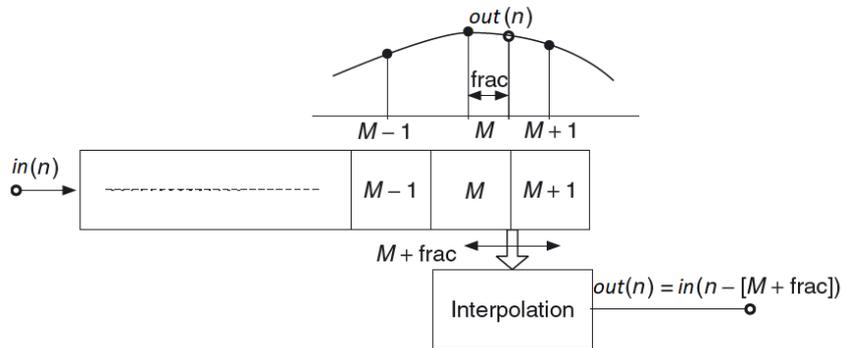


Figura 3.10: Schema dell'implementazione della lettura interpolata di una delay line.

In Figura 3.10 è mostrato lo schema generale per la realizzazione di una delay line con lettura interpolata. Il problema dell'implementazione di valori di delay frazionari è già stato affrontato in letteratura ([8]). In questo lavoro si è scelto di

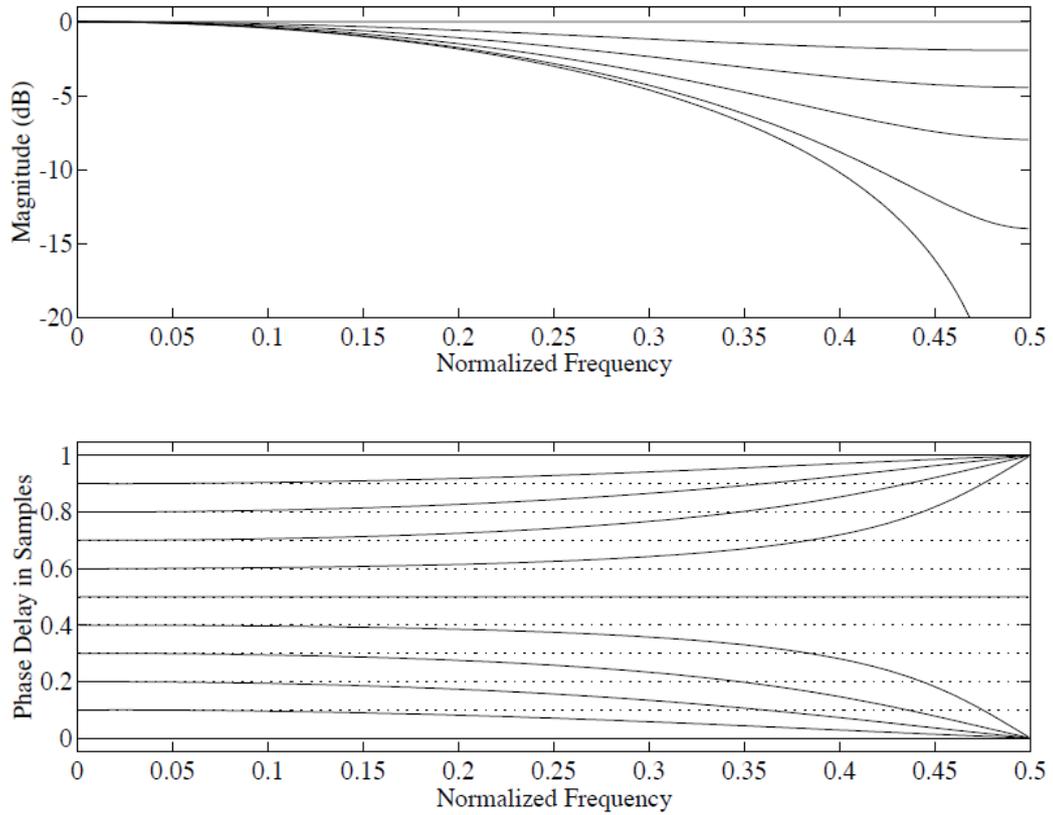


Figura 3.11: Risposta in ampiezza (in alto) e in fase (in basso) di un interpolatore lineare per undici diversi delay frazionali ($D = 0, 0.1, 0.2, \dots, 1.0$). Si noti che ci sono solo sei curve nel grafico in alto perché le risposte in ampiezza sono uguali per le parti frazionali d e $1 - d$. Nel grafico in basso, le linee tratteggiate indicano la risposta in fase ideale.

utilizzare l'interpolazione lineare per ridurre il costo computazionale. Questo tipo di interpolazione ha l'effetto di un filtraggio passa-basso sul segnale (Figura 3.11). Questo non rappresenta un problema visto che il Leslie è utilizzato prevalentemente con segnali di organo, che non hanno un elevato contenuto in frequenza al di sopra dei 10 kHz. Facendo riferimento a Figura 3.10, nel caso di interpolazione lineare l'uscita della delay line si esprime come:

$$out(n) = in(n - (M + 1)) \cdot frac + x(n - M) \cdot (1 - frac) \quad (3.8)$$

dove M è la parte intera del delay e $frac$ la parte frazionale.

Di seguito è mostrato un esempio di codice C++ per l'implementazione efficiente di una delay line con buffer circolare, letto con interpolazione lineare. Il metodo `process` può essere richiamato ad ogni istante di campionamento per scrivere e leggere la delay line. La dimensione del buffer deve essere scelta in base al massimo valore di delay (in campioni) che si deve simulare nell'applicazione.

```

struct delayLine {
    float buffer[BUFFER_SIZE] = {};
    float wIdx;

    delayLine() {
        wIdx = 0;
    }

    float process(float in, float D) {
        int rIdx0, rIdx1;
        float out;
        float rIdx = (float)wIdx - D;
        rIdx0 = (int)floorf(rIdx);
        rIdx1 = rIdx0 + 1;
        float frac = rIdx - (float)rIdx0;

        if (rIdx < 0) rIdx += BUFFER_SIZE;
        if (rIdx0 < 0) rIdx0 += BUFFER_SIZE;
        if (rIdx1 < 0) rIdx1 += BUFFER_SIZE;

        buffer[wIdx] = in; // scrittura del buffer
        wIdx++;
        if (wIdx == BUFFER_SIZE) wIdx = 0;

        // lettura del buffer
        out = (1.f - frac)*buffer[rIdx0] + frac*buffer[rIdx1];

        return out;
    }
}

```

3.2.2 Delay line con scrittura interpolata

La scrittura interpolata di una delay line è stata proposta da Välimäki ed è indicata con il termine *deinterpolazione*. In [9] viene spiegato come un deinterpolatore può essere ottenuto da un interpolatore FIR trasponendo la sua struttura e invertendo la sequenza dei coefficienti. Queste operazioni portano ad una struttura che è una sequenza di addizioni cumulative nella delay line.

L'implementazione utilizzata in questo lavoro è stata sviluppata a partire dall'approccio utilizzato in [10], in cui l'interpolazione è eseguita sulla sequenza in ingresso alla delay line. Uno schema dell'implementazione è mostrato in Figura 3.12. I coefficienti moltiplicativi h_0 e h_1 sono calcolati in modo da effettuare l'interpolazione tra i punti $(phase(n); in(n))$ e $(phase(n-1); in(n-1))$. Si noti che se il valore del delay cresce nel tempo l'incremento della variabile $phase$ è, in ogni istante di campionamento, superiore a 1. Di conseguenza, in alcuni istanti di campionamento è necessario scrivere su più posizioni del buffer in modo da non lasciare spazi vuoti (ad esempio, nel caso di un incremento costante pari ad 1.5 si devono effettuare 3 scritture ogni 2 istanti di campionamento), aumentando così il costo computazionale. Di seguito è riportato un esempio di codice C++ per l'implementazione di una delay line con buffer circolare, scritto con interpolazione lineare. Grazie al ciclo `while`, quando necessario, si effettuano scritture successive per riempire tutte le posizioni del buffer.

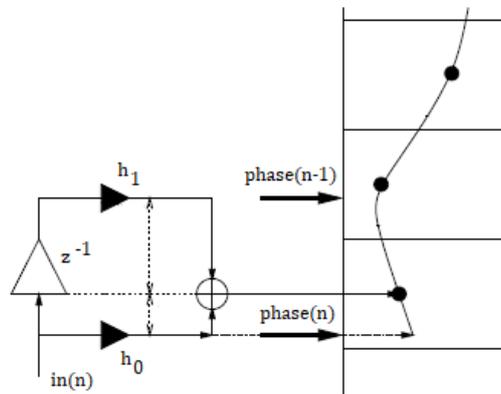


Figura 3.12: Schema dell'implementazione della scrittura interpolata (nel caso di interpolazione lineare) di una delay line.

```

struct delayLine {
    float buffer [BUFFER_SIZE] = {};
    float rIdx, delay_old, phase_old;

    delayLine(float initial_delay) {
        rIdx = 0;
        delay_old = phase_old = initial_delay;
    }
}
    
```

```

float process(float in, float delay) {
    int ph = fph + 1;
    float phase_inc = delay - delay_old + 1.f;
    float phase = phase_old + phase_inc;
    fph = (int) floorf(phase);

    // scrittura del buffer
    int wIdx;
    float h0, h1;
    while (ph <= fph) {
        wIdx = ph;
        if (wIdx > (BUFFER_SIZE-1)) wIdx -= BUFFER_SIZE;
        h0 = ((float)ph - phase_old)/phase_inc;
        h1 = 1 - h0;
        buffer[wIdx] = h1*in_old + h0*in;
        ph++;
    }

    if (fph > (BUFFER_SIZE-1)) {
        fph -= BUFFER_SIZE;
        phase -= BUFFER_SIZE;
    }

    phase_old = phase;
    delay_old = delay;
    in_old = in;

    // lettura del buffer
    rIdx++;
    if (rIdx == BUFFER_SIZE) rIdx = 0;
    out = buffer[rIdx];

    return out;
}
}

```

3.3 Crossover

Come indicato in [1], il Leslie ha un crossover passivo con pendenza di 12 dB/ottava (40 dB/decade). Per questo, il crossover è stato simulato con dei filtri passa-alto e passa-basso di Butterworth del secondo ordine con frequenza di taglio a 800 Hz. Il metodo di progetto utilizzato per i filtri è quello indicato in [11]. In Figura 3.13 sono mostrate le risposte in ampiezza ottenute. Si noti che per un crossover di Butterworth la risposta complessiva non è piatta, ma si ha un guadagno di 3 dB in corrispondenza della frequenza di incrocio.

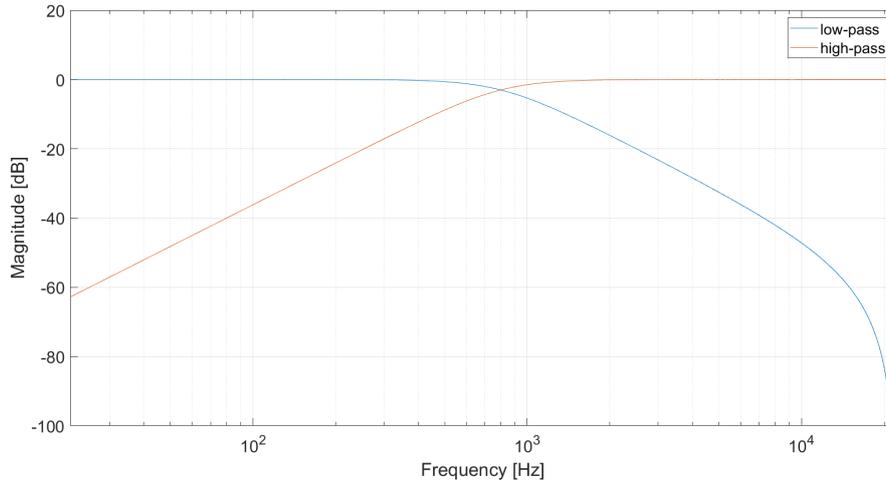


Figura 3.13: Risposte in ampiezza passa-basso e passa-alto del crossover.

3.4 Filtro di assorbimento dell'aria

Il filtro di assorbimento dell'aria è stato progettato utilizzando un filtro IIR high-shelving del secondo ordine, che rappresenta un buon compromesso tra accuratezza percettiva e performance in tempo reale. La tipologia di filtro utilizzata è quella descritta in [12], la quale è stata originariamente progettata per equalizzatori parametrici, che è caratterizzata da due parametri di controllo: guadagno e frequenza di taglio. Per simulare l'assorbimento dell'aria è necessario ricavare una relazione tra questi parametri e la distanza tra sorgente e ascoltatore. Per farlo, in [2] è stata utilizzata una procedura di ottimizzazione basata sul metodo di minimizzazione dei minimi quadrati, che restituisce i valori per il guadagno e la frequenza di taglio che minimizzano $\|A(f) - H(f)\|^2$, dove $A(f)$ è la funzione di trasferimento dell'aria da approssimare, calcolata per una certa distanza sorgente-ascoltatore x , con le formule indicate in [13], e $H(f)$ è la funzione di trasferimento del filtro high-shelving. Come risultato, si è visto che la frequenza di taglio dipende debolmente dalla distanza x ed è stata settata al valore di 10 kHz. Questo porta ad un unico parametro di controllo: il guadagno G . La relazione ricavata tra la distanza, in metri, e il guadagno, in dB, è la seguente:

$$G(\text{dB}) = -0.5x(\text{m}) \quad (3.9)$$

Le funzioni di trasferimento calcolate e le risposte in ampiezza del filtro simulate sono riportate in Figura 3.14 per distanze fino a 50 metri. Anche se la simulazione differisce dalla realtà, specialmente nel range in alta frequenza, in [2] si garantisce che, dal punto di vista percettivo, il risultato prodotto è soddisfacente. Inoltre, il fattore -0.5 , che lega la distanza al guadagno del filtro, può essere variato per aggiustare l'effetto della variazione del timbro.

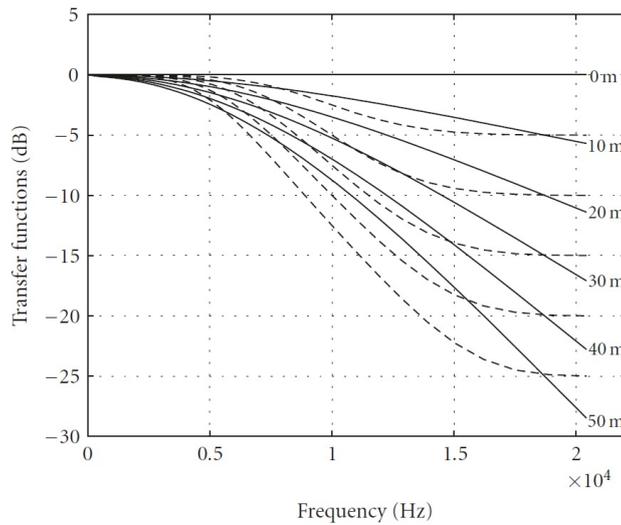


Figura 3.14: Funzioni di trasferimento dell'aria (linee continue) e funzioni di trasferimento simulate (linee tratteggiate) ottenute dal processo di ottimizzazione per varie distanze tra sorgente e ascoltatore. Le funzioni di trasferimento dell'aria sono state calcolate considerando una temperatura di 20° , pressione atmosferica pari a 1013 HPa e 45% di igrometria. La frequenza di taglio del filtro simulato è stata settata a 10 kHz e il guadagno è stato calcolato usando la Eq. 3.9.

3.5 Filtro di assorbimento del legno

Il problema della modellazione della riflessione di un'onda sonora sulla superficie di un materiale è complesso. Il comportamento temporale e spettrale del suono riflesso come funzione dell'angolo di incidenza, i fenomeni di scattering e diffrazione, ecc., rendono infatti impossibile l'uso di modelli numerici accurati sotto tutti gli aspetti.

In questo lavoro, il filtro di assorbimento del legno è stato progettato utilizzando il metodo proposto in [3], basato sull'utilizzo dei coefficienti di assorbimento del materiale. Con questa tecnica si ignora il problema della dipendenza della riflessione dall'angolo di incidenza sulla superficie e, inoltre, non si presta attenzione al fatto che i dati del materiale misurati in campo diffuso o con tubo di impedenza dovrebbero essere trattati diversamente.

