

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

DIPARTIMENTO DI
INGEGNERIA DELL'ENERGIA ELETTRICA E
DELL'INFORMAZIONE "GUGLIELMO MARCONI"

CORSO DI LAUREA IN INGEGNERIA BIOMEDICA

Titolo dell'elaborato

ADDESTRAMENTO DI UNA RETE NEURALE
BIOLOGICAMENTE PLAUSIBILE PER LA MATURAZIONE
DELLE CAPACITÀ DI PERCEZIONE MULTISENSORIALE

Elaborato in

FONDAMENTI DI STRUMENTAZIONE BIOMEDICA

Relatore

Cristiano Cuppini

Correlatrice

Eleonore Federica Di Rosa

Presentata da

Elisa Buongiorno

Anno Accademico 2024/2025

INDICE

Introduzione.....	
Capitolo 1: Materiali e Metodi.....	1
1.1: Descrizione qualitativa del modello.....	2
1.2: Descrizione matematica del modello.....	5
1.3: Regola di apprendimento hebbiana.....	9
1.4: Modulazione della soglia di attivazione.....	12
1.5: Calibrazione dei parametri α_M e α	13
1.6: Stimoli sequenziali e nuove regole di feedback inibitorio.....	14
1.7: Assegnazione dei parametri.....	16
Capitolo 2: Risultati.....	21
2.1: Influenza della rimodulazione di α_M e α sulla dinamica dei pesi sinaptici.....	21
2.2: Evoluzione dei pesi sinaptici con stimoli singoli e filtraggio tramite soglia.....	23
2.3: Evoluzione dell'attività nelle aree del modello neurale.....	31
2.4: Dinamica dei pesi sinaptici con stimoli sequenziali.....	37
2.5: Differenze nei tempi di reazione tra cambi e ripetizioni in condizioni unisensoriali e multisensoriali.....	46
2.6: Influenza degli intervalli interstimolo sui tempi di reazione.....	49
Capitolo 3: Discussione.....	51
Conclusioni.....	53

Introduzione

Nella vita quotidiana siamo costantemente esposti a una molteplicità di stimoli sensoriali, come luci e suoni, che giungono al cervello in un flusso continuo. Per affrontare in modo efficace questa complessità, il sistema nervoso si è evoluto per elaborare i segnali provenienti da ciascuna modalità e integrarli in rappresentazioni unificate, un processo noto come integrazione multisensoriale. Tale capacità non è innata, ma si sviluppa progressivamente nel corso della vita postnatale precoce, dipendendo in maniera critica da diversi fattori, tra cui le relazioni spaziali e temporali tra gli stimoli cross-modalità.

Evidenze neurofisiologiche hanno mostrato che la risposta di un neurone multisensoriale a coppie di stimoli spazialmente e temporalmente congruenti risulta potenziata rispetto a quella evocata dai singoli stimoli unimodali (Stein e Meredith, 1993). Viceversa, un disallineamento temporale tra le componenti sensoriali può attenuare il potenziamento, o addirittura indurre una depressione della risposta, fenomeno riscontrato quando gli stimoli sono separati da intervalli interstimolo (ISI) superiori a 200–400 ms (Diederich e Colonius, 2004a; Meredith et al., 1987). Questi cambiamenti fisiologici si traducono in modificazioni percettive e comportamentali: la congruenza spazio-temporale degli stimoli è infatti associata a una maggiore accuratezza nella rilevazione e localizzazione, tempi di reazione più rapidi e un processo decisionale migliorato (Bolognini et al., 2007; Diederich e Colonius, 2004a; Senkowski et al., 2007; von Saldern e Noppeney, 2013).

Negli anni, diversi approcci neurocomputazionali sono stati sviluppati per comprendere i meccanismi alla base dell'integrazione multisensoriale. Modelli biologicamente plausibili hanno fornito descrizioni dell'architettura sinaptica ipotizzata e del ruolo delle configurazioni spaziotemporali nella modulazione delle risposte multisensoriali (Ursino et al., 2014). Ad esempio, Cuppini et al. (2020) hanno proposto un modello competitivo tra sistemi uditivi e visivi per spiegare l'effetto di switch di modalità; analogamente, Shaw et al. (2020) e Crosse et al. (2022) hanno mostrato che i tempi di risposta aumentano quando cambia la modalità sensoriale rispetto a quando si ripete. Tuttavia, questi studi hanno generalmente considerato le sinapsi come già mature, trascurando i processi dinamici che ne regolano la maturazione durante lo sviluppo.

Il modello qui presentato affronta questa limitazione, ponendosi l'obiettivo di descrivere i principi computazionali che governano lo sviluppo delle capacità di integrazione multisensoriale. Esso simula i processi attraverso cui le connessioni

sinaptiche tra aree unimodali e multisensoriali si organizzano ed evolvono in risposta all'esperienza. L'addestramento delle sinapsi, implementato secondo regole di plasticità hebbiana, è concepito come un'analogia della maturazione funzionale che caratterizza la crescita individuale: l'esposizione ripetuta a stimoli concorrenti guida la progressiva stabilizzazione e specializzazione dei circuiti, favorendo l'acquisizione di competenze percettive sempre più complesse. In questa prospettiva, il modello non intende solo descrivere l'integrazione audiovisiva in termini statici, ma mostrare come tale capacità emerga dinamicamente dall'interazione tra plasticità sinaptica e stimolazione ambientale. Questo approccio fornisce inoltre un quadro interpretativo utile non solo per comprendere il funzionamento fisiologico delle aree multisensoriali, ma anche per analizzare le possibili deviazioni in condizioni di sviluppo atipico o in disturbi neurocognitivi.

1. Materiali e metodi

Il modello trattato, di natura neurocomputazionale, descrive una rete di unità neuronali definite da parametri specifici, quali soglie di attivazione, costanti di tempo per l'integrazione del segnale e guadagni sinaptici iniziali. Le unità vengono stimulate da tre principali sorgenti: input feedforward, responsabili della trasmissione dei segnali primari; connessioni cross-modali, deputate all'integrazione delle informazioni multisensoriali; e interazioni inibitorie, provenienti sia da altre aree sia dalla stessa area, che modulano l'attività della rete garantendone la stabilità funzionale.

Il codice MATLAB utilizzato implementa la fase di addestramento del modello, nella quale i pesi sinaptici vengono aggiornati in base a specifiche regole di plasticità. Questa fase è stata introdotta per simulare, in termini fisiologici, i meccanismi di adattamento sinaptico che sottendono apprendimento e memoria, riproducendo i processi attraverso cui le connessioni neuronali si rafforzano, si indeboliscono o si stabilizzano in relazione all'attività. Diventa quindi fondamentale valutare come l'efficacia delle diverse tipologie di sinapsi evolva durante l'apprendimento, analizzando i cambiamenti delle connessioni e i pattern di attività che emergono nel tempo. L'addestramento si articola in 4000 epoche, con analisi condotte a intervalli di 1000 epoche, al fine di monitorare in modo progressivo le trasformazioni della connettività sinaptica e dei pattern di attività neuronale.

1.1 Descrizione qualitativa del modello

Il modello è costruito su 4 strati: lo strato di input, lo strato competitivo, l'area multisensoriale e l'area premotoria.

Lo strato di input

È composto da due regioni, organizzate topograficamente, che rappresentano rispettivamente le aree di elaborazione uditiva e visiva (A e V nella Fig. 1).

Ogni unità neurale in queste aree riceve 4 input:

- Uno stimolo esterno, della corrispondente modalità sensoriale, simulato da una funzione gaussiana e filtrato dai campi recettivi (RFs) delle unità neurali. Il punto centrale della gaussiana coincide con il punto di applicazione dello stimolo nel mondo esterno, mentre l'effetto combinato dell'ampiezza dei RF dei neuroni unisensoriali e dell'affidabilità degli stimoli esterni è rappresentato dalla deviazione standard (σ_a e σ_v). Come mostrato in Ursino et al., 2017, entrambi i termini influenzano la posizione percepita degli input neurali.

L'inizio e la durata sono stati scelti per riprodurre l'assetto sperimentale di Crosse et al., 2022.

Per affrontare la variabilità sperimentale, l'efficacia degli input esterni (I_0^v e I_0^a , per la modalità visiva e uditiva, rispettivamente) è scelta casualmente da una distribuzione uniforme (vedi Tabella 1).

- Un input laterale da elementi nella stessa regione, che è eccitatorio per le unità prossimali e inibitorio per quelle distali, ed è mediato da sinapsi intra-area con forma a cappello messicano (L_v e L_a).
- Un input cross-modale, proveniente da unità corrispondenti nell'altra regione di input. In questo modo, le unità neurali che coprono porzioni adiacenti dello spazio esterno nelle due regioni scambiano sinapsi cross-modali eccitatorie (W_{av} e W_{va} in Fig. 1), caratterizzate da distribuzione gaussiana e dinamica rapida.

L'efficacia di queste connessioni è sufficientemente forte da indurre aumenti significativi nell'attività dell'altra area sensoriale in caso di stimolazione multisensoriale, ma non in condizioni unisensoriali. Ciò è in linea con i risultati di studi su animali (Ghazanfar & Schroeder, 2006).

- Un input inibitorio da unità interneurali dell'altra modalità sensoriale. Quando stimoli di diverse modalità sensoriali sono presentati in sequenza, questo input crea una competizione cross-sensoriale. Le sinapsi inibitorie di

feedback che mediano questo meccanismo hanno forma gaussiana e sono caratterizzate da dinamica lenta ($L_a I_v$ e $L_v I_a$ in Fig.1).

La posizione percepita degli stimoli esterni viene inferita nelle regioni di input: in particolare, viene calcolata come il baricentro delle attività evocate in ciascuna area. A causa della presenza delle proiezioni cross-modali, la localizzazione spaziale inferita degli stimoli uditivi o visivi è influenzata dalla presentazione contemporanea dello stimolo nell'altra modalità sensoriale, anche se i due eventi unisensoriali vengono elaborati separatamente.

Le sinapsi eccitatorie a forma di gaussiana ($W I_a$ e $W I_v$) proiettano dalle aree di input verso aree interneurali specifiche per modalità (I_a e I_v), che fanno parte dello strato competitivo descritto più avanti.

Connessioni eccitatorie a lungo raggio (W_{ma} e W_{mv} in Fig. 1) proiettano verso un'area multisensoriale (regione M nella Fig. 1).

Lo strato competitivo

È formato da due aree di unità interneurali, una per ciascuna modalità.

Gli elementi in queste aree:

- Sono eccitati da elementi della propria area di input della stessa modalità;
- Ricevono anche un input inibitorio dall'altra area interneurale, secondo un meccanismo competitivo Winner-Takes-All (il vincitore inibisce il perdente).

Con questa struttura, la modalità sensoriale vincente inibisce l'altra modalità, attraverso le sinapsi inibitorie di feedback ($L_a I_v$ e $L_v I_a$) che comunicano con le regioni unisensoriali.

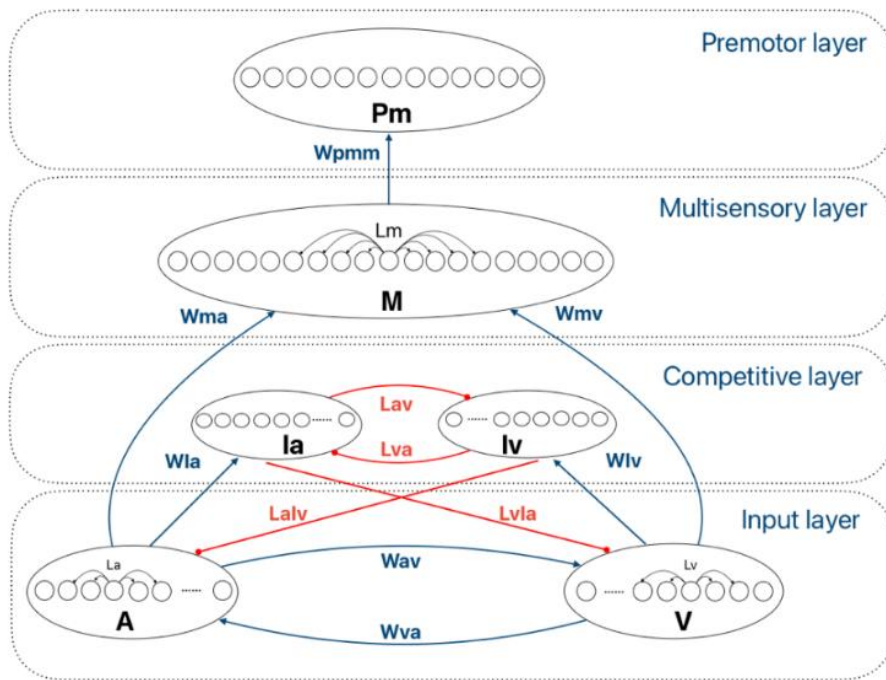
Esistono evidenze sperimentali a favore della competizione tra modalità sensoriali, ma i meccanismi neurali che potrebbero sostenerla non sono ancora noti. Questo strato competitivo potrebbe trovare una implementazione fisiologica in regioni di ordine superiore, come suggerito da Huang et al., 2015 e da Hairston et al., 2008 in uno studio sull'analisi di stimoli uditivi nel dominio temporale. Tuttavia, proiezioni dirette tra aree visive e uditive centrali e cortecce associative (Clavagnier et al., 2004), oppure una modulazione sensoriale funzionale (Bizley & King, 2009; Yu et al., 2013), potrebbero anch'esse essere responsabili di questo tipo di competizione.

L'area multisensoriale

Le unità in quest'area ricevono sinapsi eccitatorie feedforward a forma di gaussiana (W_{mv} , W_{ma}) da elementi delle aree unisensoriali sensibili alla stessa porzione di spazio, e un input laterale generato da altre unità nella regione, tramite sinapsi intra-area a cappello messicano (L_m).

L'area premotoria

In quest'area, ogni unità neurale è connessa all'unità corrispondente nell'area multisensoriale. Questa area è utilizzata per simulare le risposte comportamentali dei soggetti agli stimoli esterni, e la sua attività evocata viene confrontata con una soglia fissa (30% dell'attività massima dei neuroni) per valutare i tempi di reazione (RT) agli stimoli esterni.



1.2 Descrizione matematica del modello

Ogni area è composta da 180 unità neurali, distanti 1° l'una dall'altra, e ciascuna sensibile a una porzione specifica dello spazio esterno.

Dal punto di vista computazionale, ogni unità è descritta da una dinamica di primo ordine, che simula le proprietà integrative della membrana cellulare, e di una relazione sigmoidale allo stato stazionario che simula la presenza di una soglia inferiore e di una saturazione superiore per l'attivazione neurale.

Di seguito, ogni elemento sarà indicato con un apice, r , riferito a una regione specifica del modello ($r = a, v, m, ia$ o iv). $U(t)$ e $y(t)$ sono usati per rappresentare l'input e l'output netti di un dato elemento neurale al tempo t , rispettivamente.

Durante la simulazione, ad ogni passo temporale dt , il potenziale di ciascun neurone viene aggiornato secondo tale equazione differenziale:

$$\tau \frac{dy^r(t)}{dt} = -y^r(t) + F(u^r(t)) \quad (1)$$

dove il termine $-y^r(t)$ rappresenta il decadimento intrinseco, cioè la perdita spontanea di carica che tende a riportare il potenziale verso il valore di riposo in assenza di stimoli. La funzione di trasferimento sigmoidale $F(u^r(t))$ rappresenta invece l'input totale ricevuto dal neurone, proveniente da altre unità o da stimoli esterni, e agisce ad aumentare il potenziale:

$$F(u^r) = \frac{1}{1 + e^{-s(u^r - \theta)}} \quad (2)$$

dove u è l'input netto, s è la pendenza della curva e θ la soglia di attivazione. Questa funzione restituisce valori prossimi a zero per potenziali molto inferiori alla soglia, valori prossimi a uno per potenziali che la superano nettamente, e una transizione graduale tra i due estremi. In tal modo, il modello trasforma il segnale elettrico continuo in un output normalizzato e limitato, che può essere utilizzato come input per gli altri neuroni nel passo temporale successivo della simulazione.

Per semplicità, in questo lavoro tutti gli elementi neurali sono descritti utilizzando gli stessi parametri e la stessa costante di tempo. L'input netto che raggiunge uno specifico elemento neurale (cioè, la quantità $u^r(t)$, nell'Eq. (1)) dipende dalla regione a cui appartiene.

Le aree di input - Gli elementi in queste regioni elaborano separatamente gli stimoli esterni uditivi e visivi ($r = a, v$). Il loro input netto è il risultato di tre componenti. La prima, la componente esterna, è l'input unisensoriale $e^r(t)$, proveniente dal

mondo esterno. La seconda, la componente cross-modale, è l'input $c^r(t)$, dall'area che elabora l'altra modalità sensoriale, trasmesso al neurone bersaglio attraverso sinapsi eccitatorie. L'ultima, la componente inibitoria, è il contributo delle sinapsi inibitorie di feedback dell'interneurone attivato dall'altra modalità sensoriale, $l^r(t)$. L'input esterno è caratterizzato dalla sua efficacia I_0^r e dalla sua durata T^r . Assumendo uno stimolo della modalità sensoriale r ($r = a$ o v):

$$e^r(t) = \begin{cases} I_0^r, & 0 \leq t \leq T^r \\ 0, & t \geq T^r \end{cases} \quad (3)$$

L'input cross-modale, $c^r(t)$, è ottenuto assumendo che ciascun elemento riceva un'eccitazione dalla regione che elabora l'altra modalità definita come:

$$\begin{aligned} c^a(t) &= W_{av} \cdot y^v(t - \Delta t) \\ c^v(t) &= W_{va} \cdot y^a(t - \Delta t) \end{aligned} \quad (4)$$

Dove W_{va} e W_{av} sono i pesi di questa eccitazione reciproca. Nel modello sono uguali per entrambe le regioni ($W_{va} = W_{av} = W$), e Δt è un ritardo che rappresenta la latenza con cui gli input cross-modalità vengono scambiati tra le due regioni.