La più comune caratterizzazione acustica della superficie dei materiali è data dai coefficienti di assorbimento, forniti in bande di ottava: 125, 250, 500, 1000, 2000, 4000 Hz. Da questi dati è possibile ricavare la risposta all'impulso della riflessione $r(t)$ o la funzione di trasferimento, a valori complessi (riflettanza), $R(j\omega) = F\{r(t)\}$, dove F indica la trasformata di Fourier. Visto che il coefficiente di assorbimento è dato dal rapporto energetico dell'energia assorbita e dell'energia incidente, la relazione tra i coefficienti di assorbimento $a(\omega)$ e $R(j\omega)$ è data da:

$$a(\omega) = 1 - |R(j\omega)|^2 \quad (3.10)$$

Quindi è possibile utilizzare la relazione seguente per ottenere i valori della riflettanza dai coefficienti di assorbimento:

$$|R(j\omega)| = \sqrt{1 - a(\omega)} \quad (3.11)$$

Si noti che un valore negativo all'interno della radice non si verifica mai in pratica.

Assumendo che la propagazione del suono nell'aria e la riflessione sulla parete si comportino linearmente e che il sistema sia tempo invariante, è possibile utilizzare un filtro digitale per modellare l'assorbimento della parete. Il progetto del filtro consiste semplicemente nel cercare una corrispondenza ottimale dei dati con un determinato tipo di filtro e ordine, dato un criterio di errore. Visto che dai coefficienti di assorbimento è possibile ricavare solamente delle specifiche sulla risposta in ampiezza del filtro in bande di ottava, c'è il problema aggiuntivo di trovare delle caratteristiche appropriate per la fase, così come quello dell'interpolazione dei dati in modo acusticamente significativo.

I coefficienti di assorbimento del legno sono stati presi da [14], in cui viene fatto uno studio sulle proprietà di assorbimento di vari tipi di legno, tra cui il legno di quercia, utilizzato per la costruzione dei cabinet di alcuni modelli di Leslie. Come si nota in Figura 3.15, i legni presi in esame mostrano un maggiore assorbimento in alta frequenza, con un picco in corrispondenza dei 2 kHz.

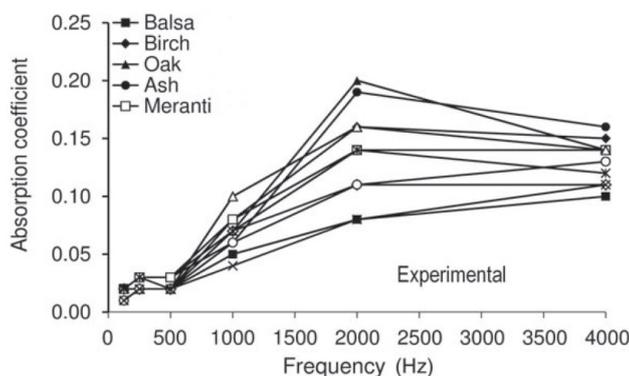


Figura 3.15: Coefficienti di assorbimento misurati per alcuni tipi di legno.

Dai coefficienti di assorbimento della quercia, utilizzando la Eq. 3.11, sono stati ricavati i punti della risposta in ampiezza da approssimare. È stato progettato un filtro di tipo IIR utilizzando la funzione `fdesign.arbmag` del Matlab, la quale prende in ingresso l'ordine del numeratore del filtro, l'ordine del denominatore, il vettore delle frequenze, contenente i valori di frequenza per i quali vengono forniti i valori della risposta in ampiezza, e il vettore delle ampiezze, contenente i corrispondenti punti della risposta in ampiezza, e restituisce i coefficienti del filtro ordinati in sezioni del secondo ordine. Non c'è stato bisogno di fare una interpolazione dei dati visto che l'algoritmo funziona bene anche con pochi punti. Si è visto che la risposta desiderata viene approssimata adeguatamente da un filtro IIR del quarto ordine. Il risultato del

progetto è mostrato in Figura 3.16. Il filtro risultante è a fase minima, come si vede in Figura 3.17, caratteristica desiderata quando nel progetto non vengono fornite specifiche sulla fase.

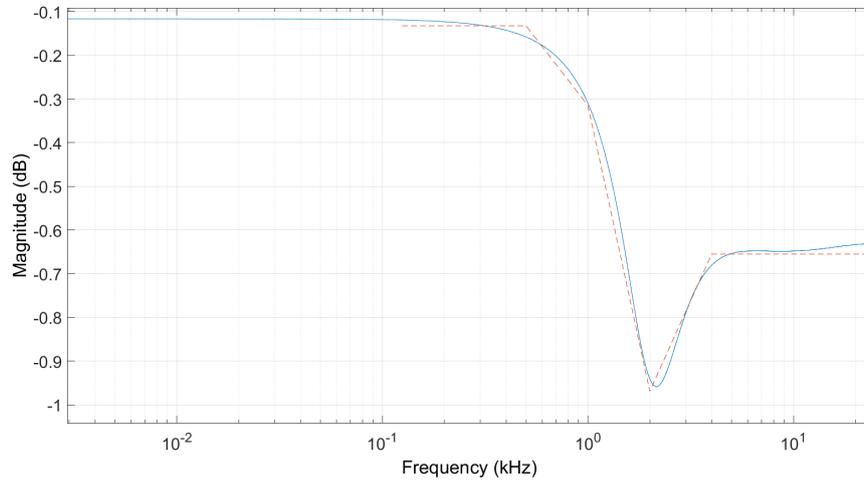


Figura 3.16: Risposta in ampiezza del filtro progettato (linea continua) e risposta in ampiezza da approssimare (linea tratteggiata).

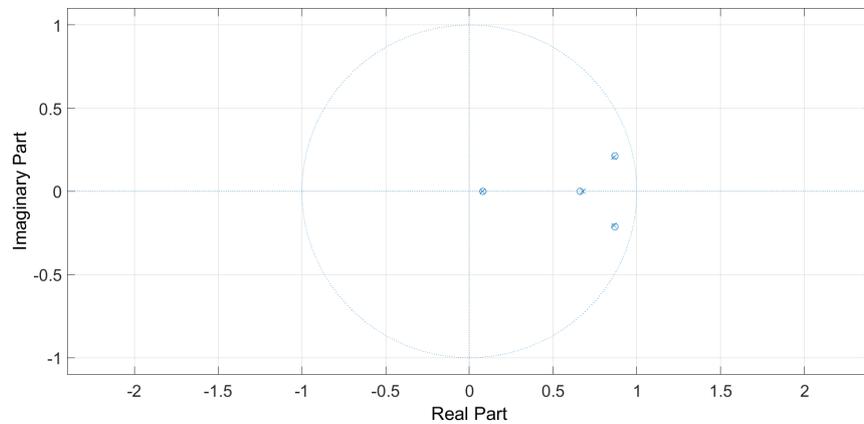


Figura 3.17: Poli e zeri del filtro progettato.

3.6 Modulazione in ampiezza

Come si è visto in precedenza, nel modello proposto in [2] la modulazione in ampiezza consiste nel moltiplicare il segnale generato dalla sorgente per un guadagno $g_{AM} = 1/x$, dove x indica la distanza tra la sorgente e l'ascoltatore. In questo lavoro, la modulazione in ampiezza è stata leggermente complicata in modo da

tener conto della direttività delle sorgenti. Infatti, mentre il pattern di radiazione della tromba è quasi omnidirezionale (come spiegato nel capitolo 1), questo può essere utile per l'elaborazione del segnale del woofer, visto che il deflettore presenta una direttività maggiore. L'algoritmo proposto è controllato da due parametri chiamati *directivityAngle* e *directivityLevel*. Si faccia riferimento a Figura 3.18, che mostra una sorgente sonora S che ruota intorno al punto H, con un raggio di rotazione r , e un microfono. Il parametro *directivityAngle* setta l'ampiezza

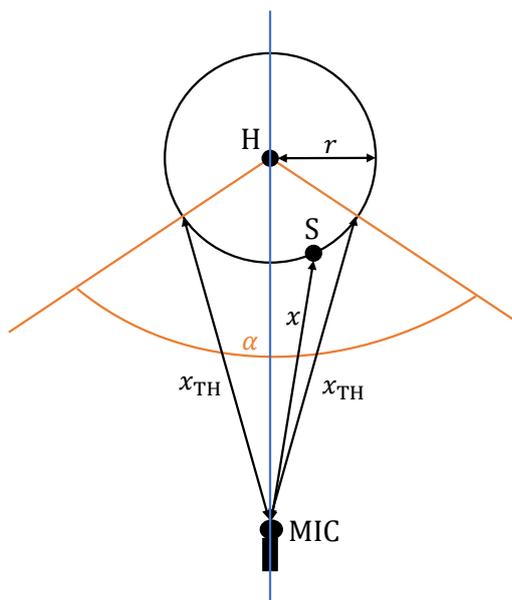


Figura 3.18: Schema di una sorgente sonora S che ruota intorno ad un punto H ad una certa distanza da un microfono.

dell'angolo α in Figura 3.18, ovvero l'angolo centrato sulla retta che congiunge il centro di rotazione della sorgente al microfono. In base all'ampiezza di questo angolo è possibile calcolare una distanza di soglia, che in Figura 3.18 è indicata come x_{TH} . L'idea di base dell'algoritmo è che se la distanza corrente tra sorgente e microfono è inferiore a x_{TH} (la sorgente S si trova all'interno dell'angolo α), allora il volume della sorgente deve essere aumentato rispetto al caso $1/x$. In particolare, in ogni istante di campionamento, il calcolo del guadagno da applicare al segnale viene fatto in questo modo:

$$g_{AM} = \begin{cases} 1/x & \text{se } x > x_{TH} \\ (1/x - 1/x_{TH}) \cdot \text{directivityLevel} + 1/x_{TH} & \text{se } x \leq x_{TH} \end{cases} \quad (3.12)$$

con $\text{directivityLevel} \geq 1$. Si noti che se $\text{directivityLevel} = 1$ si ritorna al caso più semplice in cui il guadagno applicato al segnale è sempre $1/x$. In Figura 3.19 è mostrato l'andamento della modulante al variare dei parametri.

Nel modello ogni sorgente virtuale ha una sua distanza di soglia e quindi un diverso andamento della modulazione. I parametri *directivityAngle* e *directivityLevel* sono

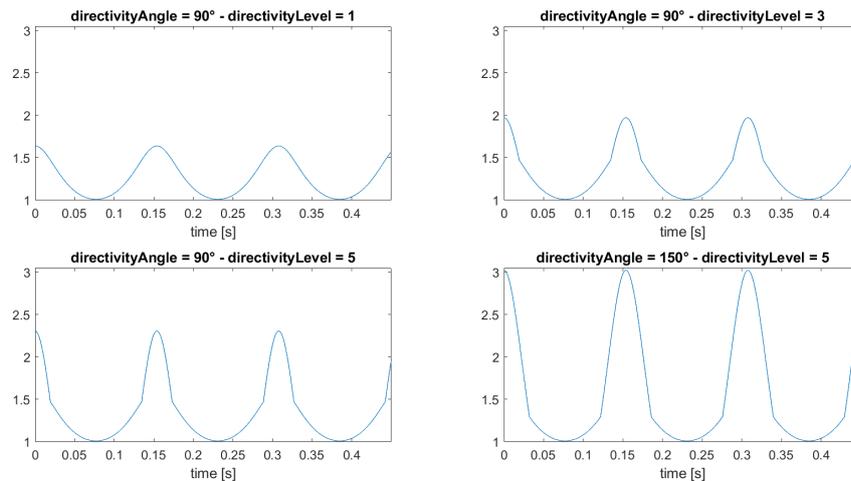


Figura 3.19: Andamento della modulante per differenti valori dei parametri *directivityAngle* e *directivityLevel*. Le curve sono riferite ad una sorgente che ruota con una frequenza di 6.5 Hz e un raggio di 0.19 m attorno ad un punto distante 0.8 m da un microfono.

diversi per la tromba e per il woofer in modo da simulare direttività differenti.

3.7 Dinamica dei diffusori rotanti

Come spiegato nel capitolo 1, il Leslie ha tre modi di funzionamento: *brake*, *slow* e *fast*. Il deflettore e la tromba hanno comportamenti molto diversi durante i cambi di velocità. In particolare, i tempi di transizione della tromba sono molto veloci (circa 1 secondi per tutte le transizioni), mentre i tempi del deflettore sono più lunghi (il deflettore ha una inerzia maggiore rispetto alla tromba) e variano molto da una transizione all'altra. Ad esempio, la transizione *brake* → *slow* dura circa 3 secondi, mentre la transizione *fast* → *brake* dura circa 25 secondi. Per questo, è importante modellare accuratamente l'andamento della velocità di rotazione del deflettore durante tutte le transizioni. Per misurare le transizioni del deflettore, tra le frequenze associate ai modi di funzionamento, è stata inviata in ingresso ad un Leslie 122 una nota stazionaria, utilizzando un organo Hammond B-3, e l'uscita è stata misurata posizionando un microfono davanti al deflettore e scollegando la tromba. Il segnale registrato è stato processato con un rivelatore di involuppo e l'andamento della frequenza di rotazione è stato ottenuto determinando il tempo tra feature successive dell'involuppo (picchi e/o passaggi per lo zero). In Figura 3.20 è mostrato un esempio di segnale registrato, il suo involuppo e i picchi dell'involuppo individuati.

I punti registrati per le frequenze istantanee sono stati utilizzati per fare il fitting di un polinomio, sfruttando la funzione `fit` del Matlab. Il fitting è stato eseguito sulle curve misurate, normalizzate tra 0 e 1. Quindi, prima di applicare l'algoritmo

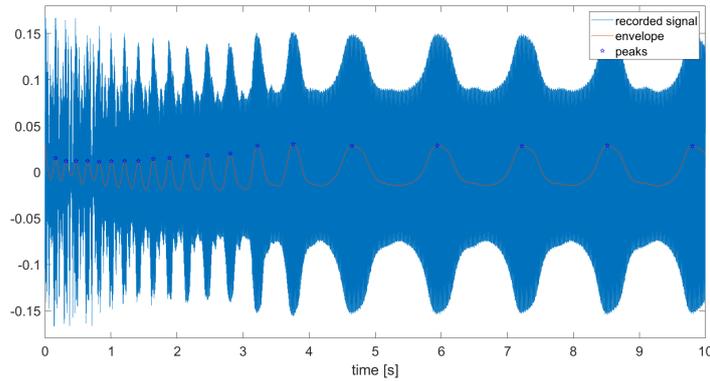


Figura 3.20: Sono mostrati l'andamento nel tempo del segnale registrato per la transizione $fast \rightarrow slow$, l'involuppo calcolato e le posizioni dei picchi dell'involuppo.

di fitting, i dati misurati sono stati riscaldati ed è stato applicato uno shift nel caso delle transizioni $fast \rightarrow slow$ e $slow \rightarrow fast$. Per modellare accuratamente tutte le transizioni è stato necessario utilizzare un polinomio del terzo ordine. I risultati del fitting sono mostrati in Figura 3.21 - Figura 3.26.

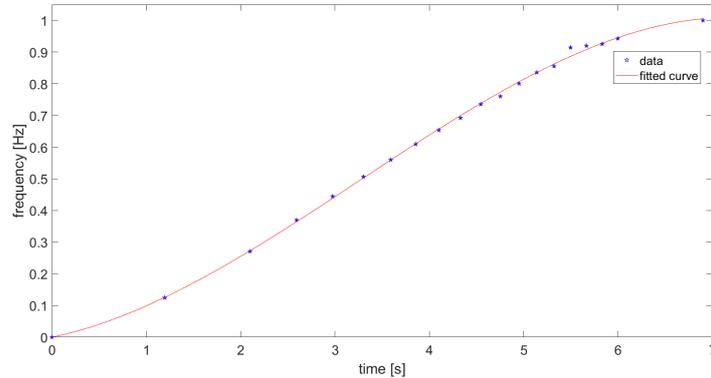


Figura 3.21: Risultato del fitting per la transizione $brake \rightarrow fast$.

Nel modello ogni transizione è caratterizzata da diversi coefficienti del polinomio. Se consideriamo, ad esempio, una transizione da una frequenza f_{start} ad una frequenza f_{stop} , la legge utilizzata per calcolare la frequenza istantanea di rotazione è:

$$f(t) = f_{offset} + f_{delta} \cdot p_T(t) \quad (3.13)$$

dove t è la variabile temporale, $f_{offset} = \min(f_{start}, f_{stop})$, $f_{delta} = |f_{start} - f_{stop}|$ e p_T indica il polinomio relativo alla transizione T .