Infine, l'input inibitorio, $l^r(t)$, è l'effetto dell'interneurone eccitato dall'altra regione sensoriale che interagisce con l'elemento target attraverso sinapsi inibitorie. Pertanto, l'input che un elemento unisensoriale riceve dall'interneurone dell'altra modalità è definito come:

$$\begin{aligned} l^a(t) &= L_a I_v \cdot y^{iv}(t) \\ l^v(t) &= L_v I_a \cdot y^{ia}(t) \end{aligned} \quad (5)$$

dove $y^{ia}(t)$ e $y^{iv}(t)$ sono le attività degli interneuroni uditivi e visivi presinaptici rispettivamente, mentre $L_a I_v$ e $L_v I_a$ sono le forze delle sinapsi inibitorie. Queste sinapsi sono simmetriche ($L_a I_v = L_v I_a = LI$). Nel modello, non incorporiamo un ritardo per l'inibizione cross-sensoriale, perché la dinamica dell'effetto inibitorio è stata scelta molto più lunga di qualsiasi altro meccanismo della rete; quindi, l'effetto di un ritardo è già incluso nella costante di tempo scelta per queste proiezioni.

Interneuroni inibitori - gli elementi in queste due regioni ($r = ia, iv$) sono rispettivamente eccitati dalle aree di input uditivo e visivo, e scambiano proiezioni inibitorie, implementando un meccanismo di «Winner-Takes-All» (WTA). Il loro

input netto è il risultato di uno stimolo eccitatorio, $I_{ex}^r(t)$, proveniente dalla corrispondente regione di input unisensoriale attraverso sinapsi eccitatorie; e di una componente inibitoria, $I_{in}^r(t)$, prodotta da sinapsi inibitorie dell'altro interneurone. Le componenti eccitatorie, $I_{ex}^r(t)$, che mirano agli interneuroni uditivi e visivi sono definite come:

$$\begin{aligned} I_{ex}^{ia}(t) &= WI_a \cdot y^a(t) \\ I_{ex}^{iv}(t) &= WI_v \cdot y^v(t) \end{aligned} \quad (6)$$

Dove WI_a e WI_v sono i pesi delle connessioni eccitatorie da una regione di input unisensoriale al suo corrispondente elemento interneuronale, assunti gli stessi per entrambe le modalità sensoriali ($WI_a = WI_v = WI$).

L'input inibitorio, $I_{in}^r(t)$, che un interneurone riceve dall'interneurone dell'altra modalità, attraverso sinapsi inibitorie, è definito come:

$$\begin{aligned} I_{in}^{ia}(t) &= L_{av} \cdot y^{iv}(t) \\ I_{in}^{iv}(t) &= L_{va} \cdot y^{ia}(t) \end{aligned} \quad (7)$$

Dove $y^{ia}(t)$ e $y^{iv}(t)$ sono le attività degli interneuroni uditivi e visivi presinaptici rispettivamente, mentre L_{av} e L_{va} sono le forze delle connessioni inibitorie reciproche. Queste sinapsi simmetriche ($L_{av} = L_{va} = L$) implementano il meccanismo WTA tra le due aree. Anche in questo caso, come nell'equazione (5), non includiamo un ritardo puro per lo stesso motivo esposto sopra.

Area di output multisensoriale - questa regione ($r = m$) riceve un input netto che è la somma degli stimoli, trasmessi da sinapsi eccitatorie a lungo raggio, dalle aree di input uditivo e visivo. Il suo input netto, $ex^m(t)$, è definito come:

$$ex^m(t) = \sum_r Wm_r \cdot y^r(t - \Delta t^m); r = a, v \quad (8)$$

Dove $Wm_a = Wm_v = Wm$ sono i pesi delle connessioni eccitatorie dalle regioni di input unisensoriali all'area multisensoriale e Δt^m è un ritardo, che rappresenta la minima latenza con cui gli stimoli provenienti dalle regioni di input sono in grado di generare risposte comportamentali.

Per semplicità e per ridurre il numero di assunzioni del modello, tutte le sinapsi precedentemente descritte sono simmetriche per le due modalità sensoriali.

Dinamiche di ciascun componente di input - Tutte le quantità precedenti (Eqs. (3) a (8)) influenzano l'input $u^r(t)$ dell'elemento postsinaptico corrispondente attraverso un'equazione differenziale di secondo ordine. Denotando con $o_i(t)$ l'uscita dell'equazione differenziale per la sorgente di input generica $i(t)$ (descritta da una qualsiasi delle Eqs. (3) a (8)), abbiamo:

$$\begin{cases} \frac{d}{dt} o_i(t) = \delta_i(t) \\ \frac{d}{dt} \delta_i(t) = \frac{G_i^r}{(\tau_i^r)^2} i(t) - \frac{2 \cdot \delta_i(t)}{\tau_i^r} - \frac{o_i(t)}{(\tau_i^r)^2} \end{cases} \quad (9)$$

Dove G_i^r rappresenta il guadagno e τ_i^r definisce la costante di tempo della dinamica, per ogni regione, r , e componente di input, i (Eqs. (3)-(8)). L'equazione (9) implementa una risposta all'impulso di secondo ordine con due poli reali coincidenti. Questo è usato frequentemente nella modellazione neurale per imitare la dinamica sinaptica (vedi Cuppini et al., 2014; Jansen & Rit, 1995; Wendling et al., 2002). Nel modello, per ridurre il numero di parametri, scegliamo gli stessi valori per G_i^r e τ_i^r , per ogni connessione (vedi Tabella 1), eccetto due casi: (1) gli stimoli esterni e, (2) le sinapsi di feedback che implementano il meccanismo inibitorio cross-sensoriale.

Secondo la descrizione precedente, l'input totale (chiamato $u^r(t)$) ricevuto da un neurone nella regione r , è calcolato come segue: (1) per le regioni di input, è la somma del componente esterno Eq. (3), del termine cross-modale Eq. (4) e del feedback inibitorio Eq. (5), filtrati attraverso l'equazione di secondo ordine (Eq. (9)):

$$u^r(t) = o_e(t) + o_c(t) + o_l(t); \quad r = a, v \quad (10)$$

(2) per gli interneuroni inibitori, è la somma dell'eccitazione dalla regione di input (Eq. (6)) e dell'effetto del meccanismo WTA (Eq. (7)), filtrata dall'Eq. (9):

$$u^r(t) = o_{l_ex}(t) + o_{l_in}(t); \quad r = ia, iv \quad (11)$$

(3) per la regione di output, è l'effetto delle sinapsi eccitatorie feedforward (Eq. (8)), filtrato dall'equazione differenziale precedentemente descritta (Eq. (9)):

$$u^m(t) = o_{ex}(t). \quad (12)$$

1.3 Regola di apprendimento hebbiana

Nel codice, l'apprendimento sinaptico si basa su una regola di tipo hebbiano: una connessione si rafforza quando entrambi i neuroni sono attivi simultaneamente. Il potenziamento è modulato da un termine di normalizzazione che impedisce ai pesi di superare un valore massimo prestabilito, mentre un termine di decadimento proporzionale ne riduce gradualmente l'entità in assenza di co-attivazione, prevenendo saturazioni e garantendo la stabilità della rete.

Questa logica viene applicata coerentemente a tutte le categorie di connessioni:

- le feedforward si rinforzano in seguito alla ripetuta esposizione agli stimoli primari;
- le cross-modalità aumentano la loro efficacia quando emergono correlazioni significative tra input di modalità diverse;
- le inibitorie si regolano dinamicamente per mantenere un bilanciamento funzionale nell'attività di rete.

Il sistema è in grado così di adattarsi progressivamente all'esperienza, mantenendo al contempo la stabilità e limitando il rischio di saturazione eccessiva.

Inoltre, per garantire che l'apprendimento sinaptico sia guidato esclusivamente da attività neuronali effettivamente presenti, i segnali di input e output vengono trasformati attraverso una rettificazione positiva (*positive rectification* o *half-wave rectification*). In questo modo, solo le componenti positive delle attività neuronali contribuiscono alla modifica dei pesi sinaptici, mentre i valori negativi vengono annullati. Dal punto di vista neurobiologico, questa scelta riflette il fatto che la frequenza di scarica di un neurone non può assumere valori negativi: quando un neurone non è attivo, la sua attività è considerata nulla e non partecipa ai processi di potenziamento o di depressione sinaptica.

Nelle equazioni sotto riportate, i simboli x^+ indicano l'attività neurale dopo l'applicazione della rettificazione positiva. L'indice in basso specifica l'area sensoriale di appartenenza: (*v*) per l'area visiva, (*a*) per l'area acustica e (*m*) per l'area multisensoriale, mentre il pedice "*pre*" o "*post*" distingue rispettivamente l'attività del neurone presinaptico e di quello postsinaptico.

I pesi sinaptici sono denotati da W^{ij} e LI^{ij} , dove *i* si riferisce ai neuroni postsinaptici, mentre *j* a quelli presinaptici. I valori W^{max} e LI^{max} indicano invece i limiti superiori che i pesi possono raggiungere, costituendo un vincolo di normalizzazione. I parametri α e α_M controllano la velocità di apprendimento, determinando l'ampiezza degli aggiornamenti sinaptici a ogni passo.

Applicazione della regola di Hebb alle feedforward:

$$x_{m_{post},i}^+ = \frac{x_{m,i}^T + |x_{m,i}^T|}{2} = \max(0, x_{m,i}^T)$$

$$x_{a_{pre},j}^+ = \frac{x_{a,j}^T + |x_{a,j}^T|}{2} = \max(0, x_{a,j}^T)$$

$$\Delta W_{ma}^{ij} = \alpha_M (W_{ma}^{max} - W_{ma}^{ij}) x_{m_{post},i}^+ x_{a_{pre},j}^+ - \alpha_M x_{m_{post},i}^+ \frac{W_{ma}^{ij}}{W_{ma}^{max}}$$

$$W_{ma}^{ij} = W_{ma}^{ij} + \Delta W_{ma}^{ij}$$

$$x_{m_{post},i}^+ = \frac{x_{m,i}^T + |x_{m,i}^T|}{2} = \max(0, x_{m,i}^T)$$

$$x_{v_{pre},j}^+ = \frac{x_{v,j}^T + |x_{v,j}^T|}{2} = \max(0, x_{v,j}^T)$$

$$\Delta W_{mv}^{ij} = \alpha_M (W_{mv}^{max} - W_{mv}^{ij}) x_{m_{post},i}^+ x_{v_{pre},j}^+ - \alpha_M x_{m_{post},i}^+ \frac{W_{mv}^{ij}}{W_{mv}^{max}}$$

$$W_{mv}^{ij} = W_{mv}^{ij} + \Delta W_{mv}^{ij}$$

Applicazione della regola di Hebb alle cross-modal:

$$x_{v_{post},i}^+ = \frac{x_{v,i}^T + |x_{v,i}^T|}{2} = \max(0, x_{v,i}^T)$$

$$x_{a_{pre},j}^+ = \frac{x_{a,j}^T + |x_{a,j}^T|}{2} = \max(0, x_{a,j}^T)$$

$$\Delta W_{va}^{ij} = \alpha (W_{va}^{max} - W_{va}^{ij}) x_{v_{post},i}^+ x_{a_{pre},j}^+ - \alpha x_{v_{post},i}^+ \frac{W_{va}^{ij}}{W_{va}^{max}}$$

$$W_{va}^{ij} = W_{va}^{ij} + \Delta W_{va}^{ij}$$

$$x_{a_{post},i}^+ = \frac{x_{a,i}^T + |x_{a,i}^T|}{2} = \max(0, x_{a,i}^T)$$

$$x_{v_{pre},j}^+ = \frac{x_{v,j}^T + |x_{v,j}^T|}{2} = \max(0, x_{v,j}^T)$$

$$\Delta W_{av}^{ij} = \alpha(W_{av}^{max} - W_{av}^{ij})x_{a_{post},i}^+ x_{v_{pre},j}^+ - \alpha x_{a_{post},i}^+ \frac{W_{av}^{ij}}{W_{av}^{max}}$$

$$W_{av}^{ij} = W_{av}^{ij} + \Delta W_{av}^{ij}$$

Applicazione della regola di Hebb alle inibitorie dagli interneuroni alle regioni A e V:

$$x_{a_{post},i}^+ = \frac{x_{a,i}^T + |x_{a,i}^T|}{2} = \max(0, x_{a,i}^T)$$

$$x_{Iv_{pre},j}^+ = \frac{x_{Iv,j}^T + |x_{Iv,j}^T|}{2} = \max(0, x_{Iv,j}^T)$$

$$\Delta L_a I_v^{ij} = \alpha_M(L_a I_v^{max} - L_a I_v^{ij})x_{a_{post},i}^+ x_{Iv_{pre},j}^+ - \alpha x_{a_{post},i}^+ \frac{L_a I_v^{ij}}{L_a I_v^{max}}$$

$$L_a I_v^{ij} = L_a I_v^{ij} + \Delta L_a I_v^{ij}$$

$$x_{v_{post},i}^+ = \frac{x_{v,i}^T + |x_{v,i}^T|}{2} = \max(0, x_{v,i}^T)$$

$$x_{Ia_{pre},j}^+ = \frac{x_{Ia,j}^T + |x_{Ia,j}^T|}{2} = \max(0, x_{Ia,j}^T)$$

$$\Delta L_v I_a^{ij} = \alpha_M(L_v I_a^{max} - L_v I_a^{ij})x_{v_{post},i}^+ x_{Ia_{pre},j}^+ - \alpha x_{v_{post},i}^+ \frac{L_v I_a^{ij}}{L_v I_a^{max}}$$

$$L_v I_a^{ij} = L_v I_a^{ij} + \Delta L_v I_a^{ij}$$

1.4 Modulazione della soglia di attivazione

Ad ogni passo temporale, viene registrata soltanto la porzione di attività presinaptica e postsinaptica che supera un valore prestabilito, mentre quella al di sotto della soglia viene scartata. In questo modo si evitano variazioni poco significative nei pesi dovute a fluttuazioni casuali o a segnali deboli. Nella versione iniziale del modello, tale filtraggio era applicato unicamente alle connessioni feedforward; successivamente è stato esteso anche alle connessioni cross-modali e a quelle inibitorie, così da uniformare il trattamento di tutte le tipologie sinaptiche e accrescere la robustezza complessiva del processo di apprendimento.

Il valore registrato corrisponde sempre alla differenza tra l'attività e la soglia, rappresentando "di quanto" il neurone supera il livello minimo di attivazione.

Prima di essere usati per l'aggiornamento, questi valori vengono ulteriormente trasformati per eliminare eventuali componenti negative, azzerando ogni contributo di segno opposto a quello eccitatorio. In questo modo, la plasticità si basa esclusivamente sulle co-attivazioni positive tra neuroni presinaptici e postsinaptici.

1.5 Calibrazione dei parametri α_M e α

È stata svolta un'attività di calibrazione dei parametri α_M e α nella regola di apprendimento hebbiana, applicata sia alle connessioni feedforward che a quelle cross-modali. L'obiettivo era regolare la velocità di potenziamento delle due tipologie di connessioni, in modo che le feedforward potessero raggiungere una configurazione stabile prima che le cross-modali entrassero in una fase di potenziamento significativo.

Nella configurazione iniziale del modello, le connessioni cross-modali tendevano a crescere più rapidamente, raggiungendo la saturazione quando le feedforward erano ancora in fase di consolidamento. Questo comportamento riduceva il margine di adattamento delle connessioni primarie, ostacolando l'ottenimento di una gerarchia funzionale ottimale.

Modificando i valori di α_M e α , che determinano il tasso di aggiornamento sinaptico, è stato possibile rallentare la dinamica di apprendimento delle connessioni cross-modali rispetto a quelle feedforward, favorendo così un ordine di consolidamento in cui queste ultime si stabilizzano per prime, seguite dalle connessioni cross-modali. Tale riorganizzazione temporale ha reso il processo di apprendimento più coerente con gli obiettivi del modello. Tuttavia, l'obiettivo non è stato ancora raggiunto in maniera del tutto soddisfacente e saranno necessari ulteriori adattamenti per perfezionare il comportamento del sistema.

1.6 Stimoli sequenziali e nuove regole di feedback inibitorio

È stata sviluppata una nuova versione del codice di addestramento, nella quale sono stati introdotti i seguenti cambiamenti:

- Nella fase iniziale, l'addestramento prevedeva la presentazione di stimoli singoli, che potevano essere visivi (V), uditivi (A) o multisensoriali (AV). Poiché il meccanismo di cross-sensory inhibition (implementato dal loop WI e dalle sinapsi LvIa/LaIv) si manifesta in presenza di stimoli consecutivi, la sola presentazione di stimoli singoli non consentiva un addestramento biologicamente plausibile delle sinapsi LvIa e LaIv. Nella versione corrente, la configurazione degli stimoli in ingresso può essere selezionata tra cinque opzioni, con stimoli separati da un intervallo temporale fisso (ISI):
 1. Repeat V: due stimoli visivi consecutivi;
 2. Switch V ($A \rightarrow V$): un acustico seguito da un visivo;
 3. Repeat A: due stimoli acustici consecutivi;
 4. Switch A ($V \rightarrow A$): un visivo seguito da un acustico;
 5. Repeat AV: stimolo cross-modale ripetuto.

Le sequenze di stimoli unisensoriali (solo uditivo, A e solo visivo, V) e multisensoriali (input audiovisivi, AV) sono state presentate in ordine randomizzato a ISI casuali tra 1000 e 3000 ms, e con nessuna discrepanza spaziale, come in Cuppini et al. (2020). I tempi di reazione simulati (RTs) sono stati calcolati come l'intervallo di tempo tra l'istante della presentazione dell'input e l'istante in cui l'attività evocata nell'area di output raggiunge la soglia percettiva. I risultati sono stati analizzati in base a ciascuna modalità di input e alla modalità sensoriale dell'input precedente, discriminando tra prove "Ripetizione" (lo stimolo precedente è della medesima modalità sensoriale dello stimolo attuale considerato, prova "Rp") e prove "Cambio" (lo stimolo precedente non è della medesima modalità sensoriale dello stimolo attuale considerato, prova "Sw").