Il vantaggio di questo modello è che applicando un coefficiente moltiplicativo alla variabile temporale del polinomio è possibile modificare in modo arbitrario la durata della transizione. Per quanto riguarda la tromba, visto che le transizioni hanno

Capitolo 3 Modello proposto

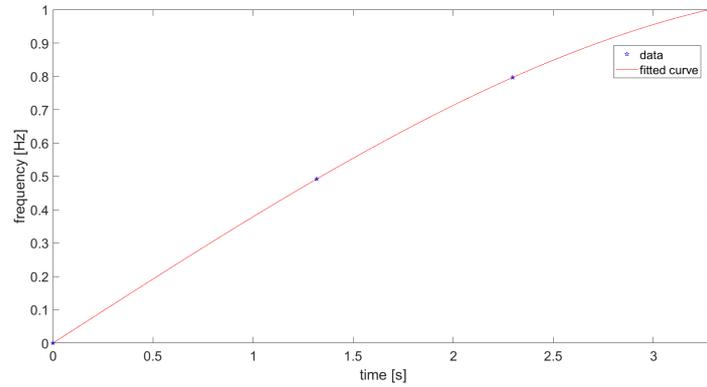


Figura 3.22: Risultato del fitting per la transizione *brake* \rightarrow *slow*.

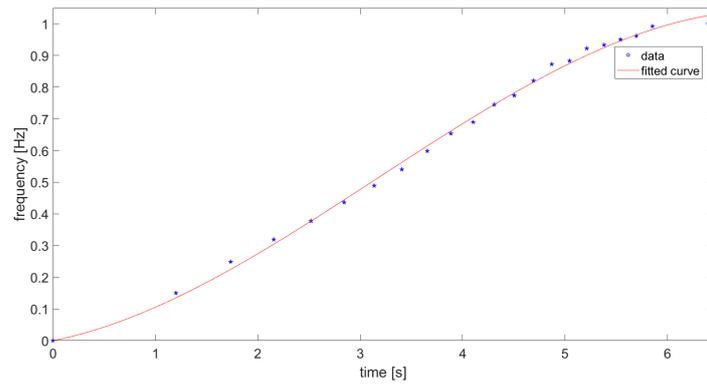


Figura 3.23: Risultato del fitting per la transizione *slow* \rightarrow *fast*.

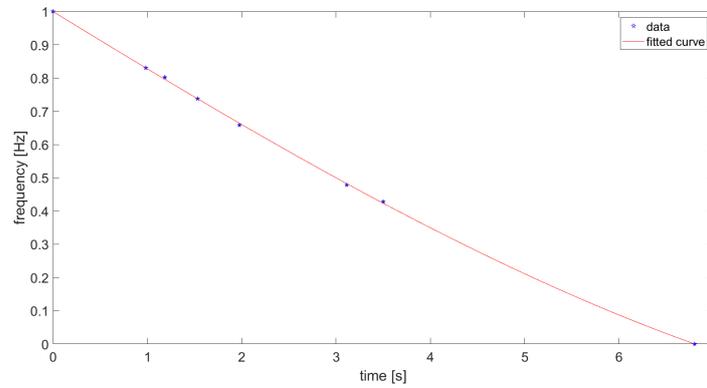


Figura 3.24: Risultato del fitting per la transizione *slow* \rightarrow *brake*.

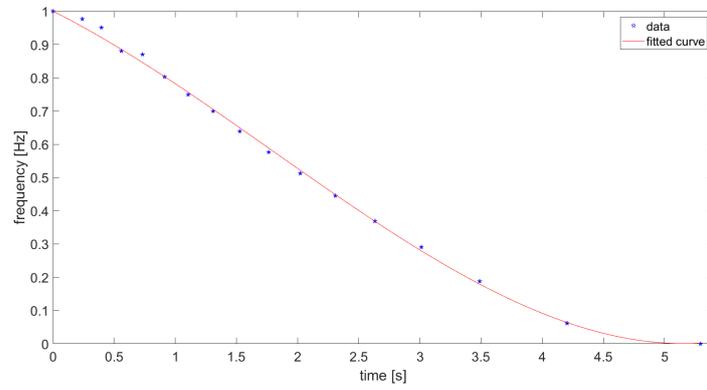


Figura 3.25: Risultato del fitting per la transizione *fast* \rightarrow *slow*.

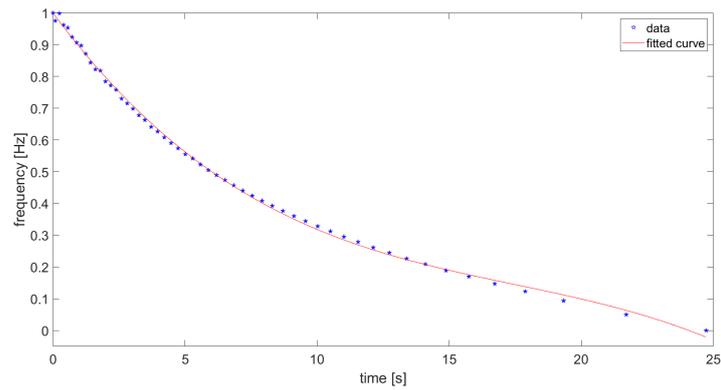


Figura 3.26: Risultato del fitting per la transizione *fast* \rightarrow *brake*.

tutte breve durata, da un punto di vista percettivo non è importante l'andamento della frequenza di rotazione durante la transizione, ma solo il tempo impiegato per completarla. Per questo, le transizioni della tromba sono state modellate utilizzando gli stessi coefficienti del polinomio ottenuti per il deflettore, ma modificando la Eq. 3.13 in questo modo:

$$f(t) = f_{\text{offset}} + f_{\text{delta}} \cdot p_T(a \cdot t) \quad (3.14)$$

dove a è un coefficiente calcolato opportunamente per fare in modo che le transizioni della tromba abbiano la durata desiderata, di circa 1 secondo.

3.8 Amplificatore

In questa sezione si descriveranno le tre tipologie di modelli che sono state testate per l'amplificatore. La prima è un semplice waveshaper mentre le altre due sono dei modelli di Wiener-Hammerstein (Figura 3.27) costituiti dalla cascata di un blocco lineare, un blocco non lineare e un ulteriore blocco lineare. Il blocco non lineare consiste in una funzione non lineare, per generare la distorsione. Il primo filtro determina l'ammontare della distorsione per ogni frequenza: se questo ha una grande attenuazione in una certa regione di frequenza, queste frequenze non produrranno molte armoniche visto che la funzione non lineare non sarà portata in saturazione abbastanza. Questo filtro controlla inoltre la quantità di distorsione di intermodulazione che si crea quando il segnale di ingresso è costituito da più di una componente tonale. Infine, il filtro finale crea le sfumature del segnale distorto andandone a modificare la forma spettrale.

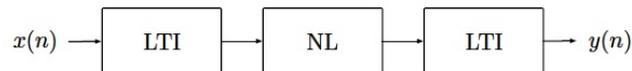


Figura 3.27: Schema a blocchi del modello di Wiener-Hammerstein.

3.8.1 Waveshaper

Il waveshaper proposto è lo stesso utilizzato in [15] ed è basato sull'equazione:

$$y(n) = \frac{1}{\arctan(k)} \cdot \arctan(k \cdot x(n)) \quad (3.15)$$

La quantità di distorsione può essere facilmente regolata variando il valore del parametro k .

L'utilizzo di un waveshaper può essere problematico perché, visto il suo comportamento non lineare, aggiunge al segnale delle componenti in frequenza. Questo potrebbe portare ad avere aliasing, che comporta un effetto sgradevole sul suono.

Se il segnale di ingresso è quello di un organo Hammond, che spesso manca di forti componenti in alta frequenza, e la distorsione è mantenuta ad un livello moderato, l'aliasing non dovrebbe essere un problema.

3.8.2 Modello di Wiener-Hammerstein allenato con backpropagation

Il secondo modello proposto è un modello di Wiener-Hammerstein implementato sulla base di [16]. In questo articolo viene spiegato un approccio innovativo per la realizzazione di alcune tradizionali strutture di filtraggio IIR come reti neurali ricorrenti (RNNs). Si consideri la rete di Elman, una delle prime forme di reti neurali ricorrenti, introdotta nel 1990 per modellare il linguaggio umano. Questa è descritta dalla seguente equazione alle differenze:

$$\mathbf{h}(n) = \sigma_h(\mathbf{W}_h \mathbf{h}(n-1) + \mathbf{U}_h \mathbf{x}(n) + \mathbf{b}_h) \quad (3.16)$$

$$\mathbf{y}(n) = \sigma_y(\mathbf{W}_y \mathbf{h}(n) + \mathbf{b}_y) \quad (3.17)$$

dove \mathbf{h} , \mathbf{y} e \mathbf{x} sono, rispettivamente, strati nascosti, uscite e ingressi del sistema. Le matrici \mathbf{W} , \mathbf{U} sono i pesi allenabili della rete e i vettori \mathbf{b} sono i biases. Le funzioni σ , generalmente non lineari, sono le funzioni di attivazione. Si nota che scegliendo delle funzioni di attivazioni lineari e settando i biases a zero si ottiene essenzialmente un filtro IIR lineare a soli poli:

$$\mathbf{h}(n) = \mathbf{W}_h \mathbf{h}(n-1) + \mathbf{U}_h \mathbf{x}(n) \quad (3.18)$$

$$\mathbf{y}(n) = \mathbf{W}_y \mathbf{h}(n) \quad (3.19)$$

Questo mostra il parallelismo tra RNNs e filtri IIR: sono entrambe strutture ricorsive che operano elemento per elemento su una sequenza di dati numerici, memorizzando informazioni nei loro stati interni. Quindi, è possibile utilizzare i filtri IIR nel contesto del machine learning trattandoli come speciali tipi di RNNs e utilizzando le tecniche e le conoscenze già sviluppate per quest'ultime.

Il training delle RNNs viene effettuato utilizzando la backpropagation-through-time (BPTT), che è un'estensione dell'algoritmo di backpropagation standard derivata dal dispiegamento temporale del grafico computazionale della rete. Visto che la complessità dell'algoritmo cresce con la lunghezza della sequenza in ingresso alla rete, il dataset è suddiviso in sequenze di breve durata (tipicamente comprese tra 512 e 2048 campioni nei problemi audio), sulle quali la BPTT è applicata indipendentemente. Questo procedimento è conosciuto come truncated BPTT (TBPTT).

L'implementazione dei filtri IIR come RNNs è resa semplice da framework come Pytorch o TensorFlow, che permettono la differenziazione automatica del grafico computazionale della rete: visto che le strutture IIR contengono solo operazioni aritmetiche di base, per le quali i framework sono già in grado di calcolare le derivate, è sufficiente implementare il forward-pass del sistema, che è equivalente alla sua equazione alle differenze.

Capitolo 3 Modello proposto

Il modello di Wiener-Hammerstein proposto in [16], in cui i blocchi lineari sono realizzati come la cascata di filtri IIR differenziabili e il blocco non lineare come un multilayer perceptron (MLP), è stato utilizzato per identificare l'amplificatore del Leslie. Questa architettura ha due proprietà desiderabili:

1. il modello è completamente differenziabile e tutti e tre i blocchi possono essere allenati simultaneamente;
2. non c'è bisogno di fare alcuna assunzione riguardo la forma del blocco non lineare perché il MLP, essendo un approssimatore universale, è in grado di apprendere ogni arbitraria non linearità statica (a patto che abbia una capacità sufficiente).

Il training è stato effettuato su diverse reti, che differiscono l'una dall'altra per il numero e il tipo di filtri utilizzati. In particolare, le strutture IIR testate sono: sezioni del secondo ordine realizzate con la forma diretta trasposta II (TDF-II), e filtri a variabili di stato (SVFs) [17]. La funzione di trasferimento, nel dominio z , per un generico filtro IIR del secondo ordine avente due poli e due zeri è data da:

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad (3.20)$$

L'equazione alle differenze per la Eq. 3.20, utilizzando la TDF-II, è data da:

$$y(n) = b_0 x(n) + h_1(n-1) \quad (3.21)$$

$$h_1(n) = b_1 x(n) + h_2(n-1) - a_1 y(n) \quad (3.22)$$

$$h_2(n) = b_2 x(n) - a_2 y(n) \quad (3.23)$$

Per mantenere la stabilità del filtro durante il training, è necessario applicare dei vincoli ai coefficienti a_1 e a_2 , in modo che i poli siano all'interno del cerchio unitario. In [16] si suggerisce di imporre $|a_1| \leq 0.5$ e $|a_2| \leq 0.5$.

Lo SVF è un filtro del secondo ordine che viene controllato da parametri di alto livello, facilmente interpretabili. Questa struttura ha il vantaggio di avere un controllo disaccoppiato per la frequenza e l'ampiezza di entrambi i poli e la possibilità di mixare liberamente le sue tre uscite per produrre zeri arbitrari. L'equazione alle differenze per lo SVF è data da:

$$y_{BP}(n) = \frac{g(x(n) - h_2(n-1)) + h_1(n-1)}{1 + g(g + 2R)} \quad (3.24)$$

$$y_{LP}(n) = g y_{BP}(n) + h_2(n-1) \quad (3.25)$$

$$y_{HP}(n) = x(n) - y_{LP}(n) - 2R y_{BP}(n) \quad (3.26)$$

$$h_1(n) = 2y_{BP}(n) - h_1(n-1) \quad (3.27)$$

$$h_2(n) = 2y_{LP}(n) - h_2(n-1) \quad (3.28)$$

$$y(n) = c_{HP} y_{HP}(n) + c_{BP} y_{BP}(n) + c_{LP} y_{LP}(n) \quad (3.29)$$

dove g è il coefficiente che controlla la frequenza di taglio del filtro (spesso derivato dalla frequenza analogica utilizzando il prewarping, ma qui usato direttamente) e c_{HP} , c_{BP} , c_{LP} sono i coefficienti di mix per le tre uscite del filtro. Il parametro R setta il damping del filtro, con l'auto-oscillazione che avviene con $R = 0$. I parametri allenabili dello SVF sono quindi c_{HP} , c_{BP} , c_{LP} , g e R . I parametri g e R devono essere inizializzati con valori positivi per produrre un filtro stabile.

La struttura utilizzata per il MLP consiste in quattro strati fully connected, con due strati nascosti di dimensione 20. Il primo strato mappa la dimensione dell'ingresso, che in questo caso è 1, con la dimensione degli strati nascosti, cioè 20. L'ultimo strato mappa la dimensione degli strati nascosti con quella desiderata per l'uscita, cioè 1. La funzione di attivazione scelta è la tangente iperbolica ed è posizionata dopo ogni strato nascosto.

Una volta effettuato il training della rete, i parametri dei filtri sono facilmente estraibili ed utilizzabili per un'implementazione real-time del modello. Il MLP, invece, ha un costo computazionale elevato. Per questo, è stata adottata una strategia per la sua sostituzione basata sull'utilizzo di una look-up-table. Il metodo consiste nel misurare la caratteristica ingresso-uscita del MLP, a training terminato. Questo può essere fatto modificando il forward-pass della rete, escludendo i filtri dei blocchi LTI, e mandando in ingresso una rampa. L'uscita registrata viene memorizzata in una tabella, che viene letta interpolandone i valori. In questo lavoro è stata utilizzata l'interpolazione lineare.

3.8.3 Modello di Wiener-Hammerstein ricavato da procedura di ottimizzazione

L'ultimo modello proposto è quello di [18]. In questo caso i blocchi lineari consistono ognuno nella cascata di 12 filtri IIR del secondo ordine (1 high-shelving, 1 low-shelving e 10 peak, del tipo descritto in [12]) che vengono adattati utilizzando le misurazioni della risposta impulsiva del sistema da identificare. I parametri del blocco non lineare (Figura 3.28) vengono ricavati da una procedura di ottimizzazione che sfrutta l'algoritmo di Levenberg-Marquardt [19, 20] per minimizzare una funzione costo basata sull'errore tra lo spettrogramma dell'uscita del modello e lo spettrogramma della sequenza di riferimento. Per i dettagli delle procedure si rimanda a [18].

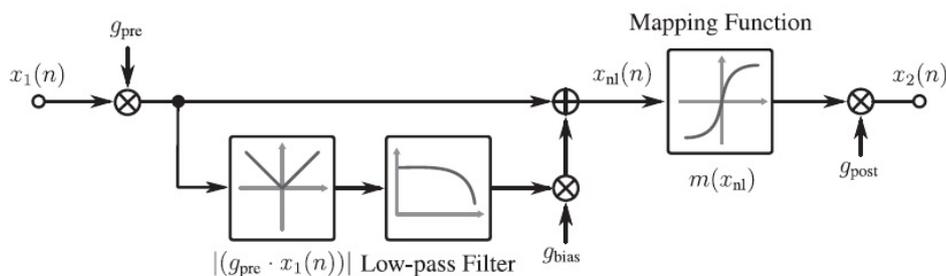


Figura 3.28: Circuito interno al blocco non lineare.