L'efficacia di ciascun input è stata scelta casualmente da una distribuzione uniforme per replicare la variabilità intra-soggetto degli stimoli sensoriali in un ambiente reale. Per ogni configurazione di stimolo, è stata calcolata la media degli RTs su 100 simulazioni, e i risultati sono stati poi confrontati con quelli di Cuppini et al. (2022) e con i dati empirici di Crosse et al. (2022).

- È stato modificato l'addestramento delle connessioni inibitorie di feedback: la sinapsi (es. LaIv, e in modo analogo LvIa) viene ora aggiornata quando

l'attività dell'area unisensoriale relativa al secondo stimolo supera la soglia, a condizione che l'interneurone eteromodale sia andato sopra soglia circa 1 s prima in seguito al primo stimolo. A tal fine è stato introdotto il parametro $delay_int_add = 200/dt$.

1.7 Assegnazione dei parametri

Il valore di tutti i parametri del modello (vedi Tabella 1 e 2) è stato adattato partendo dai dati presenti in letteratura e dai valori dei parametri del modello in Cuppini et al., 2020, considerando quanto segue:

1. L'architettura del modello è molto simile, tranne per l'aggiunta dell'area premotoria.
2. Diversamente da Cuppini et al., 2020, ogni area ora è composta da $N = 180$ elementi per tenere conto della dimensione spaziale.
3. Dato il punto 2), la maggior parte delle connessioni sinaptiche sono ora modellate con una distribuzione gaussiana, come in Cuppini et al., 2017.

Tabella 1 - Parametri del modello

Sezione	Parametro	Valore	Note
Neuroni	dim	180	
	Θ	25	
	s	0.3	
	τ	3 ms	
Input	G_i^r	75	i = e, m, pm
	τ_i^r	15 ms	i = I_ex, I_in
	τ_e^a	15 ms	
	τ_e^v	τ_e^v	
	τ_{cross}	3 ms	
	τ_i^r	3 ms	i = m, pm
	I_0^a	[1.35 – 1.97]	
	T^a	60 ms	
	I_0^v	[1.8 – 2.6]	
	T^v	60 ms	
	Δt	16 ms	
	Δt^m	50 ms	
	AV ratio	3/5	
Inibizione	G_l^r	750	r = a, v
	τ_l^r	180 ms	r = a, v

Tabella 2 - Parametri delle sinapsi

Sezione	Parametro	Valore
Proiezione feedforward all'area multisensoriale	Wma0	0.12
	Wmv0	Wma0/AV ratio
	Wma SD = Wmv SD	5
Area cross-modale	Wva0	0.12
	Wav0	Wma0/AV ratio
	Wav SD	7
	Wva SD	5
Inibizione laterale intra-area a cappello messicano	LLm0_in	0.3
	kLm	0.4/0.3
	LLm0_ex	kLm*LLm0_in
	LLm SD_in	8
	LLm SD_ex	4
	LLa0_in	0.3
	kLa	5/4
	LLa0_ex	kLa*LLa0_in
	LLa SD_in	120
	LLa SD_ex	3
	LLv0_in	4
	kLv	5/4
	LLv0_ex	kLv*LLa0_in
	LLv SD_in	120
	LLv SD_ex	3
Proiezione feedforward eccitatoria verso gli interneuroni	WIa0	0.2
	WIv0	WIa0/AV ratio
	WIa SD = WIv SD	5
Inibizione di feedback dagli interneuroni alle aree primarie	LaIv0	0.0057
	LvIa0	0.016
	LaIv SD = LvIa SD	0.5
Meccanismo WTA (Winner-Take-All) tra interneuroni con connettività uno-a-uno	L_WTA0	10
Proiezione feedforward verso l'area premotoria con connettività uno-a-uno	W_pmm	1

Nel modello analizzato, tutte le variabili temporali non sono espresse in millisecondi reali, ma in numero di passi di simulazione, perché la dinamica non viene calcolata in tempo continuo ma in tempo discreto.

In un sistema biologico, l'attività neuronale evolve in modo continuo, con potenziali di membrana, correnti sinaptiche e altre grandezze che variano senza interruzione nel tempo. Un calcolatore, invece, non può rappresentare questa continuità in maniera esatta: deve suddividere il tempo in intervalli finiti e aggiornare lo stato del sistema a intervalli regolari, definiti dal passo temporale dt .

Tutte le variabili legate al tempo (dalle costanti di decadimento delle correnti sinaptiche, ai ritardi di trasmissione, fino alle durate degli stimoli) vengono convertite da valori in millisecondi a valori espressi in multipli di dt , così che il modello possa operare in unità coerenti con la discretizzazione numerica.

I singoli neuroni

L'ascissa centrale, θ , è stata scelta in modo da produrre un'attivazione trascurabile quando l'input è nullo. La pendenza della relazione sigmoideale, s , è responsabile di una transizione graduale dall'attività nulla alla saturazione in risposta a stimoli esterni. La costante di tempo è in accordo con i valori solitamente riscontrati nelle equazioni deterministiche di campo medio (Ben-Yishai et al., 1995; Treves, 1993).

Input esterno $e_j^i(t)$

Studi fisiologici hanno dimostrato che i neuroni della corteccia uditiva presentano latenze più brevi rispetto ai neuroni della corteccia visiva. Pertanto, la costante di tempo è stata scelta più lenta per lo stimolo visivo rispetto a quella per lo stimolo uditivo ($\tau^a < \tau^v$). In particolare, τ^a riproduce un'attivazione dell'area uditiva che avviene dopo 25-30 ms dall'inizio dello stimolo esterno. Poiché le costanti di tempo rappresentano il tempo necessario affinché l'attività nelle regioni di input raggiunga il 90% del valore di regime in risposta a un input a gradino, abbiamo scelto $\tau^a = 15$ ms. D'altra parte, τ^v è stato fissato a 25 ms, per simulare l'effetto di uno stimolo visivo che produce una risposta rilevabile 45-50 ms dopo il suo inizio.

Il sistema visivo è anche noto per presentare una migliore risoluzione spaziale rispetto a quello uditivo (vedi Recanzone, 2009; Recanzone et al., 2000). Pertanto, σ^v è stato impostato per imitare una percezione più nitida nella via visiva (il che significa la capacità di rilevare la posizione spaziale corretta di un segnale esterno con pochi gradi di incertezza), mentre σ^a è stato impostato maggiore di σ^v , secondo studi computazionali precedenti (Cuppini et al., 2014; Magosso et al., 2012).

La forza degli stimoli visivi e uditivi esterni (parametri I_v e I_a) è stata scelta per evocare una risposta nelle regioni di input nella parte alta della caratteristica statica sigmoideale, cioè prossima alla saturazione.

Componente laterale intra-area $l_i^r(t)$

$L_{ex0}^r(t)$, $L_{in0}^r(t)$, $L_{ex}^r(t)$, σ_{in}^r sono stati assegnati in modo da consentire un equilibrio tra eccitazione e inibizione e quindi evitare instabilità, cioè un'eccitazione incontrollata che si propaga all'intera area, e in modo da permettere una forte competizione tra due ingressi della stessa modalità sensoriale.

Al contrario, l'ampiezza delle sinapsi inibitorie all'interno dell'area multisensoriale è minore, per permettere la discriminazione di eventi non correlati.

Componente inibitoria $Li_i^r(t)$

Per simulare un effetto inibitorio inter-sensoriale più forte a ISI vicini a 1000 ms (Crosse et al., 2022), abbiamo scelto una dinamica lenta per i meccanismi inibitori, implementata tramite costanti di tempo per le proiezioni di feedback uguali a 180 ms ($\tau_{in}^a = \tau_{in}^v$).

Componente cross-modale $c_j^r(t)$

La dinamica del termine cross-modale è simmetrica tra le modalità sensoriali. La loro costante di tempo ($\tau = 3$ ms) e il ritardo nelle sinapsi cross-modali, $\Delta t = 16$ ms, sono stati selezionati in modo che la componente cross-modale produca un effetto sull'altra regione dopo 30–40 ms. Questi parametri permettono di produrre un rapido effetto eccitatorio transitorio tra le aree di input, come numerose evidenze in letteratura suggeriscono che avvenga già nelle corteccie primarie (vedi Foxe & Schroeder, 2005; Musacchia & Schroeder, 2009; Recanzone, 2009; Stein & Stanford, 2008 per rassegne).

Componente eccitatoria interneurale $I_{ex}^r(t)$

WI è stato scelto in modo che anche una piccola attività nello strato di input possa attivare il meccanismo inibitorio inter-sensoriale, stimolando l'unità interneurale corrispondente. La costante di tempo $\tau_{lex}^r = 15$ ms è la stessa dei modelli precedenti (Cuppini et al., 2014; Cuppini et al., 2020).

Input inibitorio reciproco interneurale $I_{in}^r(t)$

L'efficacia delle sinapsi reciproche è stata scelta in modo che l'unità interneurale vincente sia in grado di spegnere quasi completamente l'elemento concorrente. La

costante di tempo è rimasta la stessa dei modelli precedenti (Cuppini et al., 2014; Cuppini et al., 2020).

Input eccitatorio all'area multisensoriale $ex^m(t)$

W_m è stato scelto in modo che anche una piccola attività di input, pari ad esempio al 30% del suo valore di saturazione, susciti una risposta nell'area di output poco al di sotto della saturazione. Il ritardo $\Delta t^m (= 50 \text{ ms})$, insieme al ritardo di attivazione $\Delta t^{pm} (= \Delta t^m)$ dell'area premotoria, è stato assegnato in accordo con la soglia assunta da Crosse et al., 2022 per rilevare anomalie veloci (tempi di reazione più rapidi, inferiori a 100 ms, sono considerati risposte anticipatorie). Questi elementi sinaptici sono stati inclusi per simulare il processo neurale che sottende la produzione di una risposta motoria nell'area premotoria.

I valori G_i^r sono stati regolati in modo che l'attività evocata nelle unità postsinaptiche rimanga nella porzione lineare della relazione sigmoideale.

2. Risultati

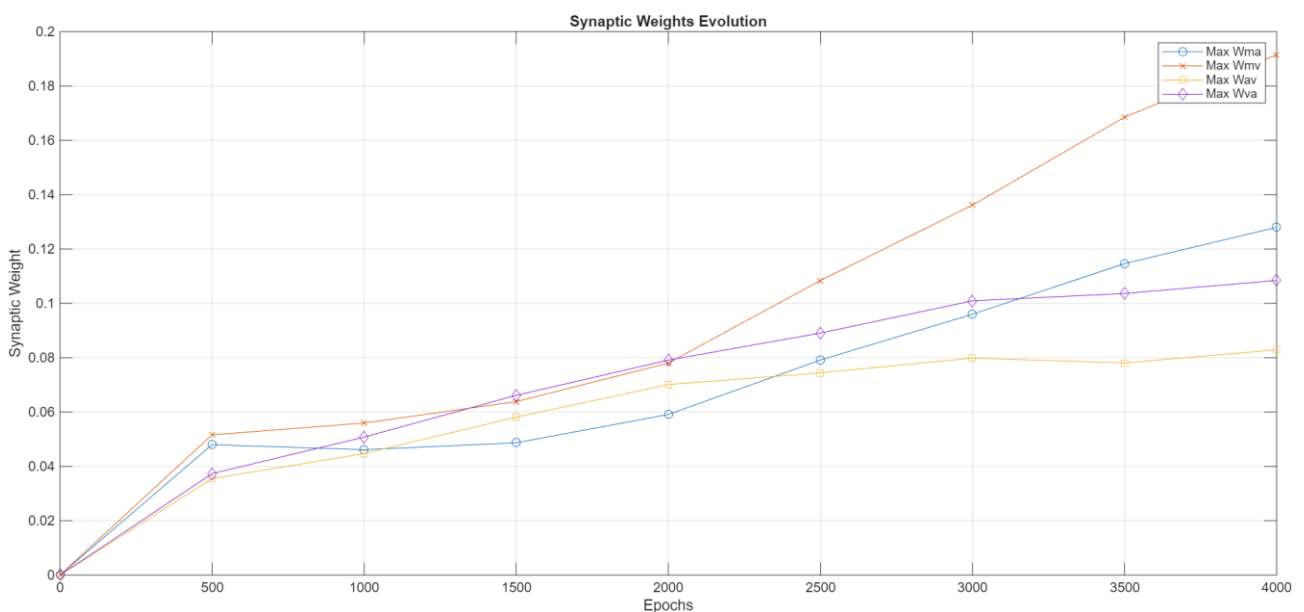
2.1 Influenza della modulazione di α_M e α sulla dinamica dei pesi sinaptici

Nei grafici che seguono è rappresentata l'evoluzione dei pesi sinaptici massimi (feedforward e cross-modal) durante le 4000 epoche di addestramento, con campionamento ogni 1000 epoche. In entrambi i casi il filtraggio tramite soglia è applicato alle connessioni feedforward, mentre per le connessioni cross-modal esso è introdotto soltanto nella seconda condizione.

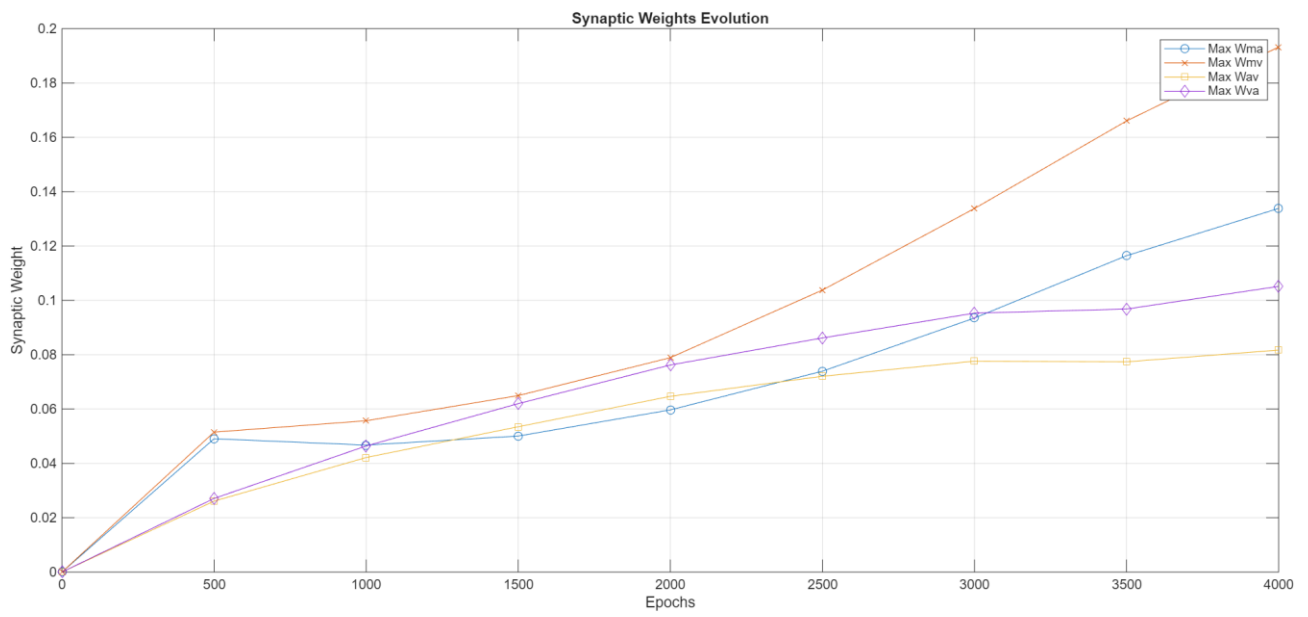
I risultati riportati sono stati ottenuti dopo aver modificato i parametri di velocità α e α_M , in modo da regolare più accuratamente la dinamica di crescita delle diverse tipologie di connessioni. Questa scelta si è rivelata fondamentale per osservare con maggiore chiarezza le differenze tra le due condizioni sperimentali e per distinguere il contributo specifico del filtraggio al processo di apprendimento.

L'analisi mette in evidenza un aspetto particolarmente rilevante legato all'andamento delle connessioni feedforward. In assenza di filtraggio, infatti, la loro crescita si manifesta in una fase successiva rispetto alle connessioni cross-modal, con un'evoluzione meno regolare e caratterizzata da una maggiore variabilità. L'introduzione del filtraggio modifica in maniera significativa questa dinamica: le connessioni feedforward tendono a stabilizzarsi prima delle connessioni cross-modal, mostrando un profilo più ordinato e robusto.