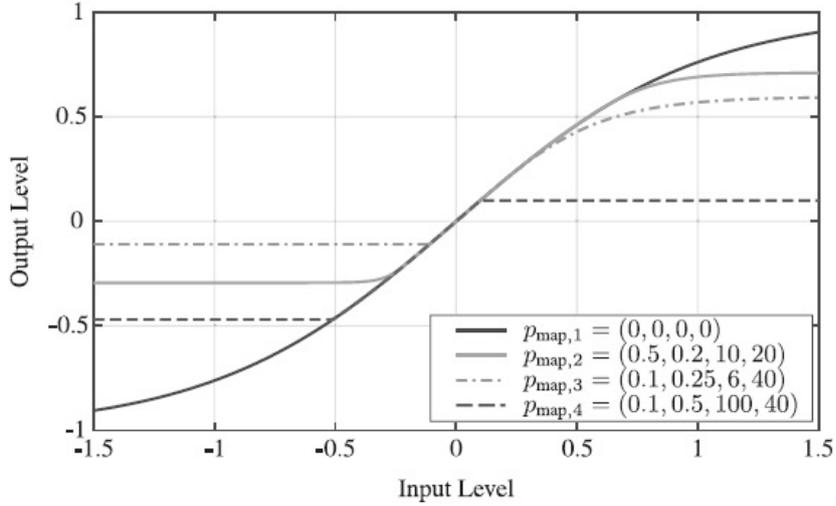


Figura 3.29: Funzione di mapping per differenti valori dei suoi parametri. Nella legenda i parametri sono ordinati come: k_p , k_n , g_p , g_n . I valori di g_p e g_n sono in dB.

La particolarità di questo modello è che il blocco non lineare è con memoria. Nel circuito di Figura 3.28 è presente infatti una diramazione in cui viene calcolato l'involuppo del segnale (si calcola prima il modulo del segnale e poi questo viene filtrato con un filtro passa-basso). L'involuppo è poi sommato al segnale, moltiplicato per un guadagno. Questo consente di simulare l'effetto particolare prodotto da un amplificatore valvolare quando il punto di polarizzazione della valvola si sposta a causa di un segnale d'ingresso con molti transienti. La funzione di mapping non lineare applicata al segnale è la seguente:

$$m(x_{nl}) = \begin{cases} \tanh(k_p) - \left[\frac{\tanh(k_p)^2 - 1}{g_p} \tanh(g_p(x_{nl} - k_p)) \right] & \text{se } x_{nl} > k_p \\ \tanh(x_{nl}) & \text{se } -k_n \leq x_{nl} \leq k_p \\ -\tanh(k_n) - \left[\frac{\tanh(k_n)^2 - 1}{g_n} \tanh(g_n(x_{nl} + k_n)) \right] & \text{se } x_{nl} < -k_n \end{cases} \quad (3.30)$$

Si noti che, per ogni campione del segnale, è necessario calcolare una sola volta la tangente iperbolica. In Figura 3.29 è mostrata la forma della funzione di mapping per diverse combinazioni dei suoi parametri. I parametri del blocco non lineare che vengono aggiustati durante il processo di ottimizzazione sono i guadagni g_{pre} , g_{bias} , g_{post} e i parametri della funzione di mapping k_p , k_n , g_p , g_n .

In Figura 3.30 è mostrato lo schema a blocchi per il calcolo della funzione costo,

Capitolo 3 Modello proposto

che è descritta dalle seguenti equazioni:

$$Y(b, k) = \text{STFT}(y(n)) \quad (3.31)$$

$$\hat{Y}(b, k, \mathbf{p}) = \text{STFT}(\hat{y}(n, \mathbf{p})) \quad (3.32)$$

$$R(b, k, \mathbf{p}) = L(|Y(b, k)|) - L(|\hat{Y}(b, k, \mathbf{p})|) \quad (3.33)$$

$$R(b, \tilde{k}, \mathbf{p}) = \text{pool}(R(b, k, \mathbf{p})) \quad (3.34)$$

$$C(\mathbf{p}) = \sum_k \sum_b R(b, \tilde{k}, \mathbf{p})^2 \quad (3.35)$$

dove \mathbf{p} è il vettore dei parametri da ottimizzare. Il segnale di riferimento $y(n)$ e l'uscita del modello $\hat{y}(n, \mathbf{p})$ sono trasformati nel dominio della STFT utilizzando una FFT a 4096 punti, un overlap di 1024 campioni e la finestra Hamming. Il modulo degli spettrogrammi è scalato logicamente in accordo con la seguente formula:

$$L(|Y(b, k)|) = \max\left(0, \frac{20 \log(|Y(b, k)|)}{L_{\text{offset}}} + 1\right) \quad (3.36)$$

con $L_{\text{offset}} = 100$ dB. In questo modo tutto quello che ha un'ampiezza inferiore a -100 dB, assumendo che sia rumore di fondo, viene trascurato durante l'ottimizzazione. Il residuo $R(b, k, \mathbf{p})$ è calcolato utilizzando la Eq. 3.33. L'operazione di pooling, indicata dalla Eq. 3.34, è effettuata calcolando il valore medio dei bin di frequenza k per una certa regione corrispondente ai semitoni. Lo spettro in semitoni inizia a $f_0 = 27.5$ Hz, andando a $f_{m+1} = f_m \cdot 2^{1/12}$, fino a $f_M = 19912.13$ Hz, con $M = 114$. Le frequenze da dc a f_0 sono mediate ad un solo valore come le frequenze da f_M a $f_s/2$ (f_s indica la frequenza di campionamento). Lo score finale $C(\mathbf{p})$ è calcolato sommando i quadrati degli elementi dello spettrogramma residuo.

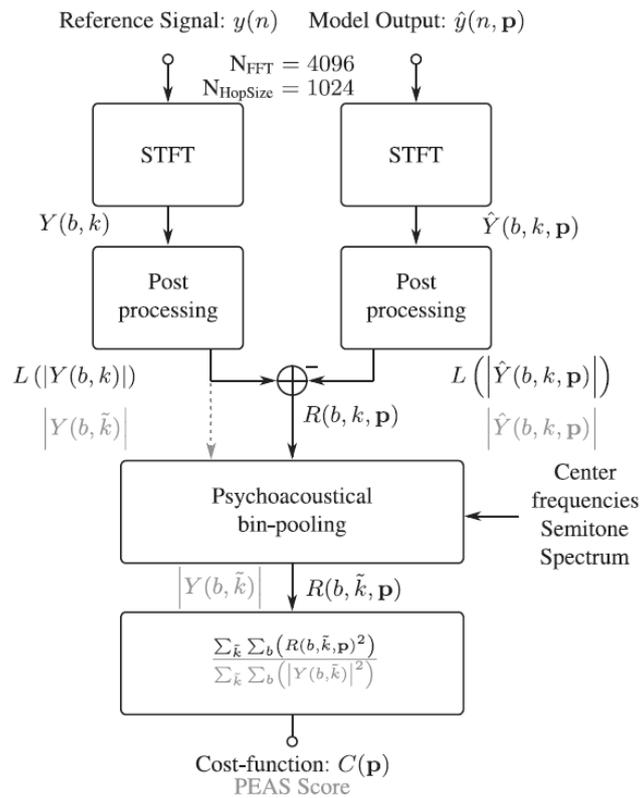


Figura 3.30: Diagramma a blocchi per il calcolo della funzione costo $C(\mathbf{p})$ e della metrica PEAS. Gli step addizionali per il calcolo della PEAS sono mostrati in grigio.

Capitolo 4

Implementazione

4.1 Cenni su VCV Rack

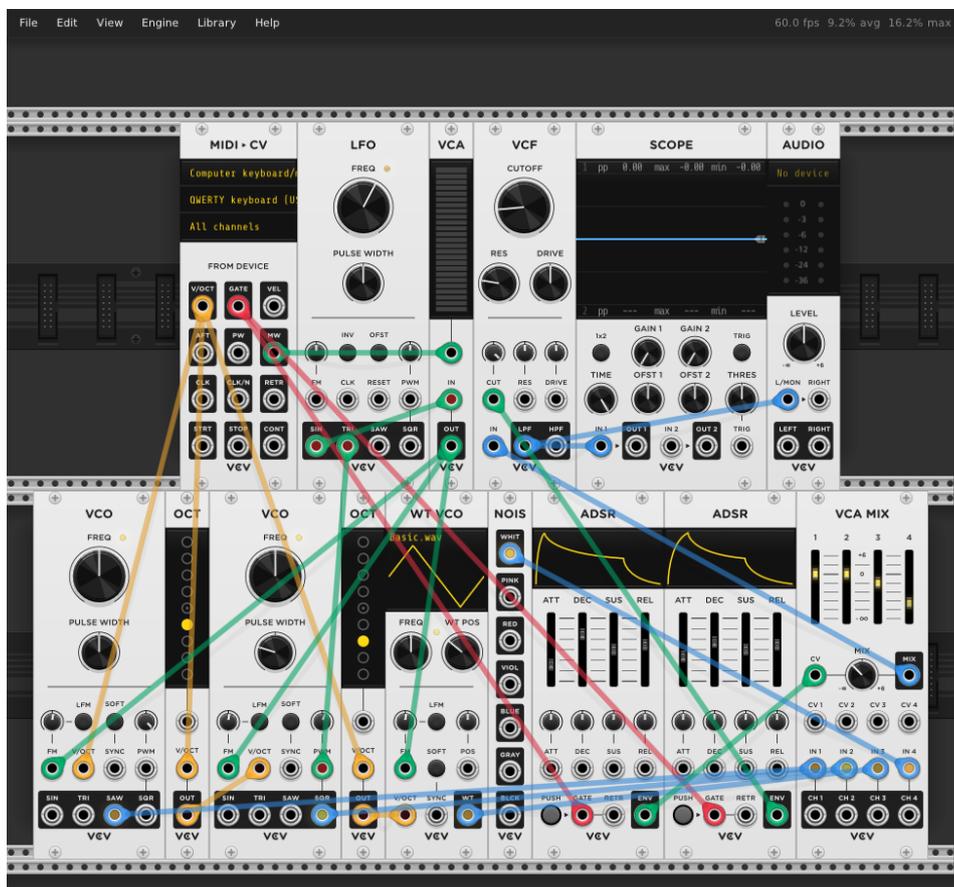


Figura 4.1: Interfaccia grafica di VCV Rack.

VCV Rack è un software di elaborazione audio open-source destinato a emulare un sistema Eurorack. In Figura 4.1 è mostrata la sua interfaccia grafica, in cui i moduli, che rappresentano i blocchi funzionali del sintetizzatore modulare virtuale, possono essere caricati e collegati tra loro. Nel linguaggio di Rack, il termine "modulo" si riferisce all'implementazione DSP di un modulo virtuale che viene compilata in una libreria dinamica caricata a runtime dall'applicazione. Il termine "plugin"

indica invece una collezione di moduli raggruppati in una cartella contenente un file compilato, la libreria dinamica.

Ogni modulo implementa una funzione di elaborazione audio, chiamata `process`, che viene invocata periodicamente dal motore audio. La particolarità di VCV Rack, rispetto ad altre piattaforme, è che solo un campione alla volta è passato alla `process` (che viene chiamata in ogni istante di campionamento). Il motivo per cui Rack utilizza questa strategia, che è meno efficiente rispetto ad un'elaborazione a buffer, è che il suo scopo principale è quello di simulare al meglio i feedback senza ritardo che ci sono nei sintetizzatori modulari analogici. Nel dominio digitale il ritardo di un feedback è pari alla dimensione del buffer che viene processato. Quindi, a parità di frequenza di campionamento, la migliore approssimazione del sistema reale si ottiene con un'elaborazione campione-per-campione.

In termini di codice, un modulo Rack consiste in una struct C++ `Module` che viene definita ereditando due oggetti principali:

- la struct `Module`, che implementa la parte DSP del modulo;
- la struct `ModuleWidget`, che implementa l'aspetto grafico del modulo.

La struct `Module` contiene i seguenti importanti metodi e membri:

```
std::vector<Param> params;
std::vector<Input> inputs;
std::vector<Output> outputs;
std::vector<Light> lights;

virtual void process(const ProcessArgs &args);
void config(int numParams, int numInputs, int numOutputs,
            int numLights = 0);
void configParam(int paramId, float minValue, float maxValue,
                 float defaultValue, std::string name = "",
                 std::string unit = "", float displayBase = 0.f,
                 float displayMultiplier = 1.f, float displayOffset = 0.f);
```

I primi quattro membri sono vettori di parametri (ad esempio knob e switch), connettori d'ingresso e d'uscita, e luci, che sono inizializzati nel costruttore. Sottoclassando la struct `Module` è quindi possibile imporre un certo numero di questi elementi. Il metodo `process` è il metodo fondamentale per l'elaborazione del segnale che viene chiamato in ogni istante di campionamento, come spiegato in precedenza. Questo prende un argomento costante come ingresso: la struttura `ProcessArgs`, che contiene informazioni globali, come la frequenza di campionamento. Il metodo `config` è necessario per specificare il numero di parametri, porte e luci da creare. Il metodo `configParam` permette di settare il range di un parametro, il suo valore di default, e altro.

Di seguito sono riportati i membri e i metodi principali dell'oggetto `ModuleWidget`:

```
Module *module = NULL;
```

```
std : vector<ParamWidget*>params ;
std : vector<PortWidget*>outputs ;
std : vector<PortWidget*>inputs ;

void setPanel ( std :: shared_ptr<Svg> svg ) ;
void addParam ( ParamWidget *param ) ;
void addOutput ( PortWidget *output ) ;
void addInput ( PortWidget *input ) ;
```

Il costruttore ha bisogno di un puntatore alla sottoclasse `Module` che deve gestire. Il metodo `setPanel` permette di caricare un file SVG e applicarlo al background del pannello del modulo. I metodi `addParam`, `addOutput` e `addInput` permettono il posizionamento sul pannello di elementi come knob, switch, porte d'ingresso e d'uscita.

Una guida per lo sviluppo di nuovi moduli con VCV Rack si può trovare in [21].

4.2 Sviluppo del modulo del Leslie

Il modulo sviluppato (Figura 4.2) è costituito da un ingresso mono, un'uscita stereo e una serie di parametri che possono essere regolati dall'utente con i knob e gli switch presenti sull'interfaccia grafica. In particolare, è possibile controllare in tempo reale:

- il modo di funzionamento (*brake*, *fast*, *slow*);
- la frequenza di taglio del crossover;
- l'accensione/spengimento dell'amplificatore e il livello di distorsione;
- le velocità delle sorgenti nelle modalità *fast* e *slow*;
- i tempi di tutte le transizioni tra i modi di funzionamento;
- i raggi di rotazione delle sorgenti;
- il valore dei parametri *directivityAngle* e *directivityLevel*;
- il numero delle sorgenti immagine;
- le posizioni dei microfoni;
- il livello della sorgente non modulata relativa al woofer;
- il volume dei segnali d'uscita di tromba e woofer.

Inoltre, sono stati aggiunti dei filtri al modello per equalizzare i segnali della tromba e del woofer. Questi sono degli IIR del secondo ordine di tipo passa-basso e peak, realizzati come descritto in [12]. I parametri di controllo di questi filtri sono la frequenza di taglio per i passa-basso e la frequenza centrale, il guadagno e il Q-factor per i peak.



Figura 4.2: Interfaccia grafica del modulo sviluppato.

4.2.1 Geometria delle sorgenti

In Figura 4.3 e in Figura 4.4 è mostrata la geometria del modello per il simulatore della tromba. Sono state considerate le 5 sorgenti immagine corrispondenti alle prime riflessioni sulle pareti dello scomparto superiore del cabinet. Il sistema di riferimento, rispetto al quale vengono effettuati i calcoli delle coordinate, ha l'origine posizionata nel centro dello scomparto superiore. I punti H indicano i centri di rotazione, mentre i punti S indicano le posizione istantanee delle sorgenti sonore (si considera che il suono venga emesso dalla bocca della tromba). Le coordinate dei centri di rotazione (Tabella 4.1) e i versi di rotazione sono il risultato dell'applicazione del teorema delle immagini. Facendo riferimento a Figura 4.5, l'angolo posizionale θ e il raggio di rotazione r della sorgente principale possono essere utilizzati per calcolare le quantità:

$$x_{\text{var}} = r \cdot \cos(\theta) \quad (4.1)$$

$$y_{\text{var}} = r \cdot \sin(\theta) \quad (4.2)$$

che, sommate opportunamente alle coordinate x e y dei centri di rotazione, consentono di ottenere le coordinate istantanee delle sorgenti (Tabella 4.2).