Feedforwad con soglia, cross-modal senza soglia



Feedforwad con soglia, cross-modali con soglia

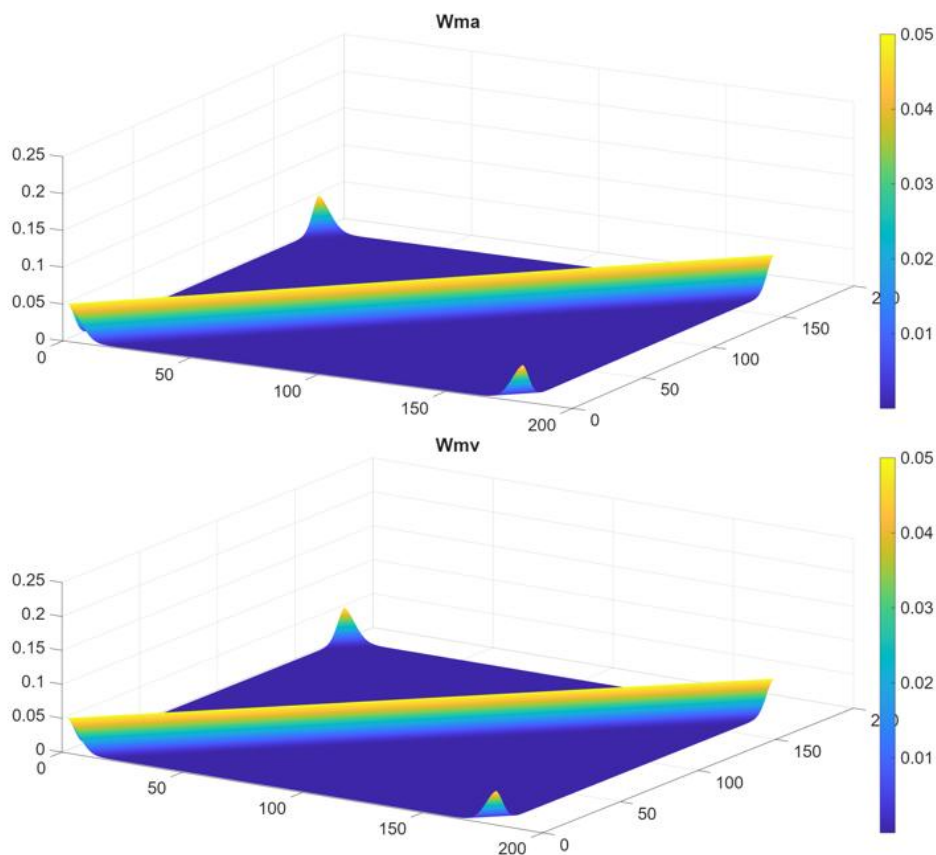


2.2 Evoluzione dei pesi sinaptici con stimoli singoli e filtraggio tramite soglia

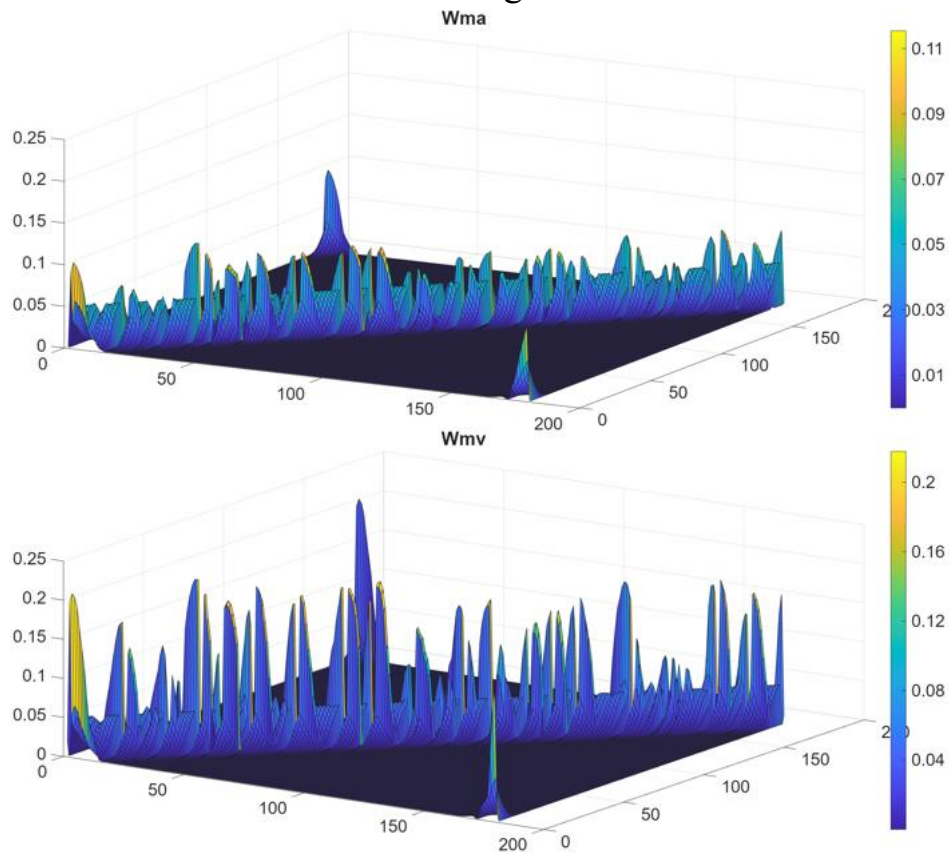
Per ottenere tali risultati è stato necessario ridurre il numero delle epoche di addestramento a 500, con un passo temporale di 125. Questa scelta è stata dettata dall'esigenza di rendere più leggibile e interpretabile l'evoluzione dei pesi sinaptici, evitando di sovraccaricare i grafici con un numero eccessivo di campioni e riducendo al contempo i costi computazionali.

I grafici riportano l'evoluzione dell'efficacia sinaptica delle connessioni feedforward, cross-modali e inibitorie, confrontando le condizioni con e senza l'applicazione del filtraggio tramite soglia. Le connessioni feedforward, per le quali il filtraggio è sempre presente, mostrano una crescita rapida e regolare che porta a una stabilizzazione precoce. Al contrario, le connessioni cross-modali e inibitorie evidenziano un incremento più graduale, con valori massimi di efficacia superiori nel caso delle cross-modali e molto più contenuti nel caso delle inibitorie. In entrambe queste tipologie di connessioni, l'introduzione del filtraggio riduce la variabilità e rende la dinamica più uniforme, pur mantenendo le differenze nei livelli finali di efficacia.