Tabella 4.1: Coordinate dei centri di rotazione.

Punto	x	y	z
H1	0	y_1	z_1
H2	0	$W - y_1$	z_1
H3	0	$-W - y_1$	z_1
H4	$-L$	y_1	z_1
H5	0	y_1	$H - z_1$
H5	0	y_1	$-H - z_1$

Per la ripresa del suono si considera una coppia di microfoni L-R posizionati

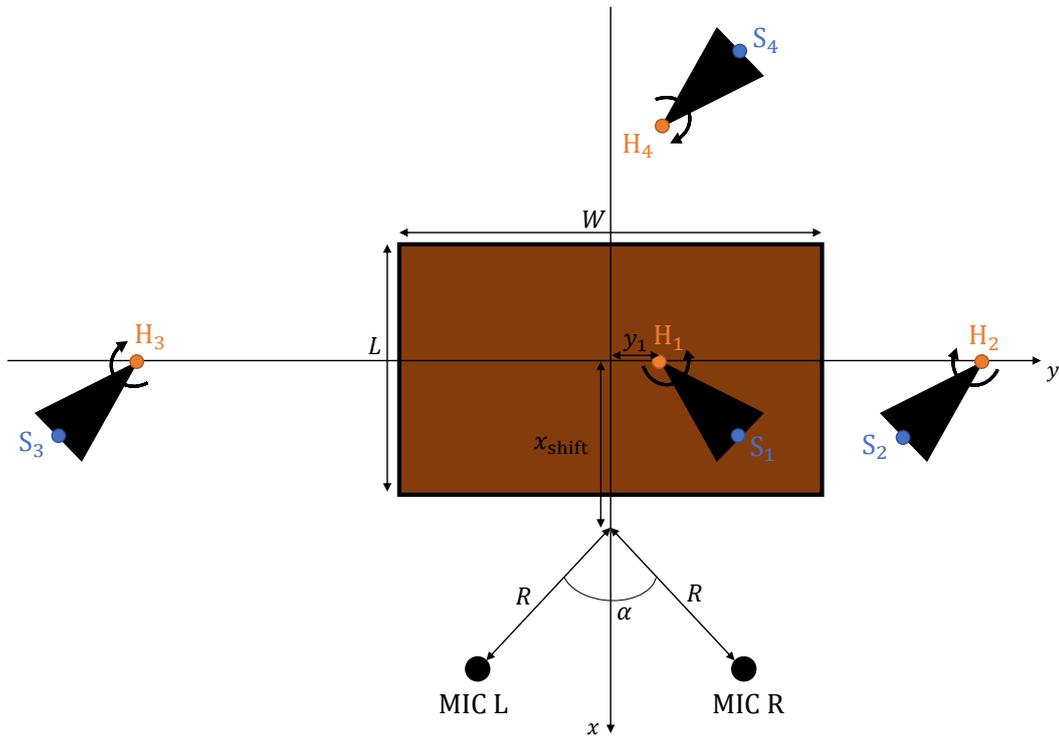


Figura 4.3: Geometria del simulatore della tromba (vista dall'alto del Leslie).

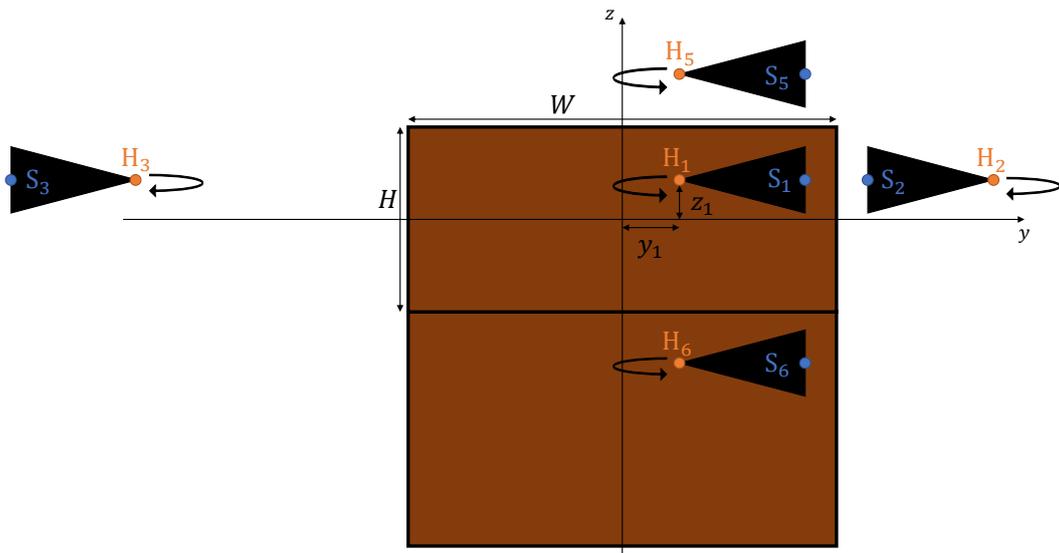


Figura 4.4: Geometria del simulatore della tromba (vista frontale del Leslie).

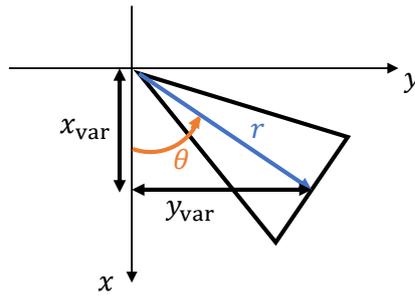


Figura 4.5: Geometria della sorgente principale.

Tabella 4.2: Coordinate delle sorgenti.

Punto	x	y	z
S1	x_{var}	$y_1 + y_{\text{var}}$	z_1
S2	x_{var}	$W - y_1 - y_{\text{var}}$	z_1
S3	x_{var}	$-W - y_1 - y_{\text{var}}$	z_1
S4	$-L - x_{\text{var}}$	$y_1 + y_{\text{var}}$	z_1
S5	x_{var}	$y_1 + y_{\text{var}}$	$H - z_1$
S5	x_{var}	$y_1 + y_{\text{var}}$	$-H - z_1$

simmetricamente rispetto all'asse x e con coordinata z nulla. Le posizioni dei microfoni sono controllate da due parametri: angolo α e raggio R . Le coordinate x e y dei microfoni possono essere calcolate come:

$$x_{\text{micL}} = x_{\text{micR}} = x_{\text{shift}} + R \cdot \cos(\alpha) \quad (4.3)$$

$$y_{\text{micL}} = -R \cdot \sin(\alpha) \quad (4.4)$$

$$y_{\text{micR}} = R \cdot \sin(\alpha) \quad (4.5)$$

dove la quantità x_{shift} garantisce una distanza minima tra i microfoni e le sorgenti.

Per il simulatore del woofer valgono le stesse considerazioni, ma il sistema di riferimento è posizionato al centro dello scomparto inferiore del cabinet. Inoltre, come nel vero Leslie, il verso di rotazione è opposto a quello della tromba.

4.2.2 Implementazione della funzione di elaborazione del modulo

In Figura 4.6 è mostrato lo schema a blocchi dell'implementazione della `process`. Di seguito sono elencate le operazioni effettuate in ogni istante di campionamento.

1. Aggiornamento dei coefficienti dei filtri degli equalizzatori. Il ricalcolo dei coefficienti di un filtro viene effettuato solo negli istanti di campionamento in cui si riscontra una variazione di uno dei corrispondenti parametri di controllo.
2. Aggiornamento delle distanze di soglia che controllano la modulazione in ampiezza. Il calcolo è effettuato solo negli istanti di campionamento in cui si

Capitolo 4 Implementazione

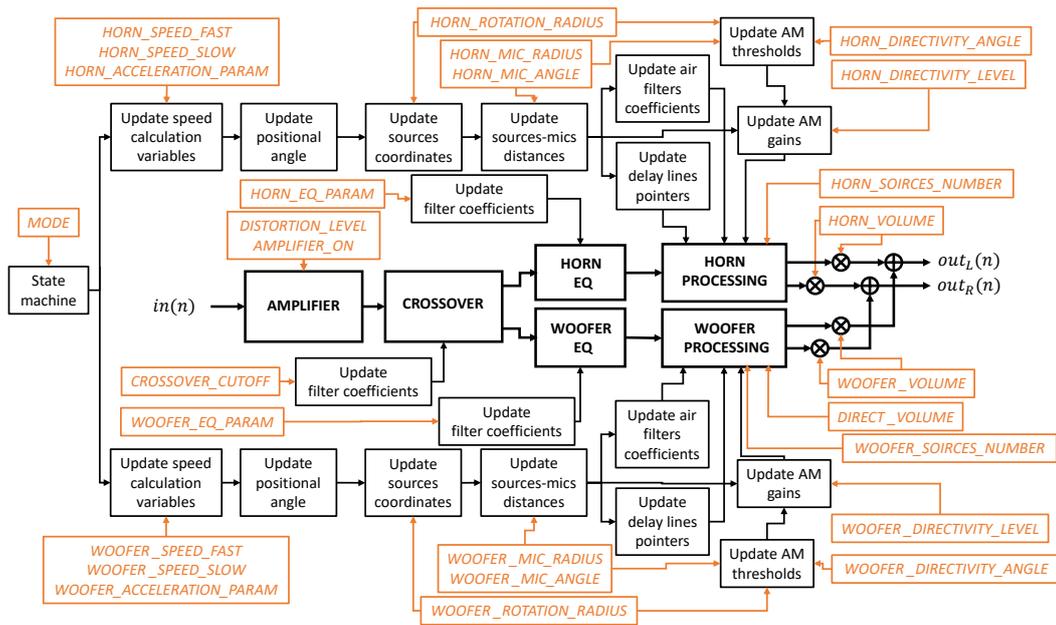


Figura 4.6: Schema a blocchi della funzione `process`. Viene mostrato come i parametri del modulo (in arancione) controllano le varie operazioni.

riscoprono variazioni dei parametri che controllano le posizioni dei microfoni o di quelli relativi alla modulazione (raggio di rotazione, *directivityAngle*).

3. Determinazione della transizione corrente con macchina a stati.
4. Aggiornamento delle variabili relative alla Eq. 3.14 per il calcolo delle frequenze di rotazione delle sorgenti. Il calcolo è effettuato solo negli istanti di campionamento in cui c'è un cambio di stato.
5. Calcolo delle frequenze di rotazione correnti e incremento degli angoli posizionali delle sorgenti.
6. Aggiornamento delle coordinate delle sorgenti. Il calcolo è effettuato come indicato nella sottosezione 4.2.1.
7. Aggiornamento delle distanze sorgente-microfono.
8. Aggiornamento dei coefficienti dei filtri di assorbimento dell'aria.
9. Aggiornamento dei guadagni AM.
10. Aggiornamento dei puntatori delle delay line.
11. Elaborazione del campione in ingresso al modulo. I blocchi `HORN PROCESSING` e `WOOFER PROCESSING` implementano lo schema di Figura 4.7 nel caso di delay line con lettura interpolata e lo schema di Figura 4.8 nel caso di delay line con scrittura interpolata. Nel primo caso è necessario usare un buffer per la sorgente

principale e uno per le sorgenti immagine. Nel secondo caso è necessario usare un buffer per il microfono destro e uno per il microfono sinistro.

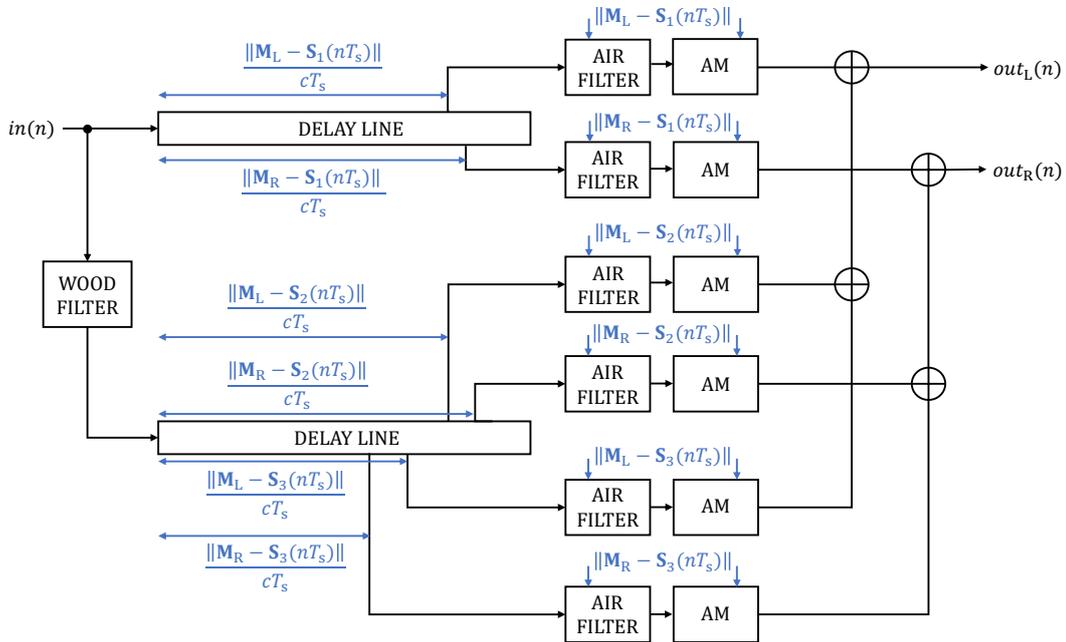


Figura 4.7: Implementazione dei blocchi HORN PROCESSING e WOOFER PROCESSING nel caso di utilizzo di delay line con lettura interpolata. Si considerano la sorgente di suono diretto e due sorgenti immagine.

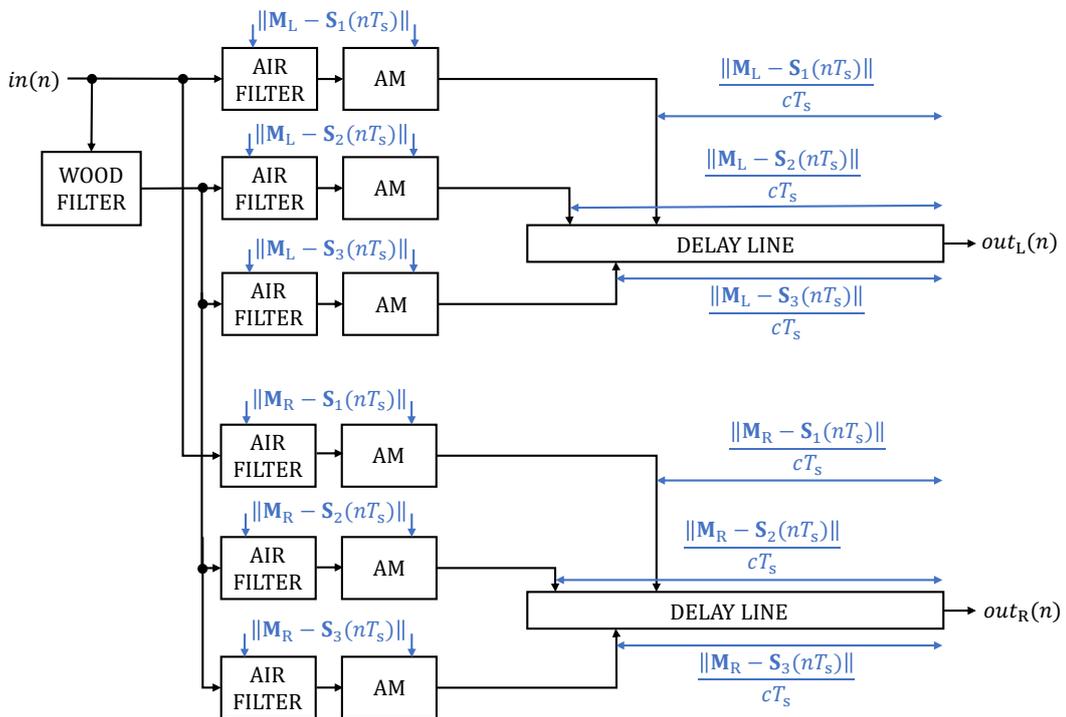


Figura 4.8: Implementazione dei blocchi HORN PROCESSING e WOOFER PROCESSING nel caso di utilizzo di delay line con scrittura interpolata. Si considerano la sorgente di suono diretto e due sorgenti immagine.

Capitolo 5

Risultati

5.1 Amplificatore

In questa sezione sono riportati i risultati dell'identificazione dell'amplificatore di un Leslie 122. Le misurazioni dell'amplificatore sono state effettuate utilizzando il setup di Figura 5.1. L'uscita dell'amplificatore è stata collegata ad un carico fittizio di 16Ω per simulare la circuiteria del Leslie. L'amplificatore lineare consente di portare la tensione di linea in uscita dalla scheda audio ad un livello adeguato per far saturare l'amplificatore del Leslie. La seconda uscita della scheda audio è direttamente connessa al secondo ingresso. Il segnale diretto registrato è stato utilizzato come segnale d'ingresso per il modello digitale durante le identificazioni. Questo è vantaggioso perché il segnale diretto misurato è automaticamente sincronizzato nel tempo con il segnale registrato in uscita all'amplificatore (requisito cruciale per l'algoritmo di identificazione).

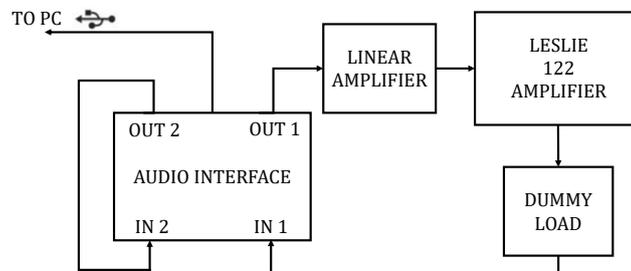


Figura 5.1: Setup utilizzato per la misurazione dell'amplificatore del Leslie.

5.1.1 Modello di Wiener-Hammerstein allenato con backpropagation

In Tabella 5.1 sono riportate le reti sulle quali è stato effettuato il training. Il dataset utilizzato consiste in circa 5 minuti di registrazioni di chitarra elettrica e organo ed è stato organizzato in batch composte da 512 sequenze di lunghezza pari a 2048 campioni. La metrica utilizzata è l'errore quadratico medio (MSE) nel tempo.

In Tabella 5.2 sono mostrati i valori finali della loss raggiunti nel training e i valori calcolati sul testing set. I risultati sono comparabili con quelli ottenuti in [16] nell'identificazione del pedale di distorsione per chitarra BOSS DS-1. Per fare

Tabella 5.1: Elenco dei modelli su cui è stato effettuato il training.

Modello	1° blocco LTI	MLP	2° blocco LTI	Numero parametri
WH_SVF11_MLP2	1 SVF	2 hidden layers (tanh activations)	1 SVF	913
WH_BIQ11_MLP2	1 TDF-II	2 hidden layers (tanh activations)	1 TDF-II	911
WH_SVF33_MLP2	3 SVF	2 hidden layers (tanh activations)	3 SVF	937
WH_BIQ44_MLP2	4 TDF-II	2 hidden layers (tanh activations)	4 TDF-II	941
WH_SVF66_MLP2	6 SVF	2 hidden layers (tanh activations)	6 SVF	973

Tabella 5.2: Risultati del training e del testing delle reti.

Modello	Training loss	Testing loss
WH_SVF11_MLP2	1.58E-02	1.14E-02
WH_BIQ11_MLP2	2.74E-02	1.45E-02
WH_SVF33_MLP2	1.29E-02	7.67E-03
WH_BIQ44_MLP2	1.29E-02	6.44E-03
WH_SVF66_MLP2	1.19E-02	6.43E-03

un confronto, è possibile considerare i risultati ottenuti dall'identificazione di un pedale di distorsione Ibanez TS9 con la rete WH_SVF11_MLP2. In questo caso, il modello riesce a simulare molto bene il dispositivo e il valore della loss calcolata sul testing set è pari a 4.42E-04. Come prevedibile, le prestazioni della rete sono peggiori nel caso dell'amplificatore del Leslie, visto che un sistema valvolare è più complesso da modellare rispetto ad un pedale di distorsione. Come si nota in Tabella 5.2, l'aumento del numero di filtri della rete porta ad un miglioramento dei risultati, ma, all'ascolto, la differenza tra il modello e il riferimento è ancora evidente. Risultati soddisfacenti potrebbero essere raggiunti aumentando di molto il numero dei filtri della rete. In questo caso però, vista l'implementazione dei filtri come layer ricorrenti, il training richiederebbe tempi molto lunghi e sarebbe preferibile utilizzare un approccio differente. Una possibilità è l'efficiente schema di training, basato sul campionamento nel dominio della frequenza, proposto in [22]. Questa parte è riservata a sviluppi futuri.

5.1.2 Modello di Wiener-Hammerstein ricavato da procedura di ottimizzazione

I risultati dell'adattamento dei filtri IIR dei blocchi lineari sono mostrati in Figura 5.2 e in Figura 5.3. Le risposte target da approssimare sono state ricavate

dalle risposte impulsive misurate per piccoli e per grandi segnali, come indicato in [18].

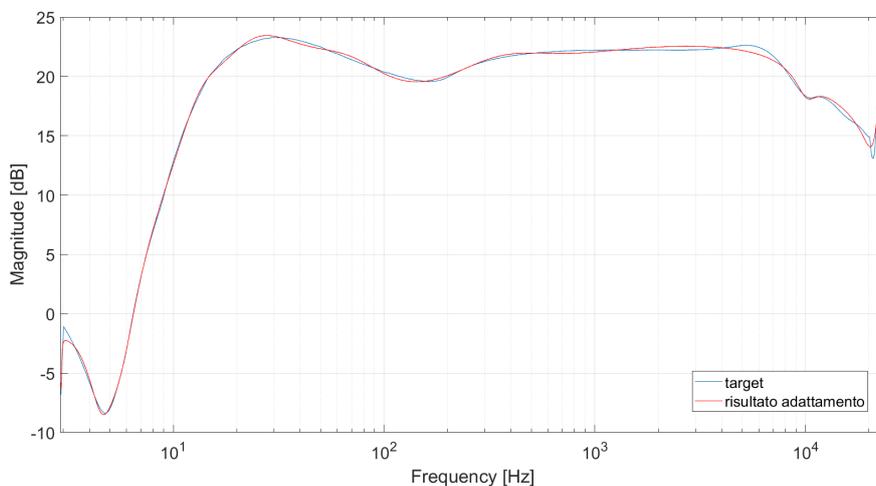


Figura 5.2: Risultato dell'adattamento dei filtri IIR che formano il blocco lineare d'ingresso del modello.

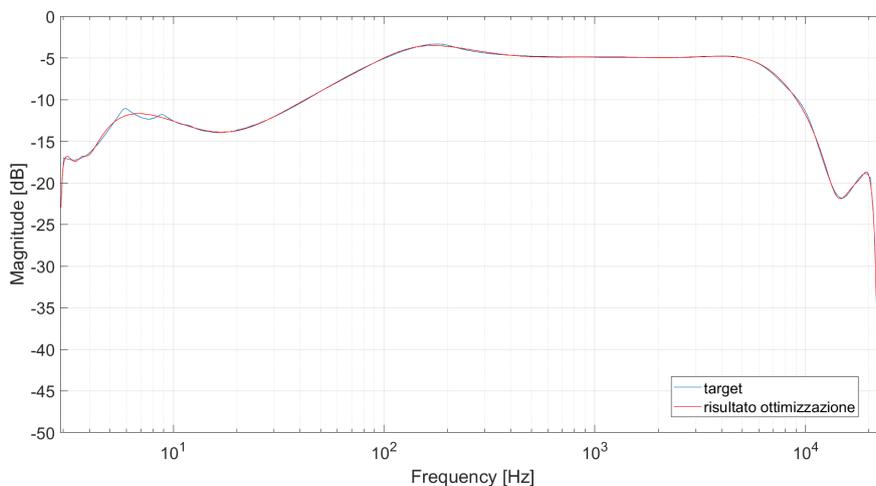


Figura 5.3: Risultato dell'adattamento dei filtri IIR che formano il blocco lineare d'uscita del modello.

Per quanto riguarda l'ottimizzazione del blocco non lineare, i segnali dati in ingresso al modello durante la procedura consistono in brevi registrazioni di chitarra e organo. Dai test effettuati, si è visto che variare le condizioni iniziali dei parametri da ottimizzare, rispetto a quelle suggerite in [18], può portare ad una migliore identificazione.

Per valutare quanto bene il modello ricavato simula l'amplificatore sono state utilizzate le seguenti metriche.

Capitolo 5 Risultati

- Perceptual evaluation of audio similarity (PEAS), simile alla funzione costo utilizzata durante l'ottimizzazione. Viene calcolata con lo schema di Figura 3.30.
- Root mean square error:

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N (y(n) - \hat{y}(n))^2} \quad (5.1)$$

- Error to signal ratio:

$$ESR = \frac{\sum_{n=1}^N (y(n) - \hat{y}(n))^2}{\sum_{n=1}^N y(n)^2} \quad (5.2)$$

- Correlation coefficient:

$$r_{y\hat{y}} = \frac{\sum_{n=1}^N (y(n) - \bar{y})(\hat{y}(n) - \bar{\hat{y}})}{\sqrt{\sum_{n=1}^N (y(n) - \bar{y})^2} \sqrt{\sum_{n=1}^N (\hat{y}(n) - \bar{\hat{y}})^2}} \quad (5.3)$$

Nelle espressioni precedenti, $y(n)$ è la sequenza di riferimento, $\hat{y}(n)$ è l'uscita del modello, \bar{y} indica il valore medio della sequenza $y(n)$ e N è la lunghezza delle sequenze. I valori calcolati delle metriche, riportati in Tabella 5.3, sono comparabili con quelli ottenuti in [18], eccetto per l'RMS che assume un valore più elevato in questo caso. I risultati possono essere confrontati con quelli ottenuti dall'identificazione di un sistema meno complesso, che consiste nel modello digitale di una valvola. In questo caso, l'algoritmo ha portato ad una identificazione quasi perfetta e, all'ascolto, l'uscita del modello è difficilmente distinguibile dal target. I valori delle metriche ottenuti dall'identificazione del modello digitale della valvola sono riportati in Tabella 5.4.

Tabella 5.3: Risultato dell'identificazione dell'amplificatore del Leslie. Sono riportati i valori calcolati delle metriche.

PEAS	RMS	ESR	$r_{y\hat{y}}$
0.0460	0.1658	0.1960	0.9132

Tabella 5.4: Risultato dell'identificazione del modello digitale di una valvola. Sono riportati i valori calcolati delle metriche.

PEAS	RMS	ESR	$r_{y\hat{y}}$
6.0881E-04	0.0028	8.8521E-04	0.9997

5.1.3 Osservazioni

La rete neurale che ha ottenuto i risultati migliori (WH_SVF66_MLP2) e il modello ottenuto dalla procedura di ottimizzazione sono stati messi a confronto. In Figura 5.4 sono mostrati gli spettrogrammi calcolati sulle uscite dei modelli e lo spettrogramma del segnale di riferimento, nel caso di un segnale d'ingresso costituito da una registrazione di organo. Si nota come il modello neurale non riesce ad identificare bene la parte più in alta frequenza, con meno energia. Questa caratteristica è stata osservata per tutte le reti allenate. Per fare un ulteriore confronto sono state utilizzate le metriche descritte nella sottosezione 5.1.2. I risultati sono riportati in Tabella 5.5 e in Tabella 5.6. I valori ottenuti nei due casi sono comparabili, ma migliori per il modello neurale. Questo significa che la rete riesce ad identificare meglio la parte dello spettro con maggiore energia. Si noti che i modelli neurali consentono di ottenere risultati comparabili con il modello ottimizzato utilizzando un numero di filtri (per i blocchi lineari) molto minore. Insieme al costo computazionale più basso si ha però un'occupazione in memoria maggiore dovuta alla look-up-table (che sostituisce il MLP). All'ascolto i modelli hanno caratteristiche differenti e risulta difficile affermare quale sia il più simile al riferimento.

Si vuole infine notare che nel vero Leslie la forma spettrale del segnale in uscita all'amplificatore è fortemente modificata dalle risposte in frequenza dei driver e dei diffusori. Nel modello proposto questa parte non viene considerata, quindi, anche nel caso di una perfetta identificazione, quando l'amplificatore viene inserito nella catena di elaborazione del simulatore l'effetto complessivo è differente da quello che si ha nel vero Leslie.

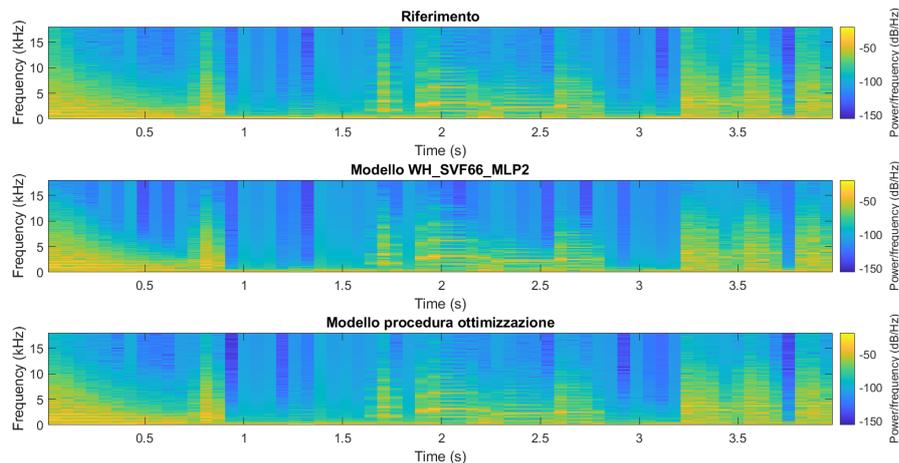


Figura 5.4: Sono mostrati: lo spettrogramma del segnale di riferimento; lo spettrogramma del modello WH_SVF66_MLP2; lo spettrogramma del modello ottenuto dalla procedura di ottimizzazione.

Tabella 5.5: Valori delle metriche per il modello WH_SVF66_MLP2. Il calcolo è effettuato sui segnali di Figura 5.4.

PEAS	RMS	ESR	$r_{y\hat{y}}$
0.0224	0.0643	0.1008	0.9685

Tabella 5.6: Valori delle metriche per il modello ricavato dalla procedura di ottimizzazione. Il calcolo è effettuato sui segnali di Figura 5.4.

PEAS	RMS	ESR	$r_{y\hat{y}}$
0.0376	0.1563	0.1570	0.9243

5.2 Simulatore degli altoparlanti rotanti

In questa sezione sono riportati i risultati ottenuti dall'analisi del simulatore degli altoparlanti rotanti. Il modello è stato valutato mettendo a confronto le caratteristiche temporali e spettrali del segnale simulato con quelle di una registrazione di un vero Leslie, ed effettuando dei test di ascolto. Le misurazioni del Leslie sono state effettuate posizionandolo all'interno di una stanza trattata con pannelli fonoassorbenti. Per misurare il segnale della tromba, il woofer è stato scollegato ed il microfono è stato posizionato ad una distanza di 30 cm dal centro dello scomparto superiore del cabinet. Per la misurazione del woofer, la tromba è stata scollegata e il microfono è stato posizionato ad una distanza di 10 cm dal centro dello scomparto inferiore. In entrambi i casi, il lato del cabinet sul quale è posizionato il microfono è quello aperto (lato posteriore, sul quale è stato rimosso il pannello in legno). Il segnale d'ingresso è fornito da un organo.

5.2.1 Simulatore della tromba

In Figura 5.5-Figura 5.9 è mostrato il confronto tra il segnale misurato e l'uscita del simulatore della tromba (forme d'onda e spettrogrammi), al variare del numero di sorgenti virtuali considerate. Il segnale d'ingresso utilizzato consiste nella nota D6 dell'organo suonata in modo stazionario. I parametri geometrici del simulatore sono stati settati in modo corrispondere alle condizioni in cui è stata effettuata la misura. Il parametro *directivityLevel* è stato settato ad 1.