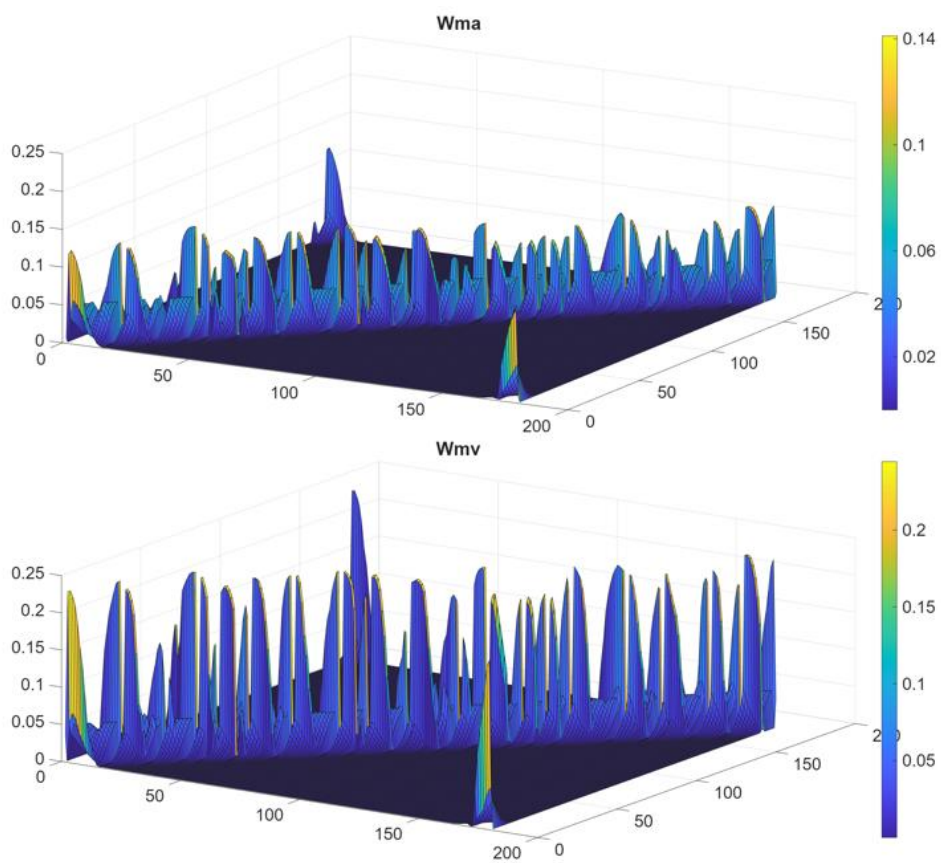
Feedforward con soglia: EPOCA 0



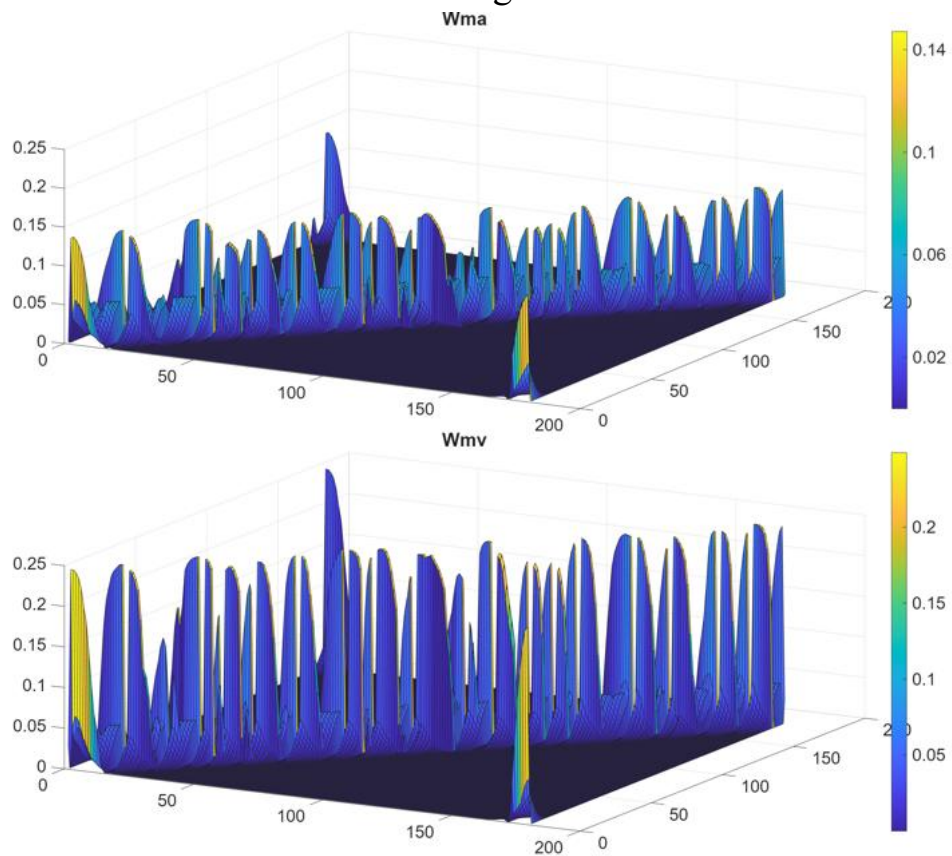
Feedforward con soglia: EPOCA 125



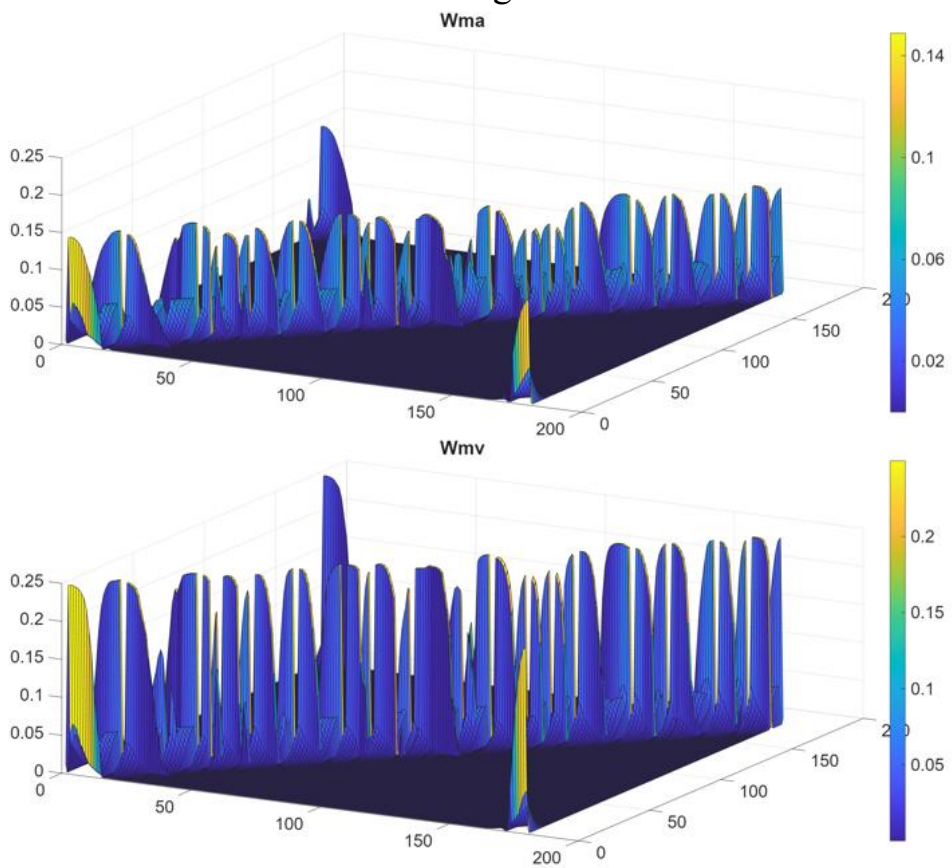
Feedforward con soglia: EPOCA 250



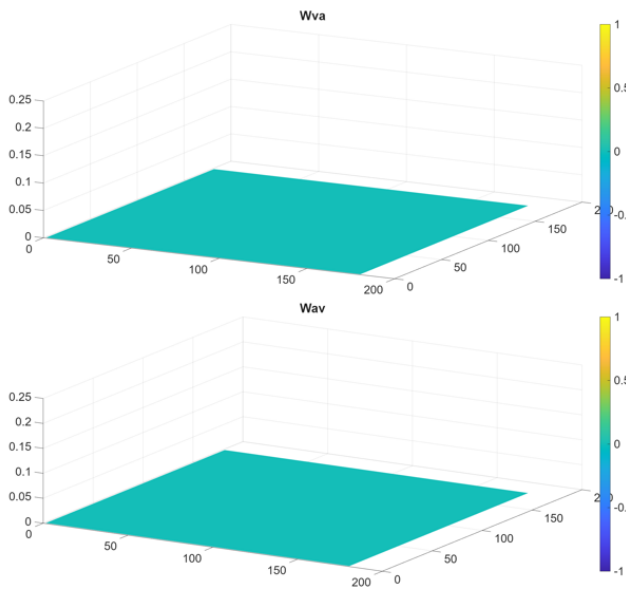
Feedforward con soglia: EPOCA 375



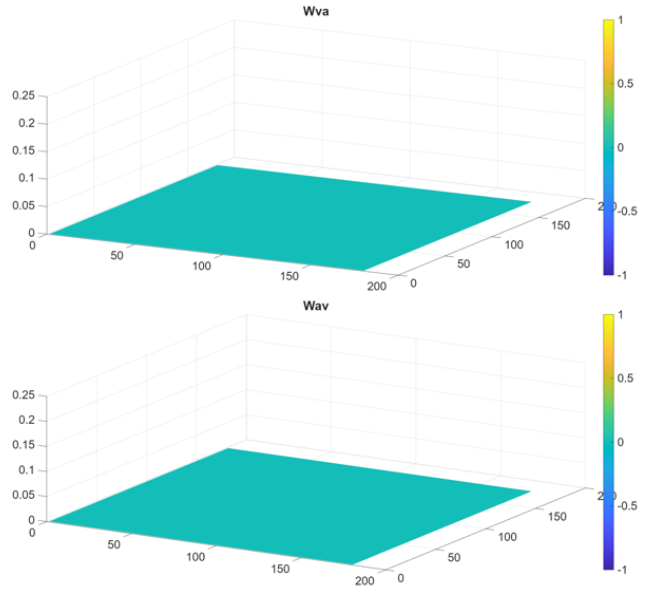
Feedforward con soglia: EPOCA 500



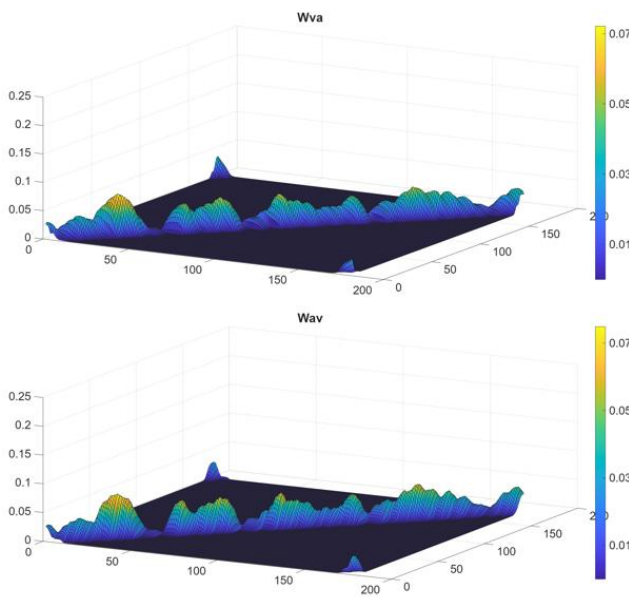
Cross-modalità senza soglia: EPOCA 0



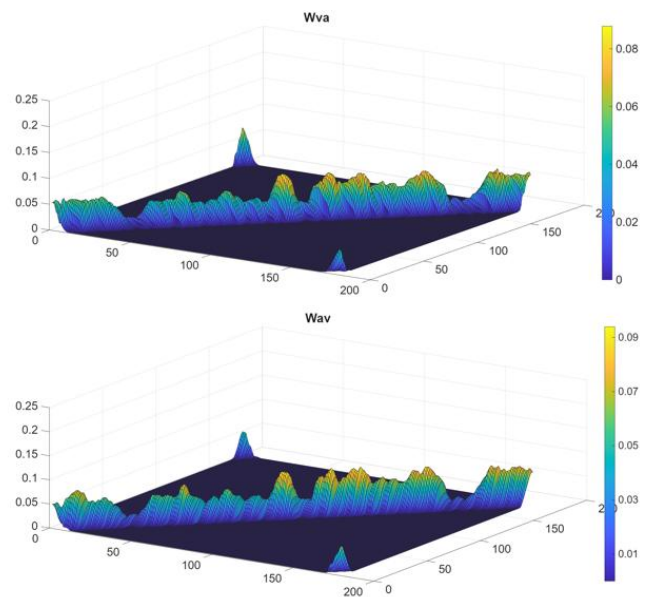
Cross-modalità con soglia: EPOCA 0



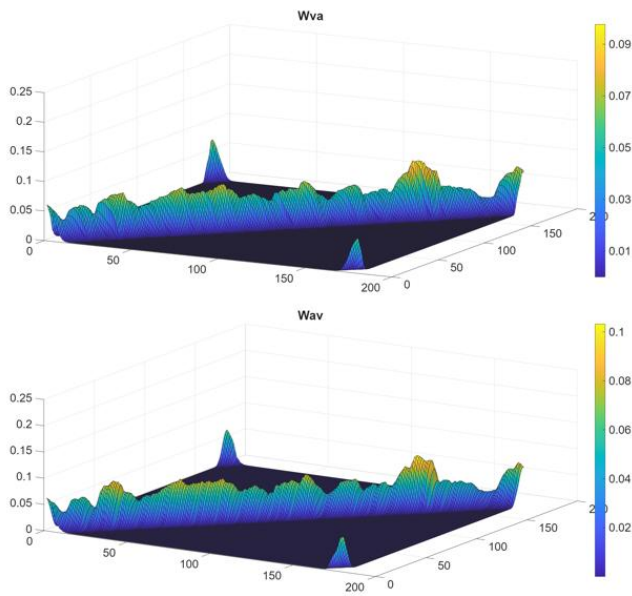
Cross-modalità senza soglia: EPOCA 125



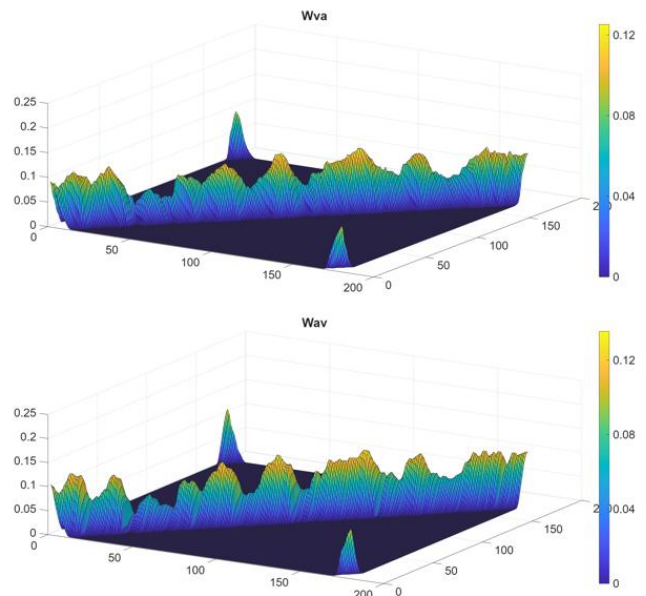
Cross-modalità con soglia: EPOCA 125



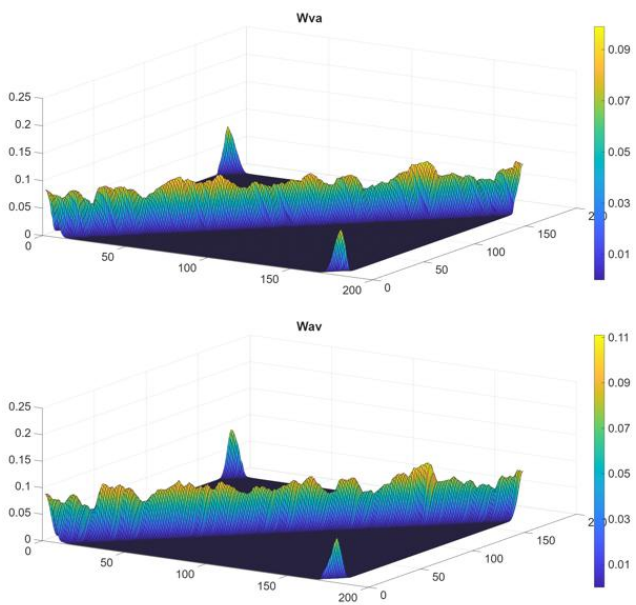
Cross-modalità senza soglia: EPOCA 250



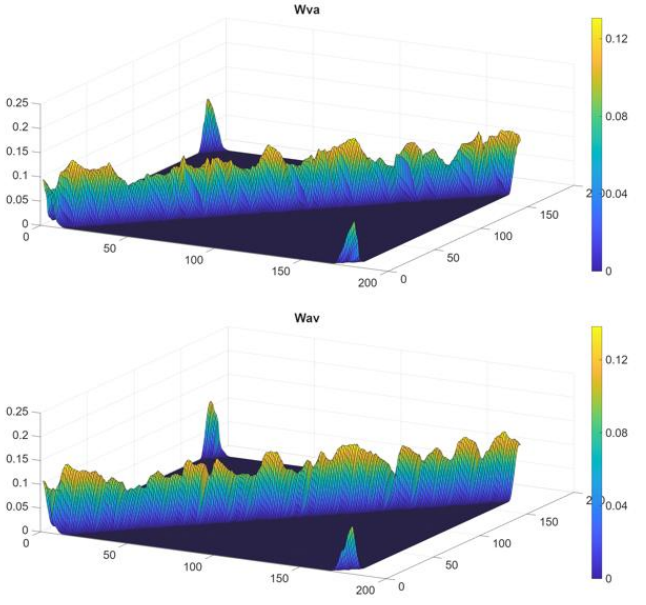
Cross-modalità con soglia: EPOCA 250



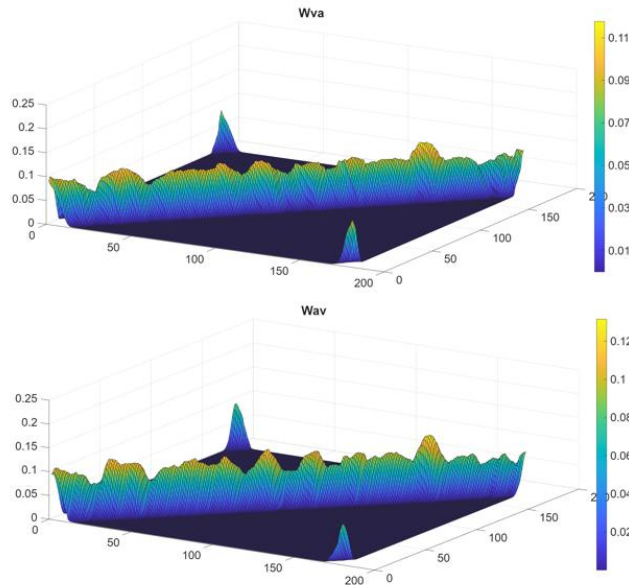
Cross-modalità senza soglia: EPOCA 375



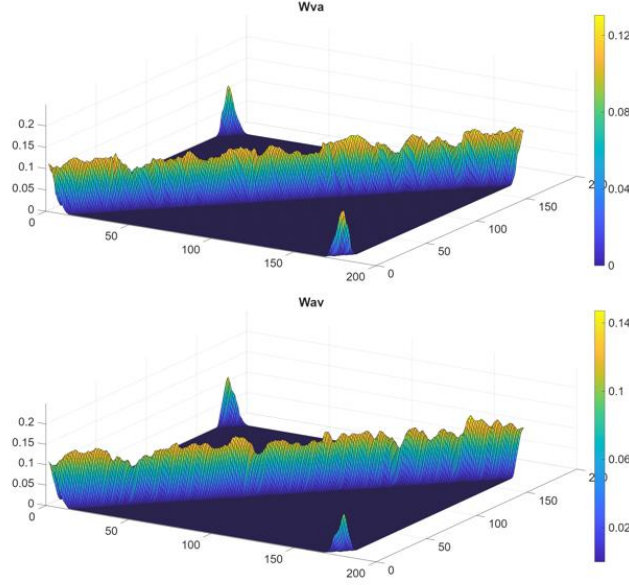
Cross-modalità con soglia: EPOCA 375



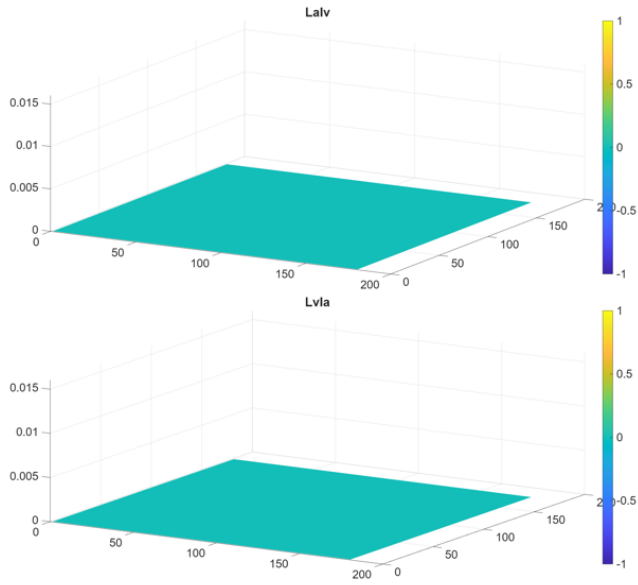
Cross-modali senza soglia: EPOCA 500



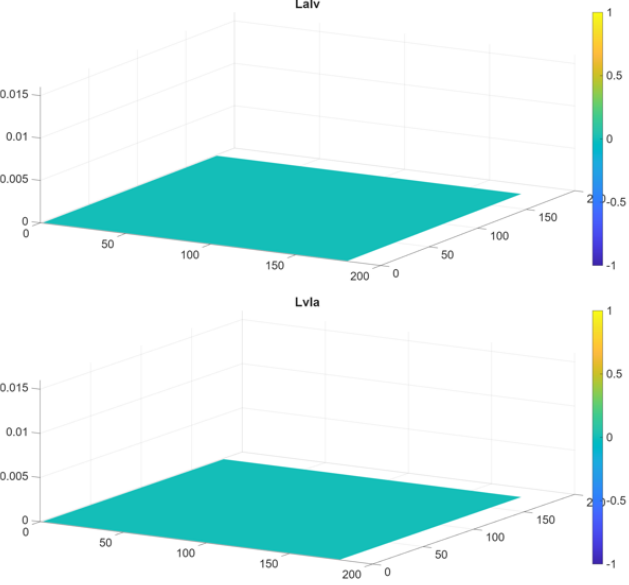
Cross-modali con soglia: EPOCA 500



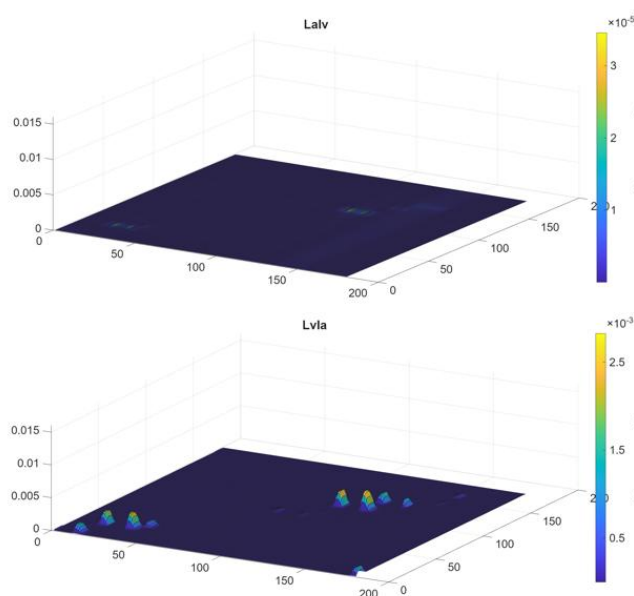
Inibitorie senza soglia: EPOCA 0



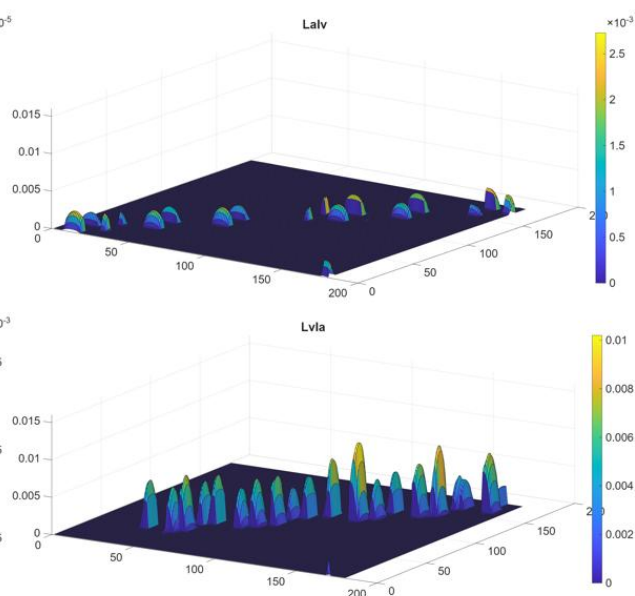
Inibitorie con soglia: EPOCA 0



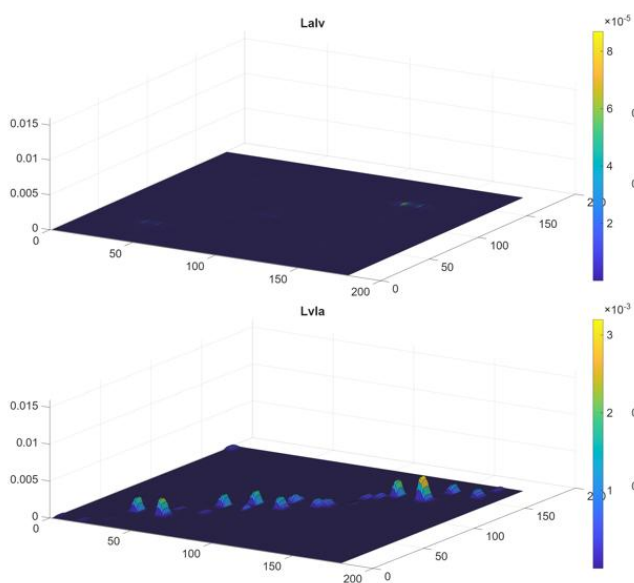
Inibitorie senza soglia: EPOCA 125



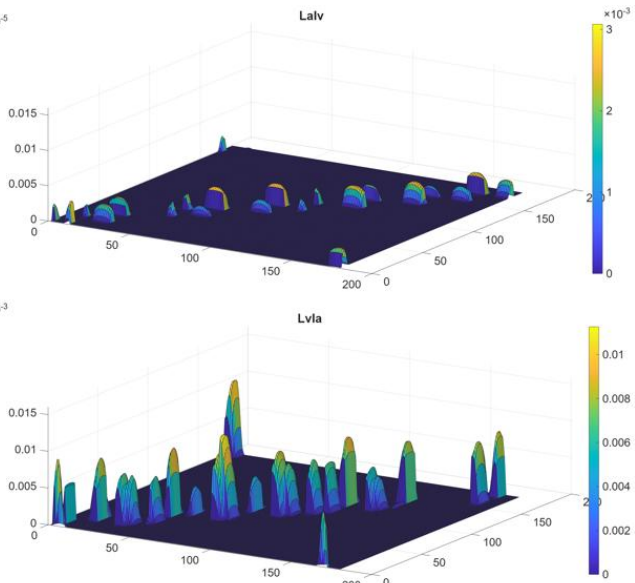
Inibitorie con soglia: EPOCA 125



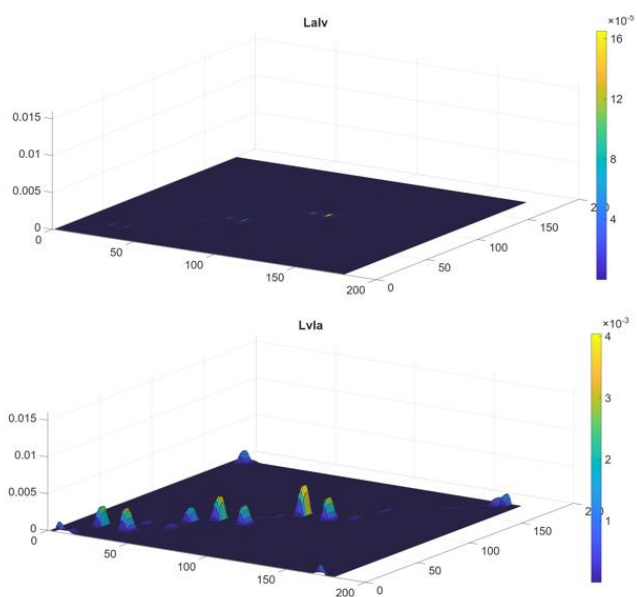
Inibitorie senza soglia: EPOCA 250



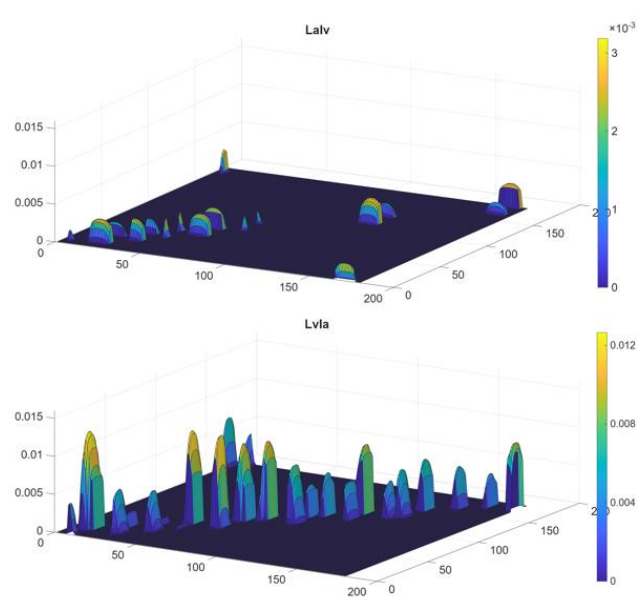
Inibitorie con soglia: EPOCA 250



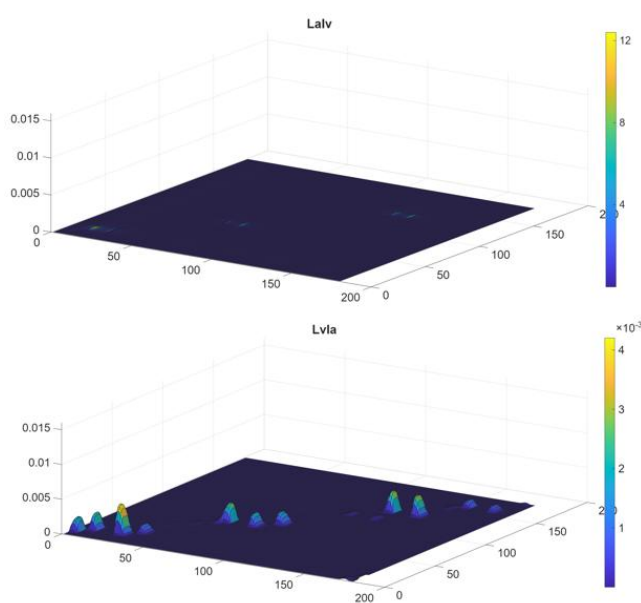
Inibitorie senza soglia: EPOCA 375



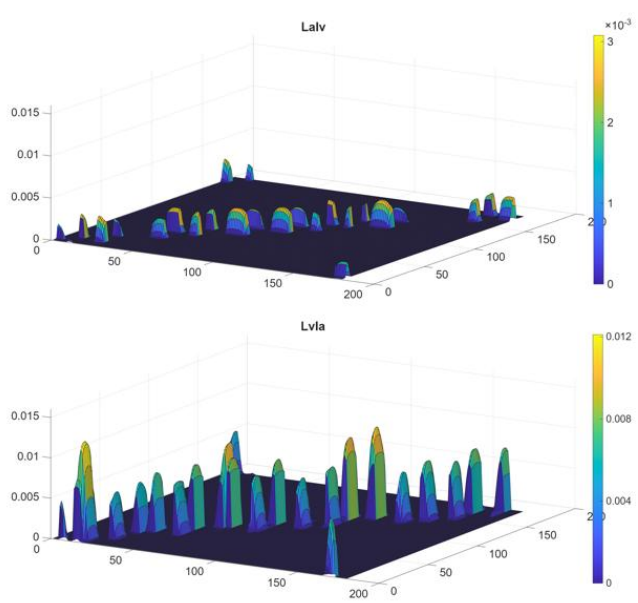
Inibitorie con soglia: EPOCA 375



Inibitorie senza soglia: EPOCA 500



Inibitorie con soglia: EPOCA 500



2.3 Evoluzione dell'attività nelle aree del modello neurale

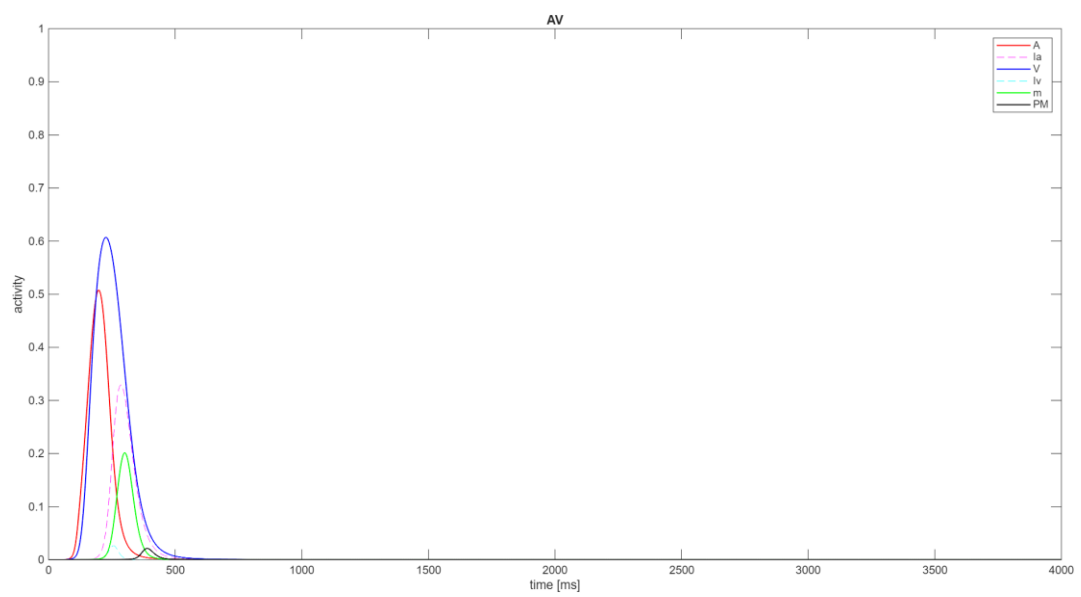
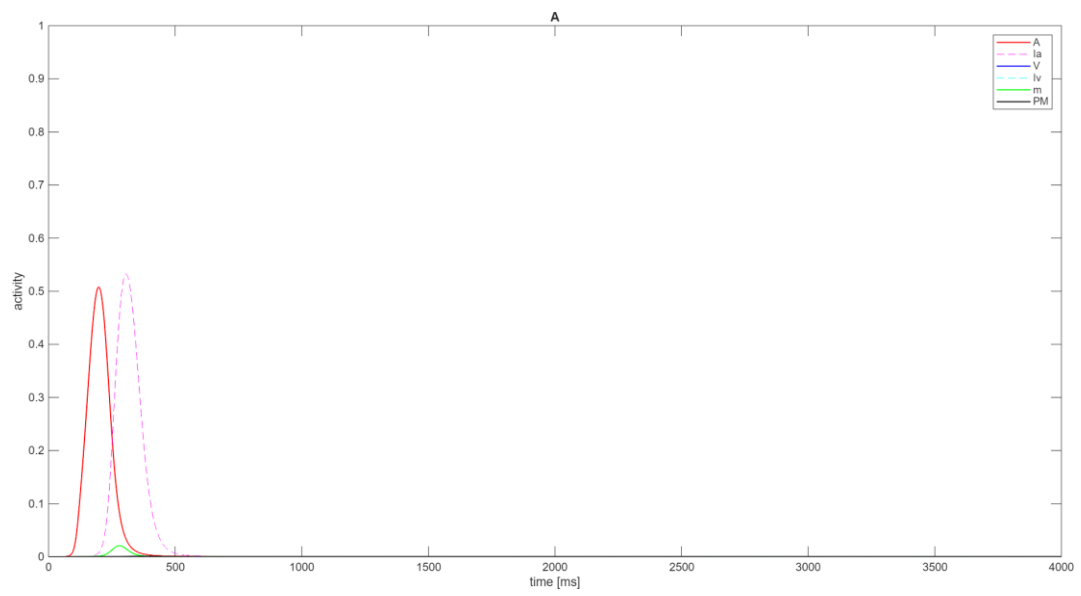
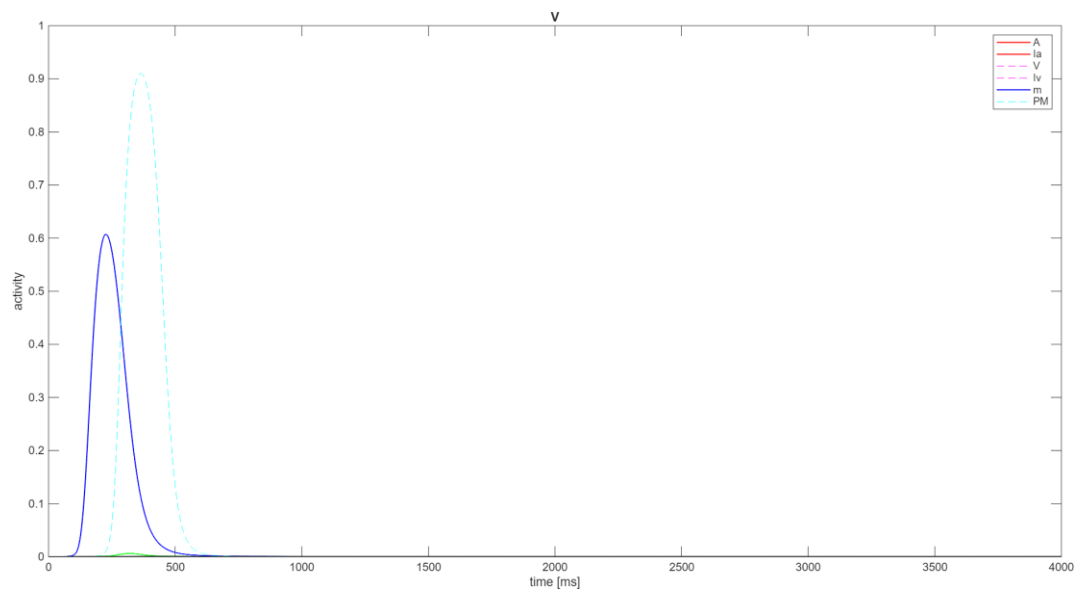
Le figure sotto riportate illustrano l'evoluzione dell'attività nelle aree visiva (V), uditiva (A), multisensoriale (m) e premotoria (PM) durante il processo di addestramento, valutata in tre condizioni sperimentali: stimolazione visiva singola, stimolazione uditiva singola e stimolazione bimodale, nel periodo compreso tra l'epoca 0 e l'epoca 4000.

All'epoca 0, l'attività è concentrata quasi esclusivamente nelle aree sensoriali unimodali corrispondenti allo stimolo presentato, con una risposta multisensoriale minima e un'assenza di attivazione premotoria. Con il progredire dell'addestramento, si osserva un graduale incremento dell'attività multisensoriale in tutte le condizioni. Tuttavia, nei casi di stimolo singolo (solo visivo o solo uditivo), questo aumento rimane modesto e non coinvolge in modo significativo l'area premotoria.

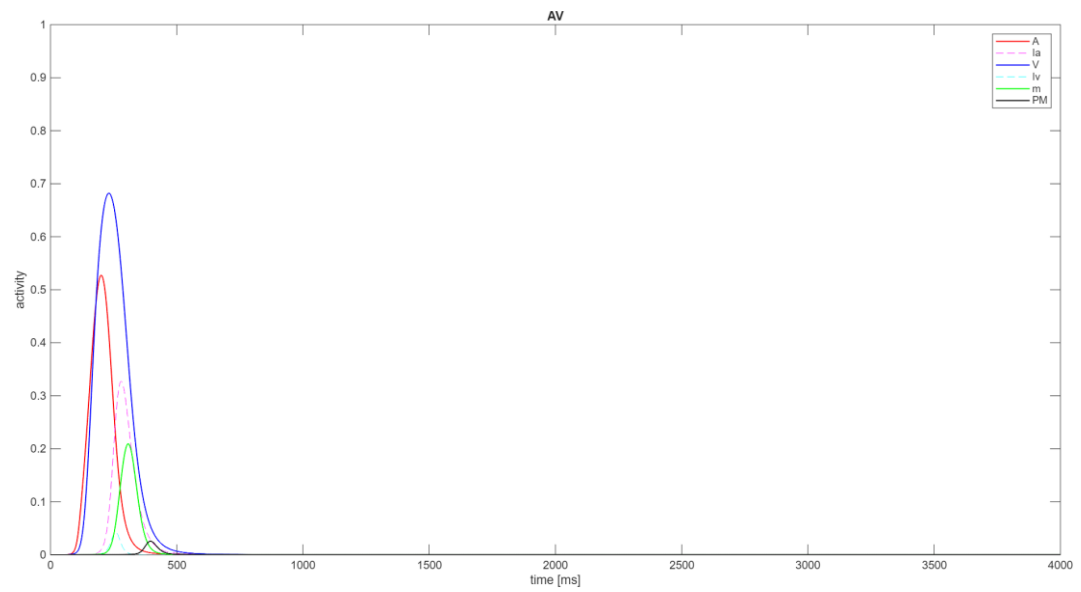
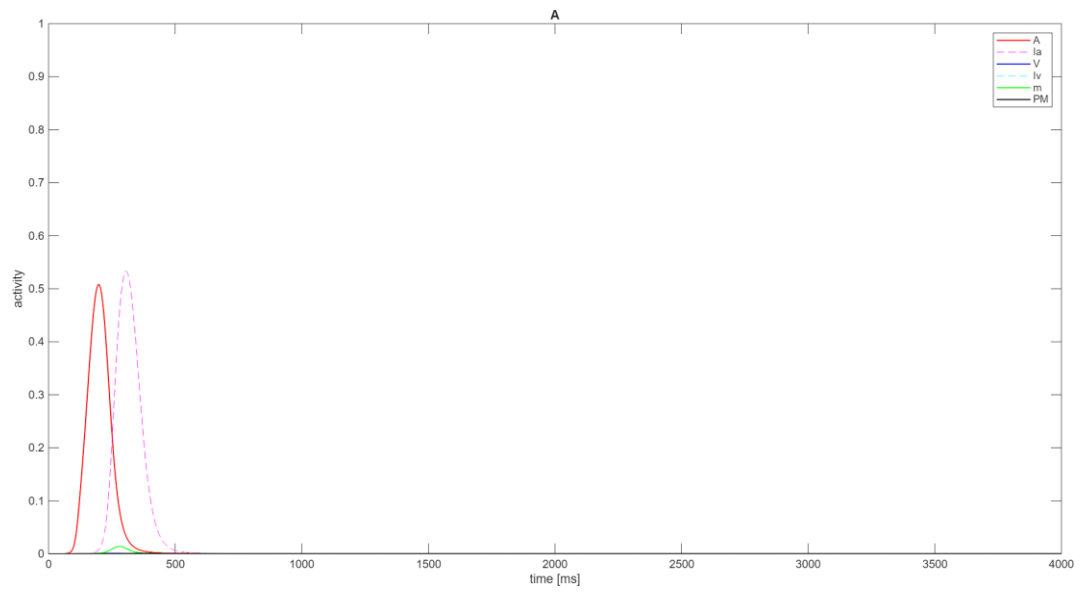
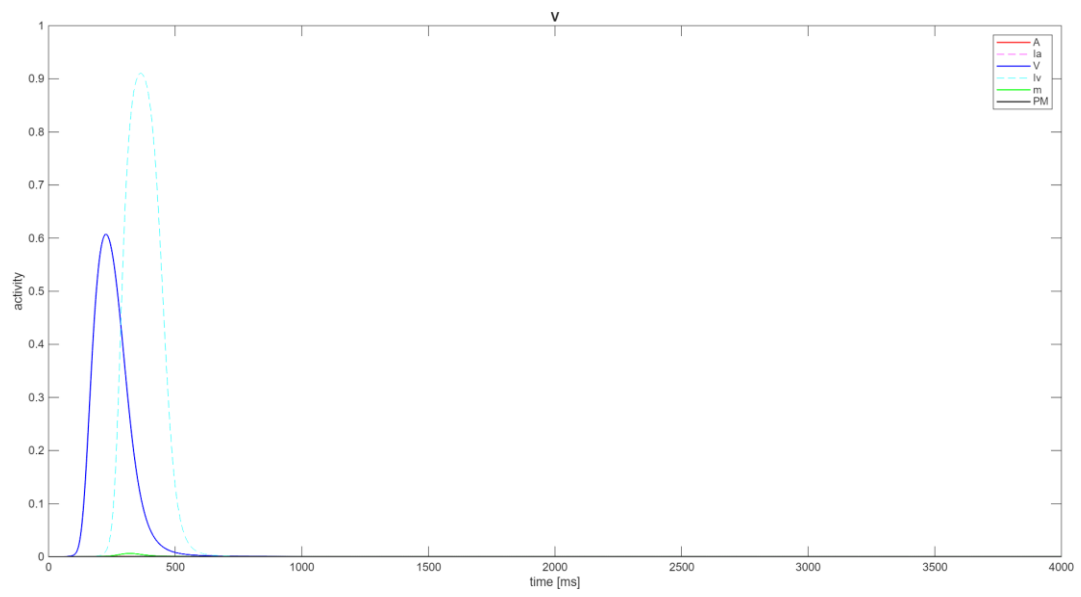
Nella condizione bimodale, invece, l'attività multisensoriale cresce in maniera marcata già dalle epoche intermedie (circa 2000–3000), accompagnata da una progressiva attivazione dell'area premotoria, che raggiunge i valori più elevati all'epoca 4000.

Questo andamento riflette il rafforzamento graduale delle connessioni sinaptiche e la crescente capacità del modello di integrare le informazioni provenienti da entrambe le modalità sensoriali, favorendo la trasmissione verso aree deputate alla pianificazione motoria.