5.2.2 Simulatore del woofer

In Figura 5.10-Figura 5.14 è mostrato il confronto tra il segnale misurato e l'uscita del simulatore del woofer (forme d'onda e spettrogrammi), al variare del numero di sorgenti virtuali considerate e del parametro *directivityLevel*. Il segnale d'ingresso utilizzato consiste nella nota C3 dell'organo suonata in modo stazionario. I parametri

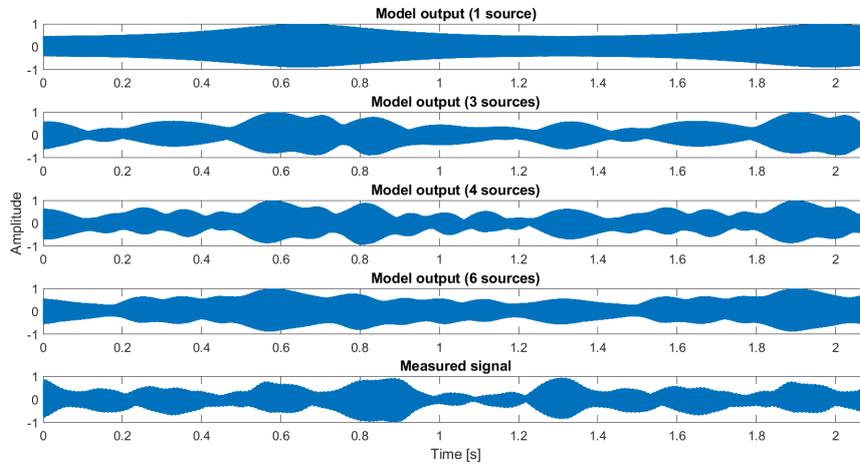


Figura 5.5: Forme d'onda del segnale in uscita dal simulatore e del segnale misurato. Il numero di sorgenti virtuali considerate viene progressivamente aumentato: nel primo grafico viene considerata solamente la sorgente di suono diretto; nel secondo grafico vengono aggiunte le immagini corrispondenti alle pareti laterali; nel terzo grafico viene aggiunta l'immagine corrispondente alla parete frontale; nel quarto grafico si considerano tutte le riflessioni del primo ordine.

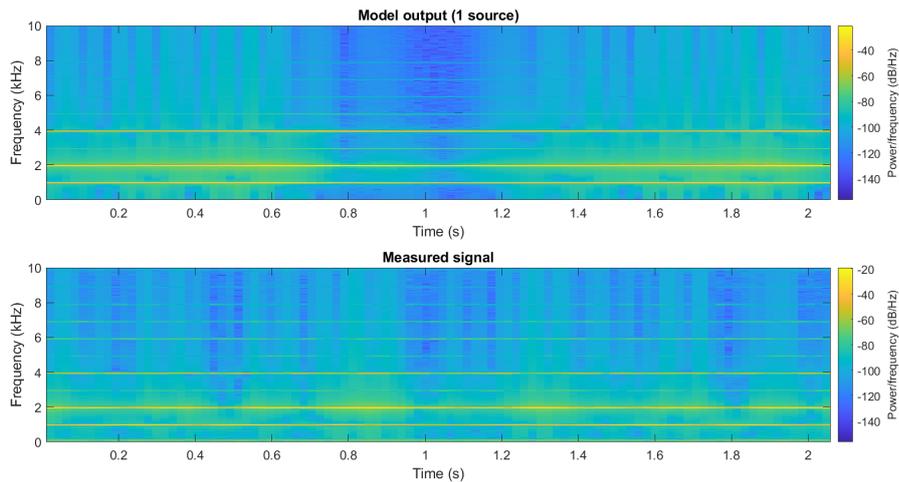


Figura 5.6: Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando solamente la sorgente di suono diretto, e lo spettrogramma del segnale misurato.

Capitolo 5 Risultati

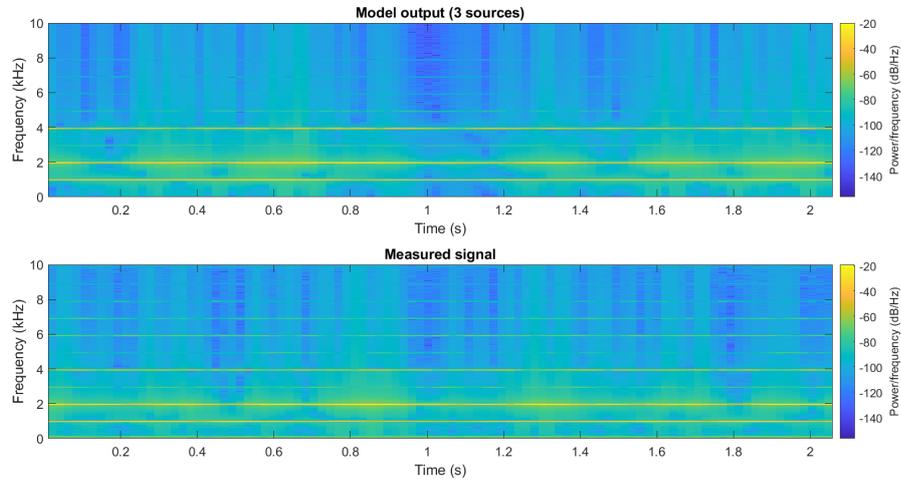


Figura 5.7: Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali, e lo spettrogramma del segnale misurato.

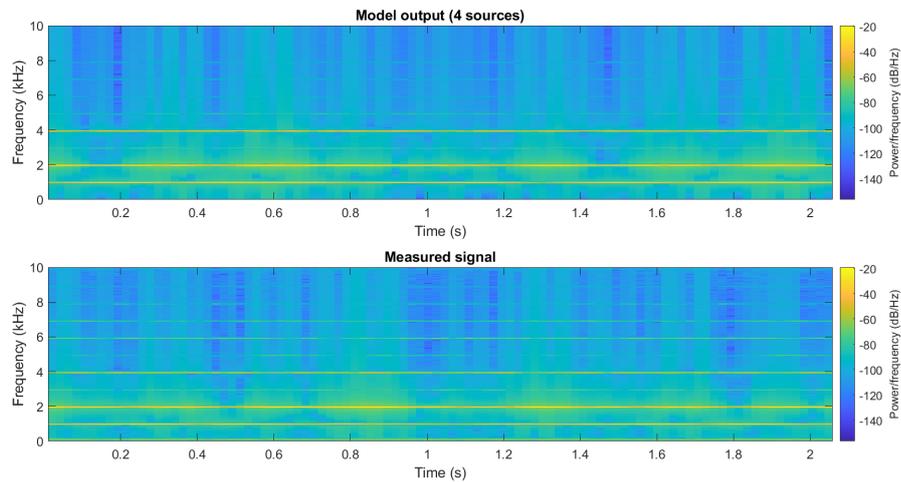


Figura 5.8: Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale, e lo spettrogramma del segnale misurato.

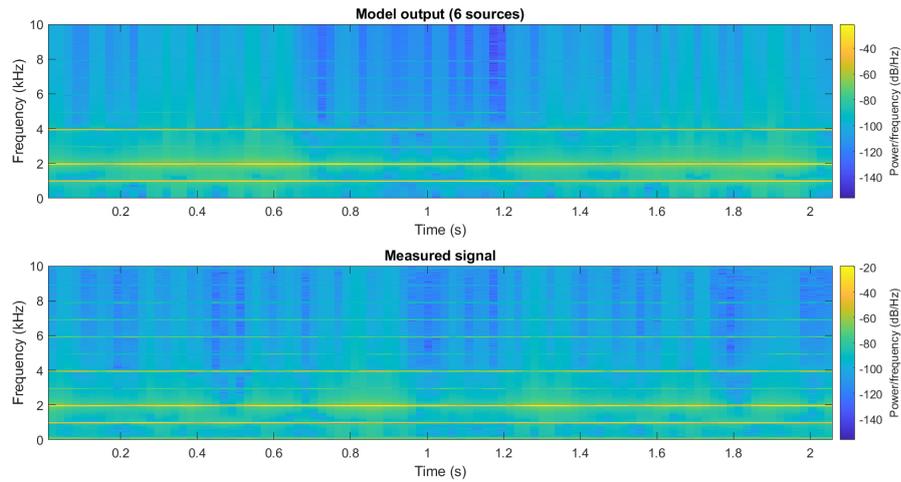


Figura 5.9: Confronto tra lo spettrogramma del segnale in uscita dal simulatore, ottenuto considerando tutte le immagini del primo ordine, e lo spettrogramma del segnale misurato.

geometrici del simulatore sono stati settati in modo corrispondere alle condizioni in cui è stata effettuata la misura.

5.2.3 Osservazioni

I risultati ottenuti mostrano che il modello proposto non riesce a simulare perfettamente le caratteristiche complesse dell'effetto Leslie. Le differenze sono dovute a diversi motivi.

- Il modello non tiene conto delle risposte in frequenza dei driver e dei diffusori.
- Il teorema delle immagini non è perfettamente applicabile per via della direttività dipendente in frequenza dei diffusori. Questo problema influisce maggiormente sul simulatore del woofer vista l'elevata direttività del deflettore.
- Le riflessioni di ordine superiore al primo non vengono considerate.
- Il modello trascura la diffusione del suono sulle pareti del cabinet e la diffrazione sulle feritoie.
- La risposta in fase del legno non è stata tenuta in conto nel progetto del filtro di assorbimento.

Comunque, con il modello è possibile ottenere uno schema di modulazione simile a quello del vero Leslie. Dai grafici si nota che la corrispondenza migliore si ottiene considerando solamente 4 sorgenti immagine (quelle corrispondenti alle pareti laterali e frontale del cabinet). Si è visto, infatti, che l'aggiunta delle sorgenti corrispondenti alle pareti superiore e inferiore peggiora il risultato. Anche i test di ascolto hanno

Capitolo 5 Risultati

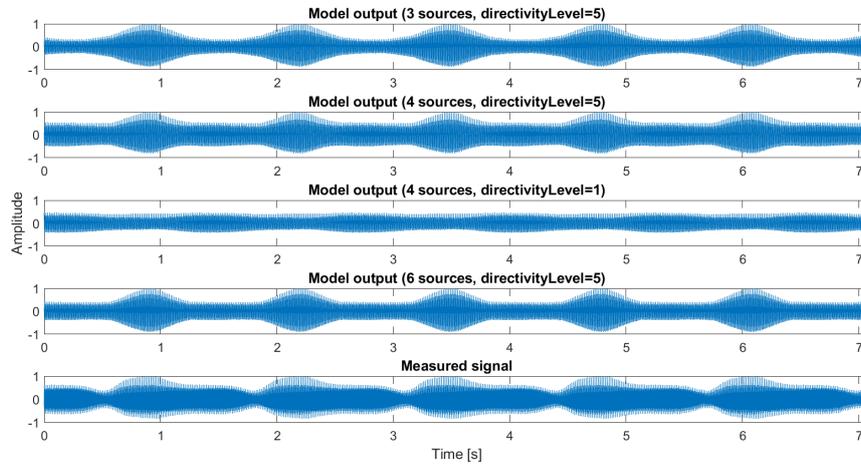


Figura 5.10: Forme d'onda del segnale in uscita dal simulatore e del segnale misurato. Nel primo grafico vengono considerate la sorgente principale e le immagini corrispondenti alle pareti laterali; nel secondo e terzo grafico viene aggiunta l'immagine corrispondente alla parete frontale; nel quarto grafico si considerano tutte le riflessioni del primo ordine.

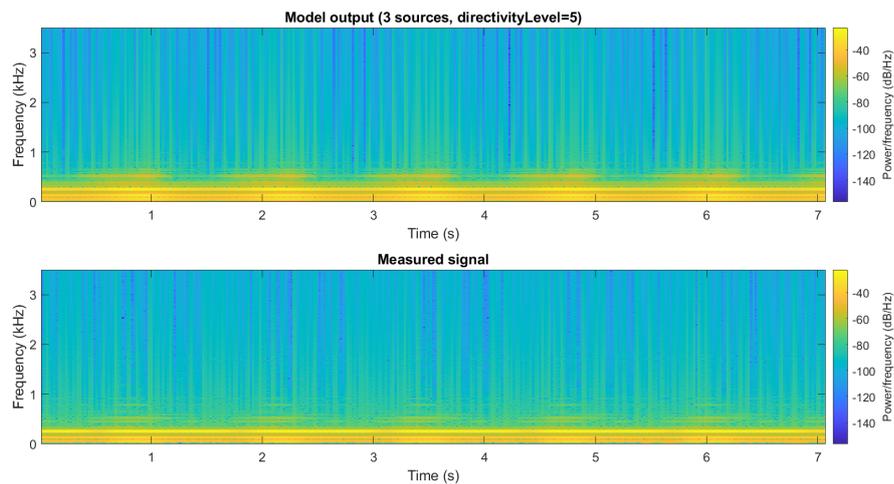


Figura 5.11: Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali e settando $directivityLevel = 5$, e lo spettrogramma del segnale misurato.

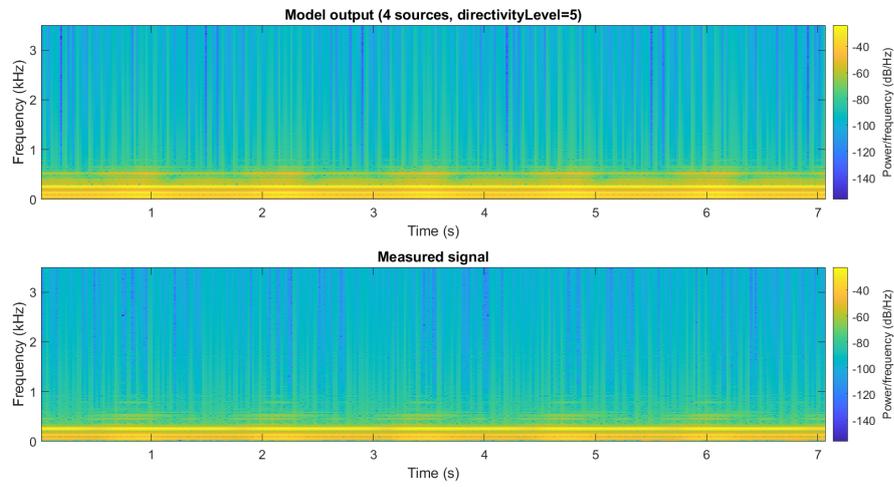


Figura 5.12: Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale e settando $directivityLevel = 5$, e lo spettrogramma del segnale misurato.

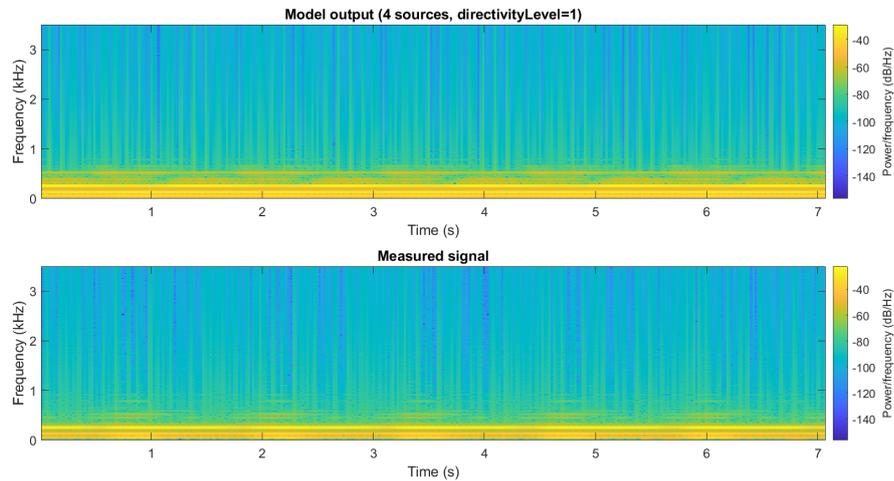


Figura 5.13: Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando le immagini corrispondenti alle pareti laterali-frontale e settando $directivityLevel = 1$, e lo spettrogramma del segnale misurato.

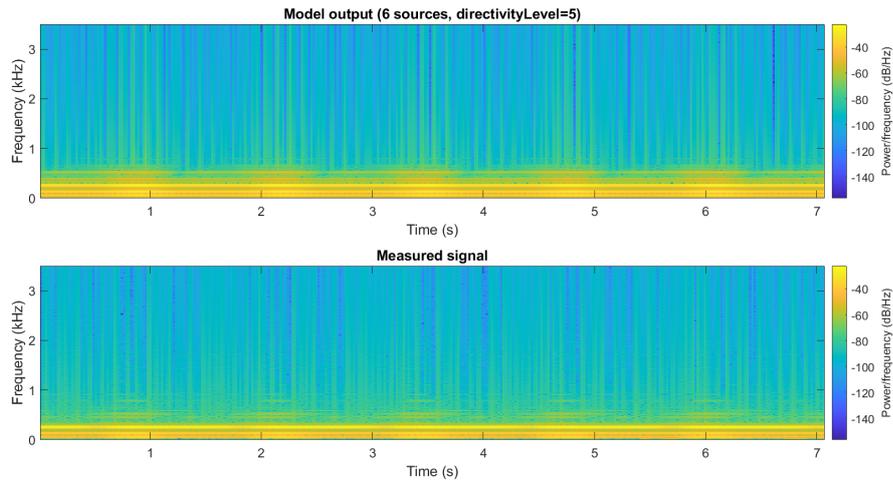


Figura 5.14: Confronto tra lo spettrogramma del segnale in uscita al simulatore, ottenuto considerando tutte le immagini del primo ordine e settando $directivityLevel = 5$, e lo spettrogramma del segnale misurato.

mostrato che l'effetto percettivo più simile si ottiene con 4 sorgenti immagine. Il motivo è dovuto al modo in cui il teorema delle immagini è applicato. Visto che il posizionamento del microfono non è effettuato all'interno del cabinet, considerare tutte le sorgenti in ogni istante di campionamento è un errore. Un esempio è mostrato in Figura 5.15 e in Figura 5.16. L'implementazione andrebbe quindi modificata, ma questo complicherebbe la parte di controllo dell'algoritmo.