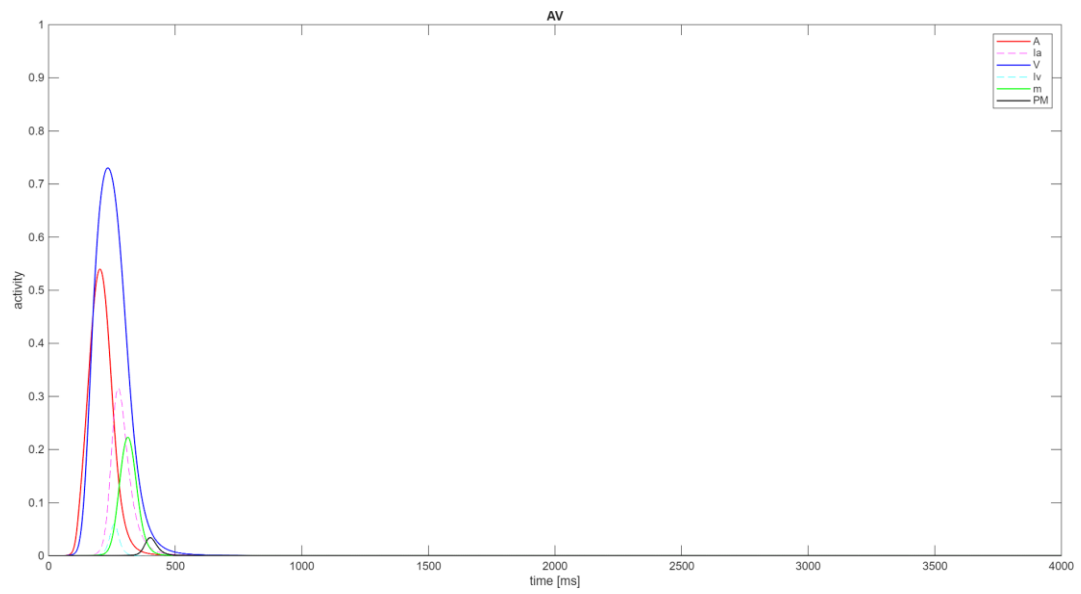
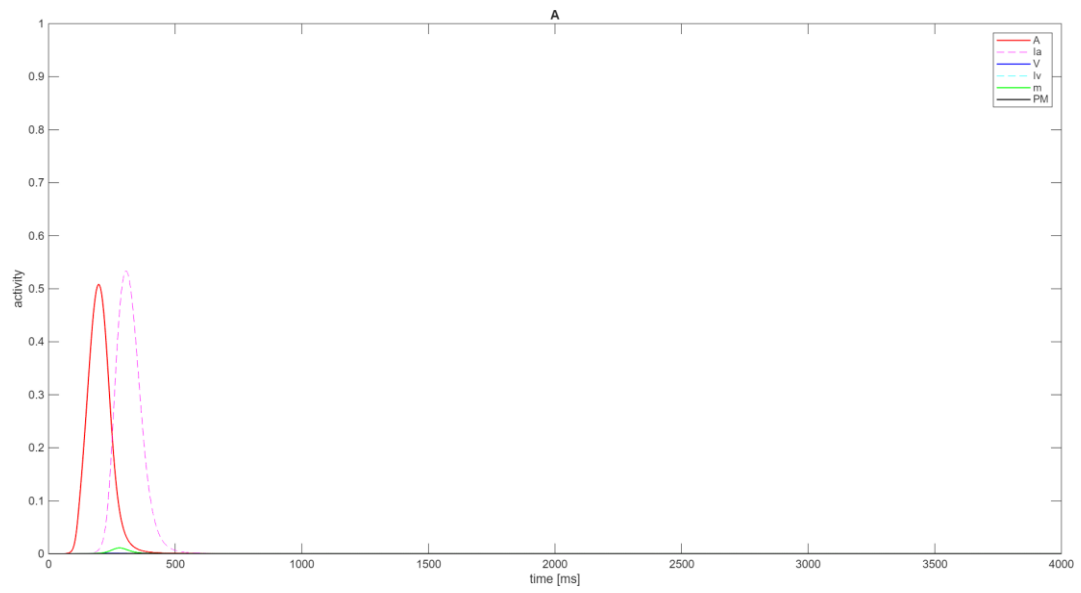
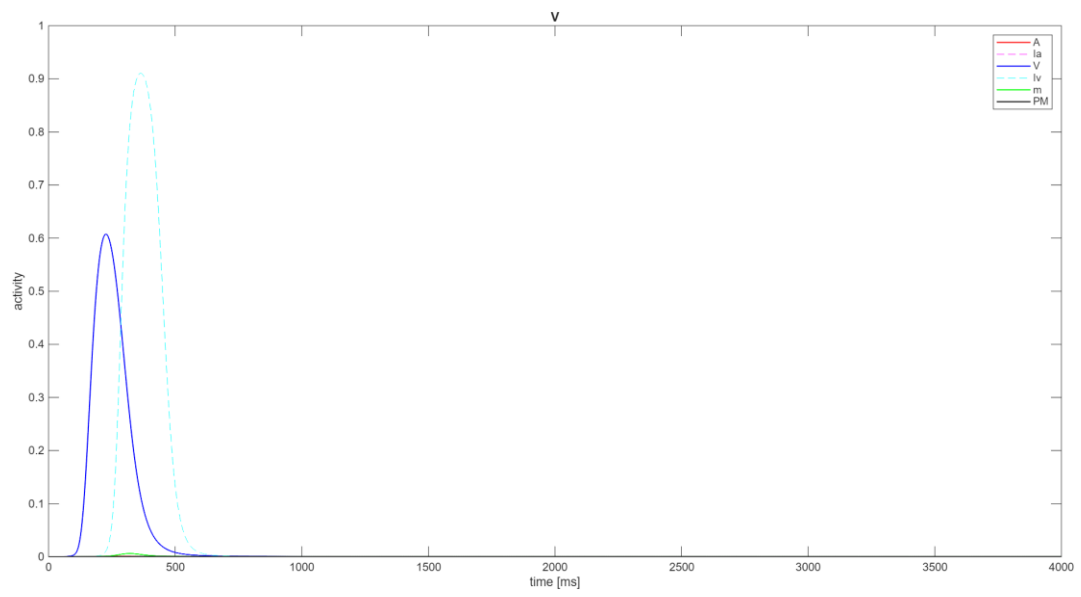
EPOCA 0



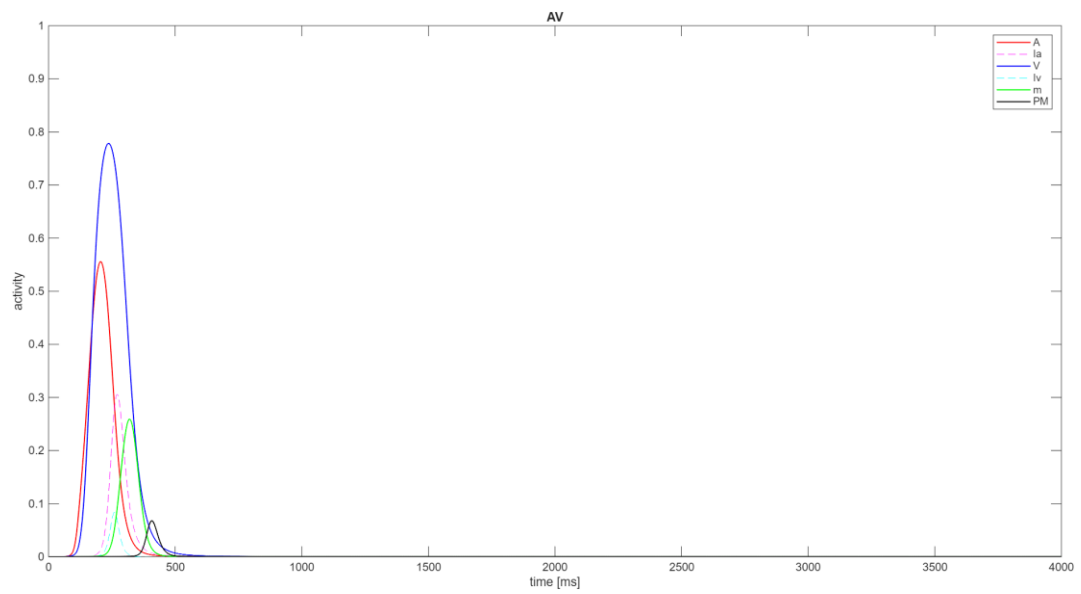
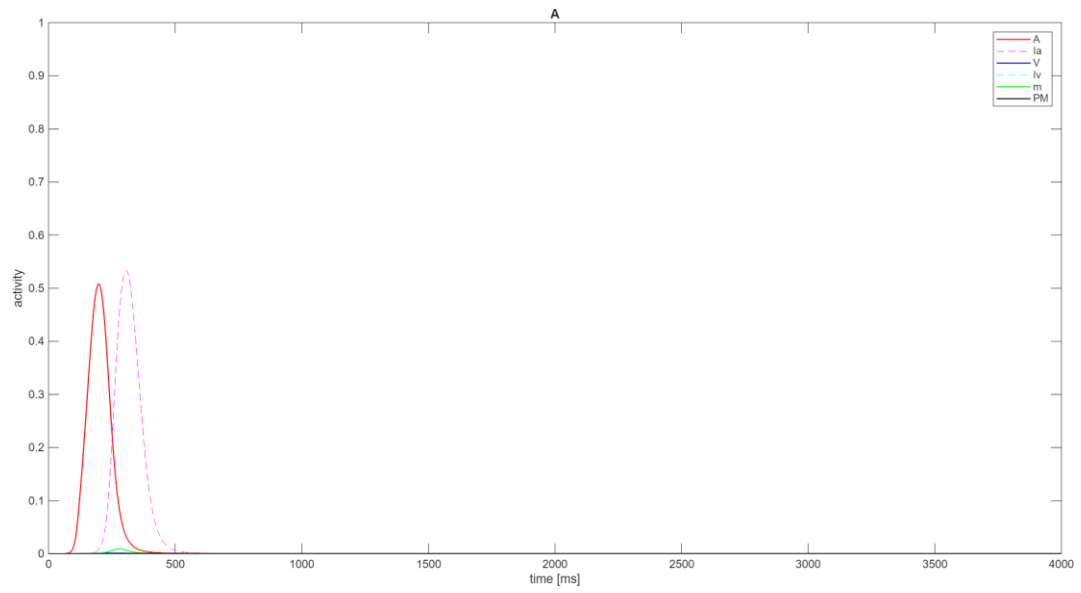
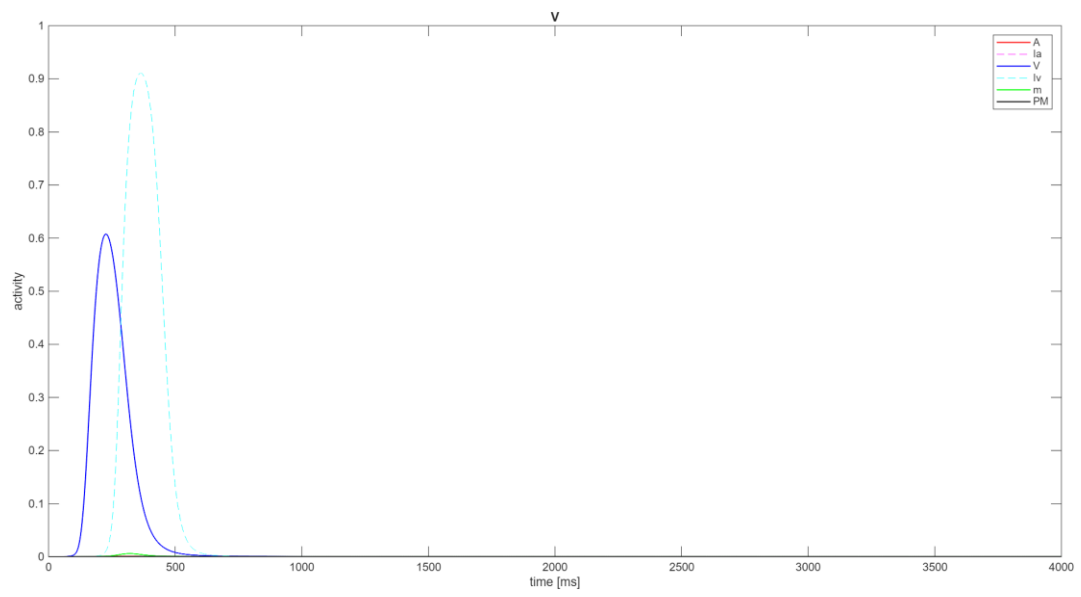
EPOCA 1000



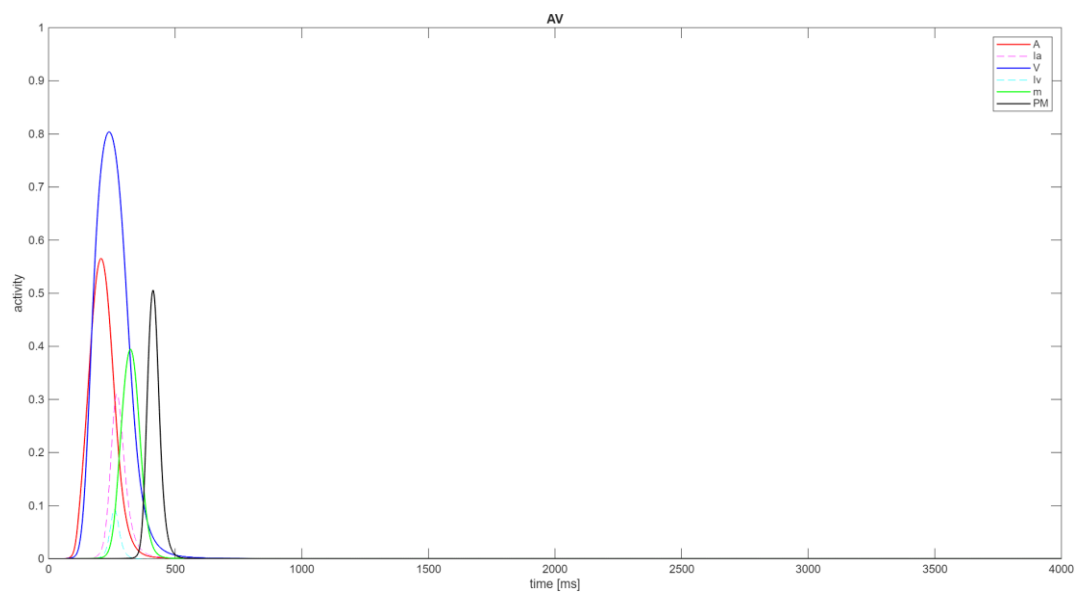
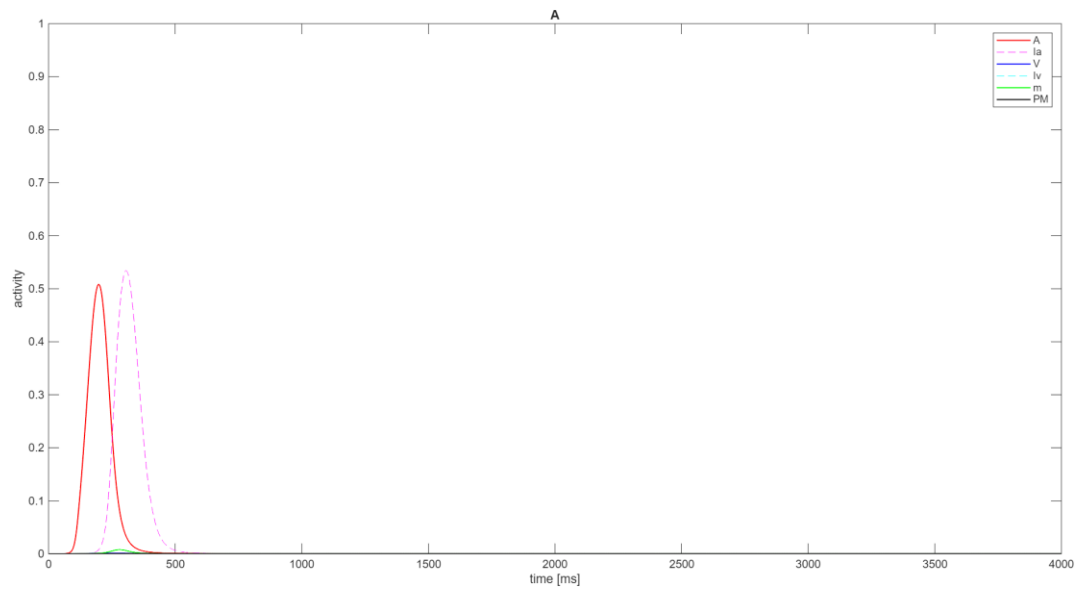
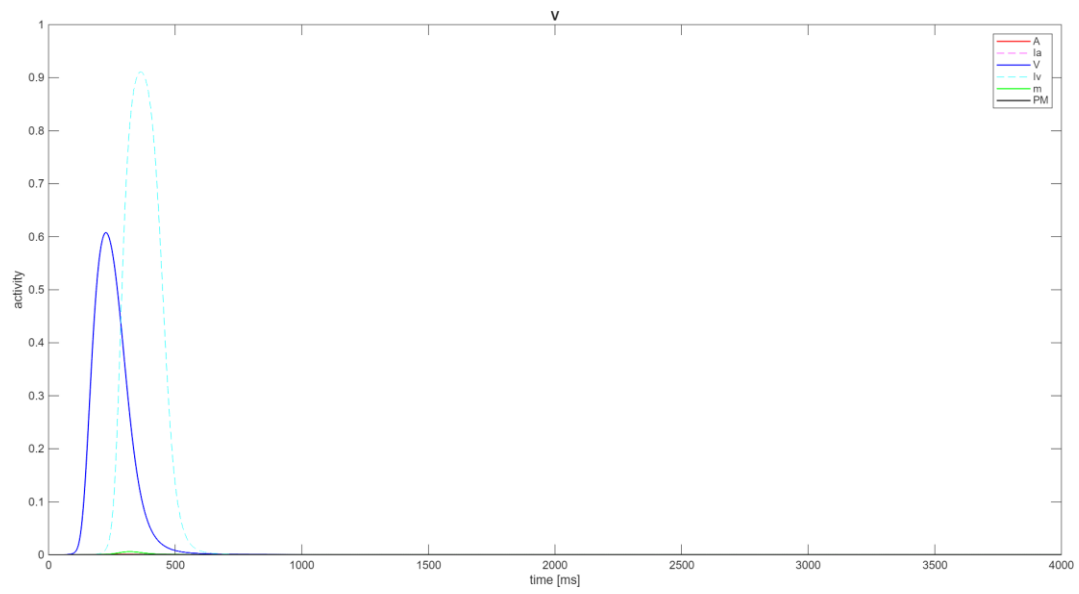
EPOCA 2000



EPOCA 3000



EPOCA 4000



2.4 Dinamica dei pesi sinaptici con stimoli sequenziali

I grafici seguenti mostrano l'evoluzione delle sinapsi, tutte sottoposte a filtraggio tramite soglia, durante un addestramento con stimoli sequenziali. L'analisi è stata condotta limitando l'addestramento a 500 epoche, con un campionamento ogni 125 epoche.