Infine, si noti che la modulazione in ampiezza proposta porta ad un miglioramento effettivo del simulatore del woofer.

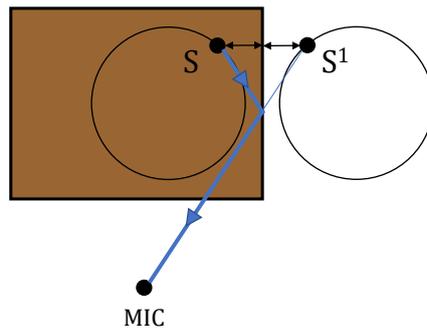


Figura 5.15: Applicazione del teorema delle immagini. S è la sorgente principale e S^1 è la sorgente immagine del primo ordine corrispondente alla parete destra. In questa situazione S^1 è una sorgente immagine valida. La linea blu indica il percorso dal suono corrispondente alla riflessione del primo ordine sulla parete destra.

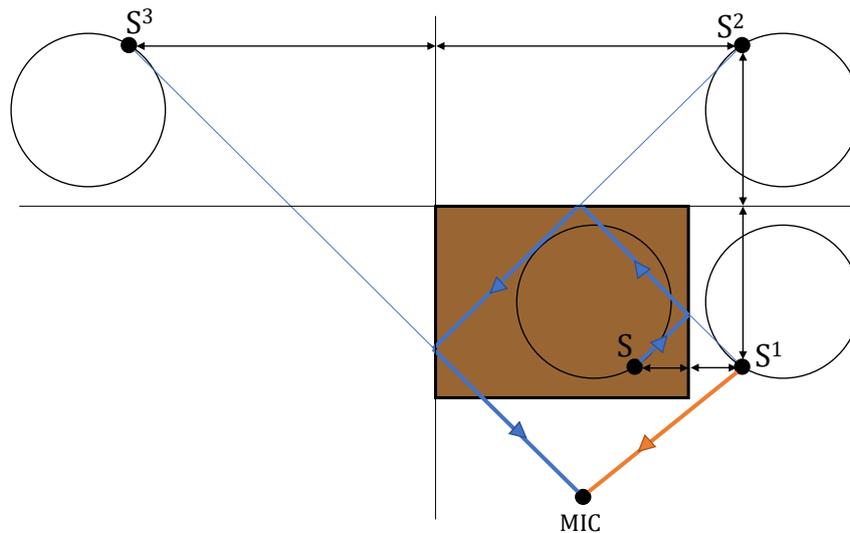


Figura 5.16: Applicazione del teorema delle immagini. S è la sorgente principale, e S^1 , S^2 , S^3 sono, rispettivamente, immagini del primo, secondo e terzo ordine. In questa situazione S^1 non è una sorgente immagine valida. Considerare il percorso arancione del suono è un errore perché la linea non intercetta la parete destra del cabinet. Comunque, nello stesso istante, dovrebbero essere considerate delle riflessioni di ordine superiore. Un esempio è la riflessione del terzo ordine data dal percorso blu.

5.2.4 Filtro di assorbimento dell'aria

Il filtro di assorbimento dell'aria è stato testato per valutarne l'utilità. Nel caso della tromba, che ruota con un raggio di circa 20 cm, dall'Eq. 3.9 si ricava che il filtro produce una variazione massima di guadagno (tra la posizione della sorgente più vicina al microfono e quella più lontana) di 0.2 dB, al di sopra di 10 kHz, che è molto piccola. I test di ascolto hanno confermato che il filtro non è rilevante dal punto di vista percettivo. Nel caso del woofer l'effetto è inutile visto che il segnale non ha componenti in frequenza al di sopra degli 800 Hz. Per questo motivo, il filtro di assorbimento dell'aria può essere rimosso dal modello in modo da evitare il ricalcolo dei coefficienti in ogni istante di campionamento, ottenendo così una significativa diminuzione del costo computazionale. I filtri potrebbero essere mantenuti effettuando il ricalcolo dei coefficienti solamente quando si varia la posizione dei microfoni, per simularne lo spostamento (considerando la distanza tra il microfono e il centro di rotazione della sorgente), ma anche in questo caso l'effetto non è rilevante perché il range di posizionamento dei microfoni copre distanze piccole.

5.2.5 Confronto tra scrittura interpolata e lettura interpolata

In Figura 5.17 e in Figura 5.18 è mostrato il confronto tra l'uscita del simulatore implementato con lo schema che utilizza la lettura interpolata (Figura 4.7) e l'uscita del simulatore implementato con lo schema che utilizza la scrittura interpolata

(Figura 4.8), per un segnale d'ingresso sinusoidale con una frequenza di 2 kHz. Si nota che le differenze tra le forme d'onda e gli spettrogrammi sono minime. Inoltre, i test di ascolto effettuati hanno confermato che i due modelli sono praticamente equivalenti dal punto di vista percettivo. L'utilizzo della lettura interpolata è quindi preferibile perché il costo computazionale è minore e l'implementazione è meno complessa.

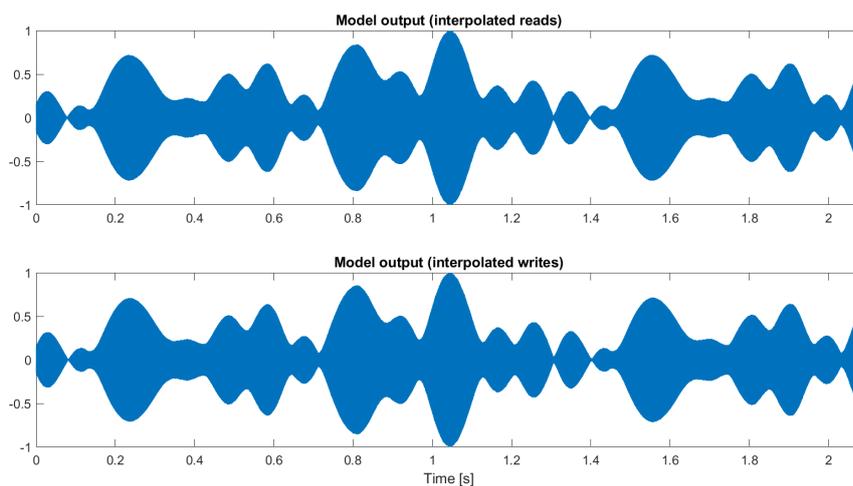


Figura 5.17: Forme d'onda ottenute in uscita del simulatore della tromba per i due possibili schemi implementativi. L'ingresso del modello è un segnale sinusoidale con una frequenza di 2 kHz.

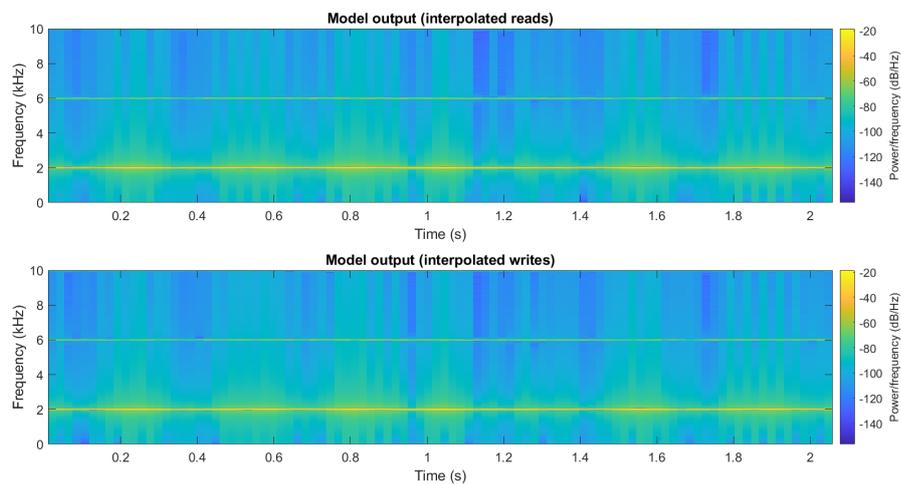


Figura 5.18: Spettrogrammi calcolati sui segnali di Figura 5.17.

Capitolo 6

Conclusioni

Il sistema di diffusione Leslie è stato studiato ed è stato sviluppato un modello computazionalmente efficiente per la sua emulazione digitale.

Nella prima parte di questa tesi, si è descritto il lavoro di Kronland et al. [2], che propongono un modello digitale del Leslie in cui ciascun altoparlante rotante è modellato come un insieme di sorgenti sonore virtuali corrispondenti al suono diretto e alle riflessioni sulle pareti del cabinet, ottenute dall'applicazione del teorema delle immagini. Ogni sorgente virtuale elabora il segnale utilizzando un filtro high-shelving, una delay line e un guadagno, tempo-varianti, che simulano, rispettivamente, le variazioni di timbro dovute all'assorbimento dell'aria, l'effetto Doppler e le variazioni di intensità del suono. Le operazioni sono controllate, in tempo reale, dalle distanze tra le sorgenti e l'ascoltatore. Il modello include anche un filtro che elabora il segnale relativo alle riflessioni in modo da simulare l'assorbimento del legno.

In seguito, il modello del Leslie utilizzato in questo lavoro è stato presentato. Per simulare gli altoparlanti rotanti si utilizza lo schema proposto in [2], in cui la parte relativa alla modulazione in ampiezza del segnale è stata modificata per tener conto della direttività delle sorgenti. Il filtro di assorbimento del legno è stato progettato adattando un filtro IIR del quarto ordine sulla risposta in frequenza ricavata dai dati di assorbimento del legno disponibili in letteratura. Il crossover utilizzato è un Butterworth del secondo ordine con frequenza d'incrocio a 800 Hz. La dinamica delle sorgenti durante i cambi di velocità è stata modellata utilizzando una funzione polinomiale del terzo ordine. Ogni transizione è caratterizzata da un set differente di coefficienti del polinomio, che sono stati ricavati dalle misurazioni effettuate sul vero Leslie. Per il progetto dell'amplificatore sono state testate tre tipologie di modelli. Il primo è un waveshaper. Il secondo è un modello di Wiener-Hammerstein che viene identificato con un approccio basato sul deep learning. I blocchi lineari sono rappresentati dalla cascata di filtri IIR, implementati come layer ricorrenti, e il blocco non lineare consiste in un MLP. Il training è stato effettuato su diverse reti che si differenziano per il numero e il tipo di filtri IIR utilizzati (sono state testate le strutture TDF-II e SVFs). A training terminato, per adattare la rete ad una implementazione real-time, il MLP viene sostituito da una look-up-table. Il terzo modello è sempre del tipo Wiener-Hammerstein ed è identificato con una procedura di ottimizzazione che sfrutta l'algoritmo di Levenberg-Marquardt. I blocchi lineari

sono costituiti, ciascuno, da 12 filtri IIR (di tipo shelving e peak) del secondo ordine e il blocco non lineare ha la particolarità di essere con memoria.

L'implementazione delle delay lines con lunghezza variabile è stata discussa approfonditamente. In particolare, sono stati analizzati due schemi differenti che utilizzano la lettura interpolata e la scrittura interpolata del buffer. Si è mostrato che, sotto l'ipotesi di sorgenti con velocità molto inferiore alla velocità di propagazione del suono, è possibile semplificare l'espressione del calcolo del delay nel metodo basato sulla lettura interpolata. Il suo utilizzo è quindi vantaggioso, rispetto a quello della scrittura interpolata, perché ha un costo computazionale minore.

Infine, il modello è stato implementato come plug-in per il software VCV Rack, in modo da effettuare una parametrizzazione, ed è stato valutato tramite test di ascolto e confrontando l'uscita del simulatore con le registrazioni effettuate su un vero Leslie. I risultati hanno mostrato che il modello riesce a ricreare le caratteristiche principali dell'effetto Leslie, ma una simulazione accurata non è possibile visto l'elevato grado di approssimazione del modello. La modulazione di ampiezza proposta si è dimostrata utile nel caso del woofer. Il filtro di assorbimento dell'aria è stato rimosso dal modello perché non ha un effetto rilevante dal punto di vista percettivo. I test hanno inoltre confermato che lo schema implementativo che utilizza la lettura interpolata è equivalente, all'ascolto, a quello che utilizza la scrittura interpolata. Per quanto riguarda l'amplificatore, i metodi di identificazione testati hanno portato a risultati comparabili. Il waveshaper rappresenta un'alternativa dal basso costo computazionale. Per migliorare la simulazione, anche la risposta in frequenza dei driver e dei diffusori dovrebbe essere modellata perché, nel vero Leslie, modifica fortemente il segnale in uscita dall'amplificatore. Questa parte è riservata a lavori futuri.

Bibliografia

- [1] Clifford A Henricksen. Unearthing the mysteries of the leslie cabinet. *Recording Engineer/Producer Magazine*, 555, 1981.
- [2] Richard Kronland-Martinet and Thierry Voinier. Real-time perceptual simulation of moving sources: application to the leslie cabinet and 3d sound immersion. *EURASIP Journal on Audio, Speech, and Music Processing*, 2008:1–10, 2008.
- [3] Jyri Huopaniemi, Lauri Savioja, and Matti Karjalainen. Modeling of reflections and air absorption in acoustical spaces a digital filter design approach. In *Proceedings of 1997 Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 4–pp. IEEE, 1997.
- [4] Julius Smith, Stefania Serafin, Jonathan Abel, and David Berners. Doppler simulation and the leslie. In *Proc. Int. Conf. on Digital Audio Effects, Hamburg*, 2002.
- [5] Nicolas Tsingos. *Simulation de champs sonores de haute qualité pour des applications graphiques interactives*. PhD thesis, Université Joseph-Fourier-Grenoble I, 1998.
- [6] John M Chowning. The simulation of moving sound sources. *Journal of the audio engineering society*, 19(1):2–6, 1971.
- [7] Jont B Allen and David A Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4):943–950, 1979.
- [8] Timo I Laakso, Vesa Valimaki, Matti Karjalainen, and Unto K Laine. Splitting the unit delay [fir/all pass filters design]. *IEEE Signal Processing Magazine*, 13(1):30–60, 1996.
- [9] VP Valimaki. Discrete-time modeling of acoustic tubes using fractional delay filters. 1998.
- [10] Davide Rocchesso. Fractionally addressed delay lines. *IEEE Transactions on Speech and Audio Processing*, 8(6):717–727, 2000.
- [11] Sophocles J Orfanidis. *Introduction to signal processing*. Prentice-Hall, Inc., 1995.

Bibliografia

- [12] Udo Zölzer, Xavier Serra, Mark Sandler, et al. Digital audio effects, 2011.
- [13] ANSI Standard. S1, 26-1995, “*Method for calculation of the absorption of sound by the atmosphere*”(Standards Secretariat of the Acoustical Society of America, New York, 1995), 1995.
- [14] Jerzy Smardzewski, Wojciech Batko, Tadeusz Kamisiński, Artur Flach, Artur Pilch, Dorota Dziurka, Radosław Mirski, Edward Roszyk, and Adam Majewski. Experimental study of wood acoustic absorption characteristics. *Holzforschung*, 68(4):467–476, 2014.
- [15] Ross Penniman. A rotary speaker modeling plug-in. *Will Pirkle. App Notes. App Note AN-9.* < <http://www.willpirkle.com/app-notes/> > Access on, 5(03):2015, 2014.
- [16] Boris Kuznetsov, Julian D Parker, and Fabián Esqueda. Differentiable iir filters for machine learning applications. In *Proc. Int. Conf. Digital Audio Effects (eDAFx-20)*, pages 297–303, 2020.
- [17] Aaron Wishnick. Time-varying filters for musical applications. In *DAFx*, pages 69–76, 2014.
- [18] Felix Eichas and Udo Zölzer. Gray-box modeling of guitar amplifiers. *Journal of the Audio Engineering Society*, 66(12):1006–1015, 2018.
- [19] Kenneth Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.
- [20] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.
- [21] Leonardo Gabrielli. *Developing Virtual Synthesizers with VCV Rack*. Focal Press, 2020.
- [22] Shahan Nercessian, Andy Sarroff, and Kurt James Werner. Lightweight and interpretable neural modeling of an audio distortion effect using hyperconditioned differentiable biquads. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 890–894. IEEE, 2021.