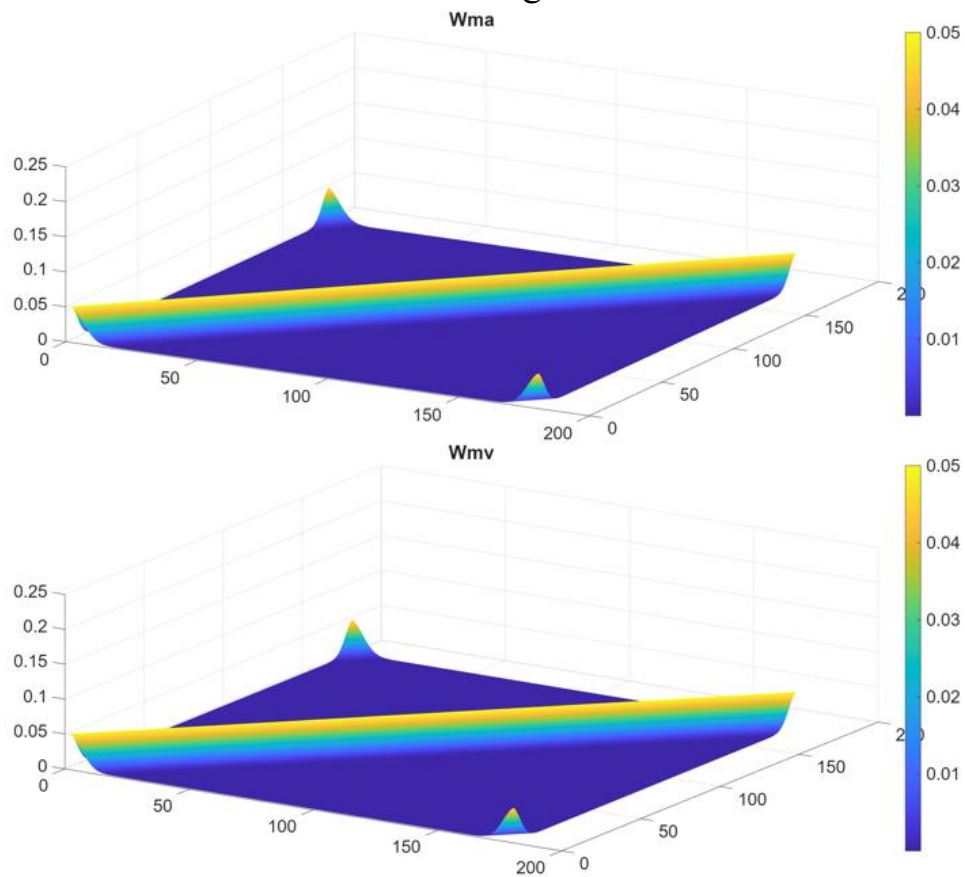
Nel caso delle sinapsi feedforward, i picchi di attività si consolidano progressivamente. Già a 125 epoche si notano dei picchi iniziali, che diventano più definiti e consistenti a 250 e 375 epoche, fino a raggiungere la loro massima espressione a 500 epoche. Questo andamento suggerisce un apprendimento rapido e mirato.

Le sinapsi cross-modali mostrano un apprendimento simile. Si osserva la formazione di picchi specifici che si intensificano con l'aumento delle epoche. L'evoluzione è chiara: da una superficie iniziale piatta, si passa a una mappa di connettività con picchi sempre più evidenti e localizzati. Anche in questo caso, il processo di apprendimento sembra efficiente e si conclude in un arco di tempo ridotto.

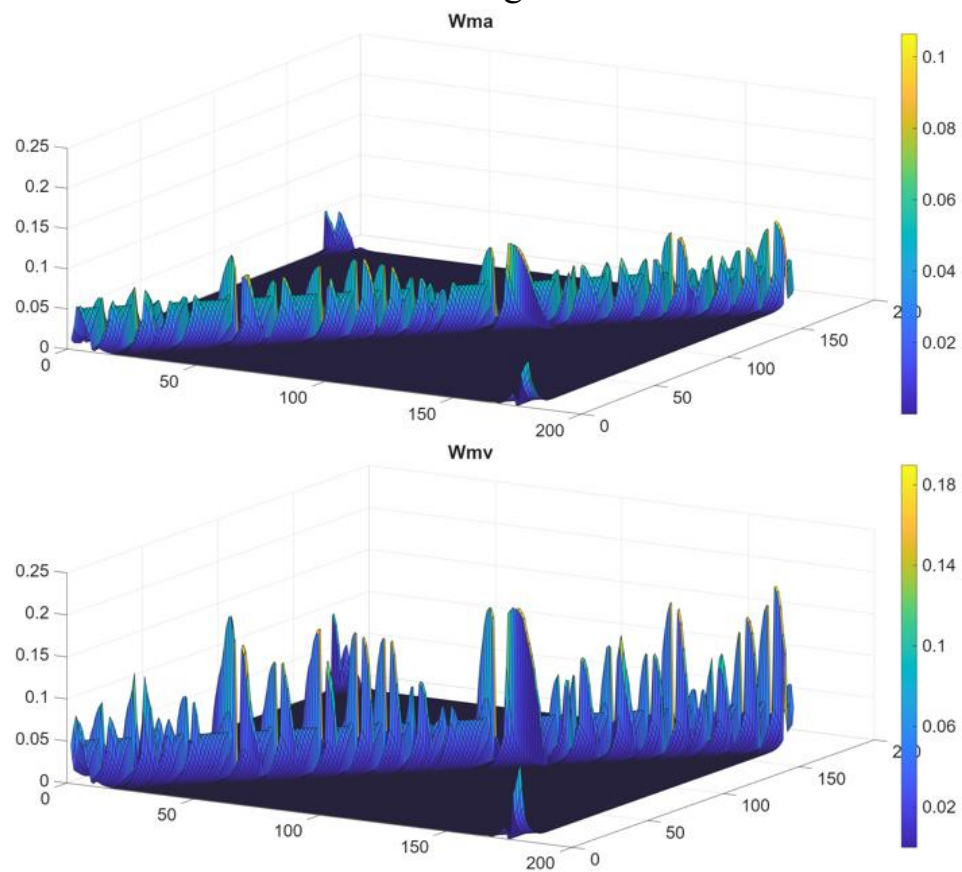
Per le sinapsi inibitorie, l'andamento è più graduale: l'apprendimento è meno marcato nelle prime epoche e i picchi si formano più lentamente. Si nota una maggiore attività dispersa e meno focalizzata rispetto alle sinapsi feedforward e cross-modali, anche se con il passare delle epoche i picchi tendono a diventare più definiti.

Nel complesso, l'addestramento con stimoli sequenziali dimostra che il modello è in grado di apprendere e di specializzare le sue connessioni, consentendo inoltre un addestramento biologicamente plausibile delle sinapsi $LvIa$ e $LaIv$, con un'efficacia variabile a seconda della tipologia considerata.

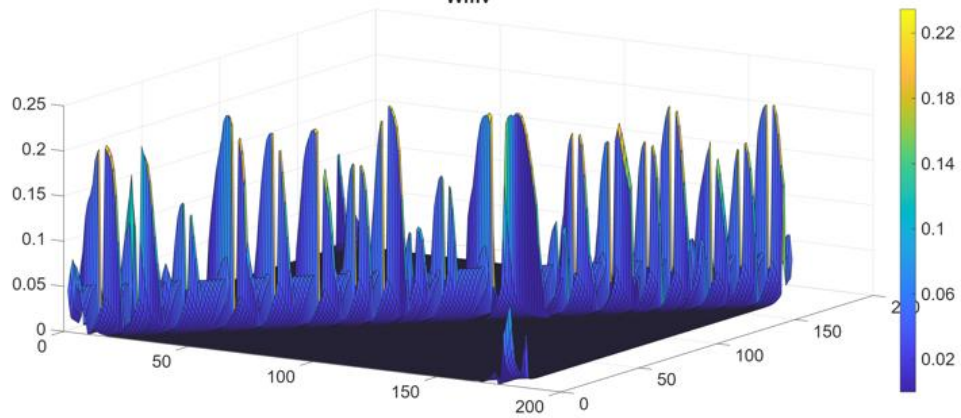
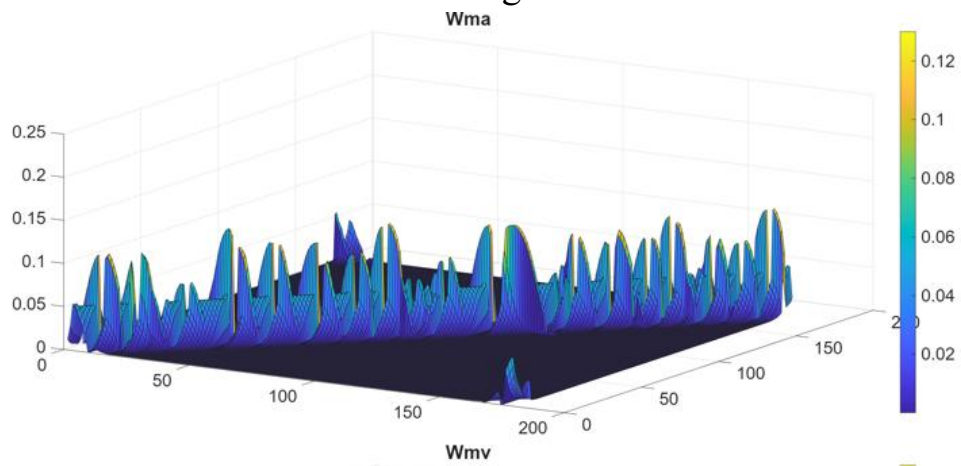
Feedforward con soglia: EPOCA 0



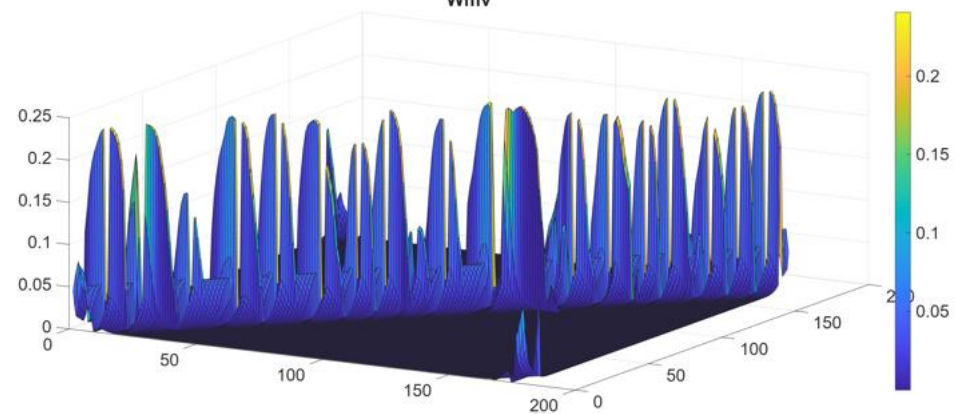
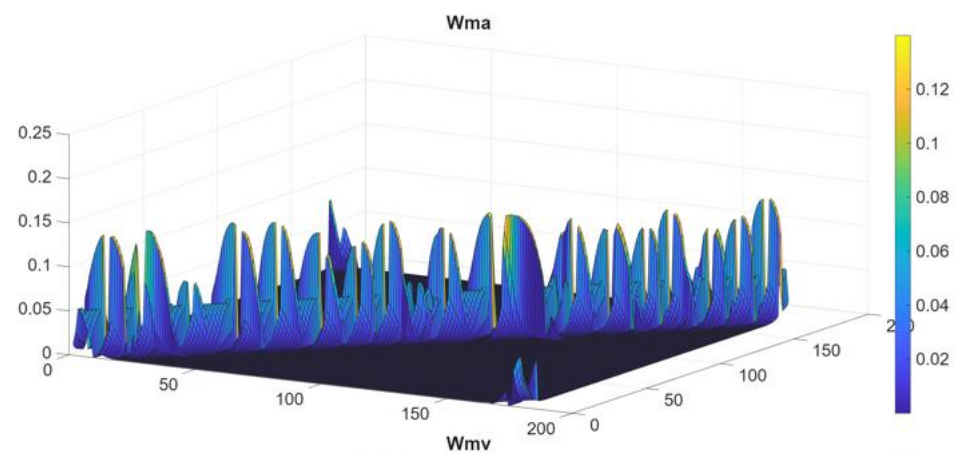
Feedforward con soglia: EPOCA 125



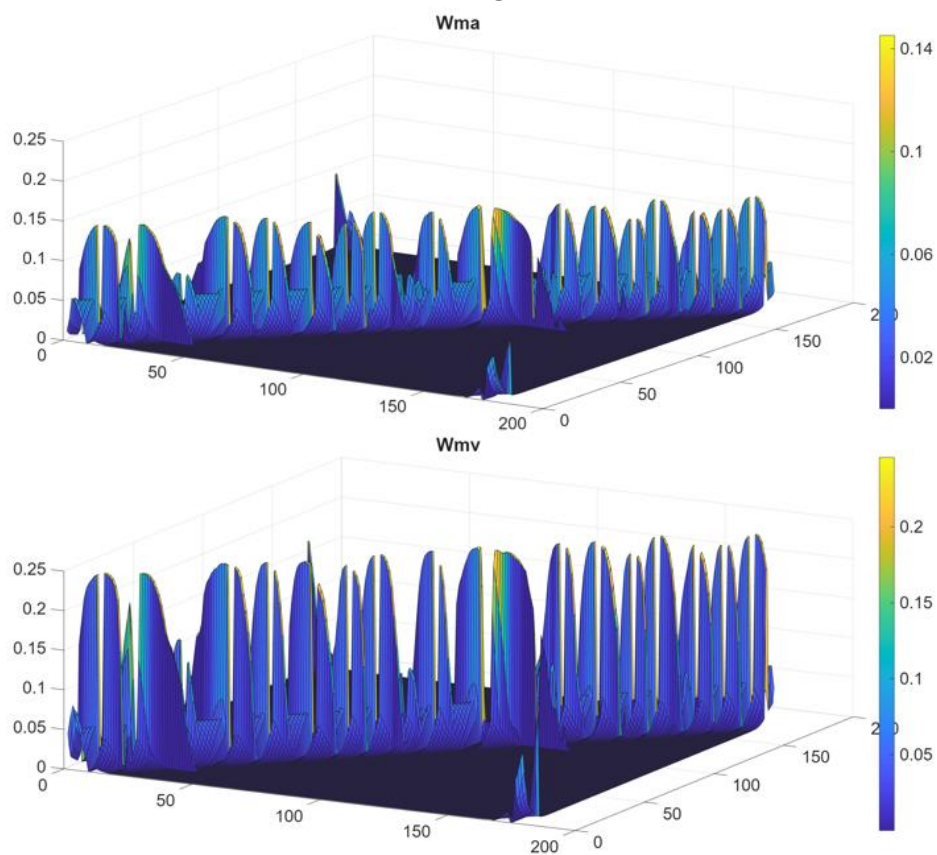
Feedforward con soglia: EPOCA 250



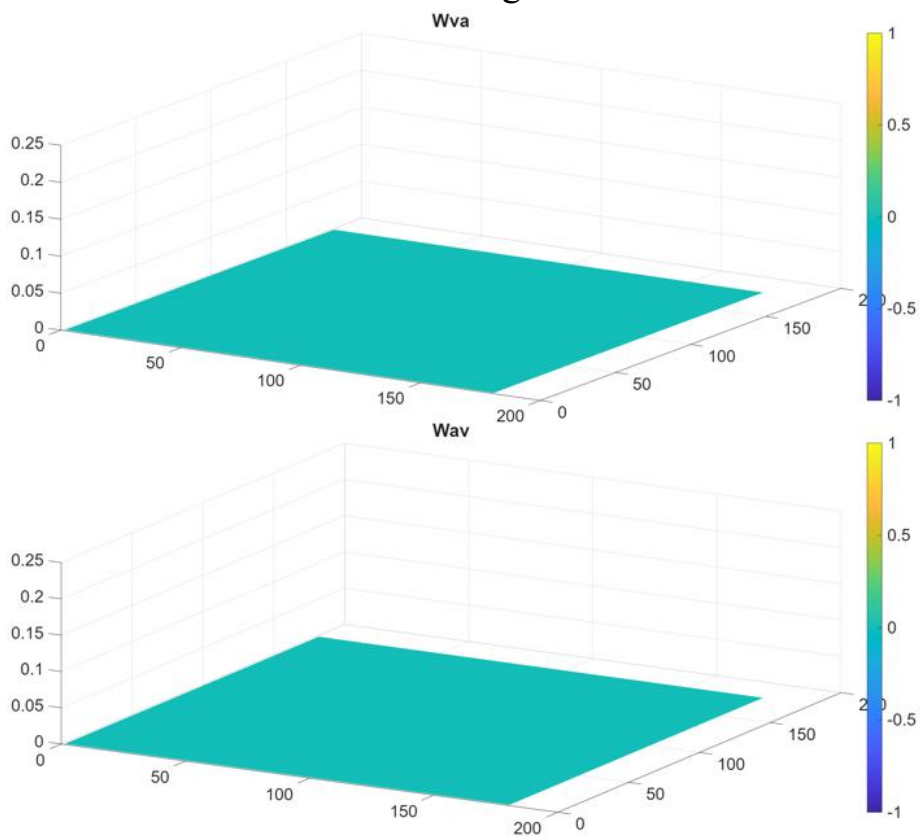
Feedforward con soglia: EPOCA 375



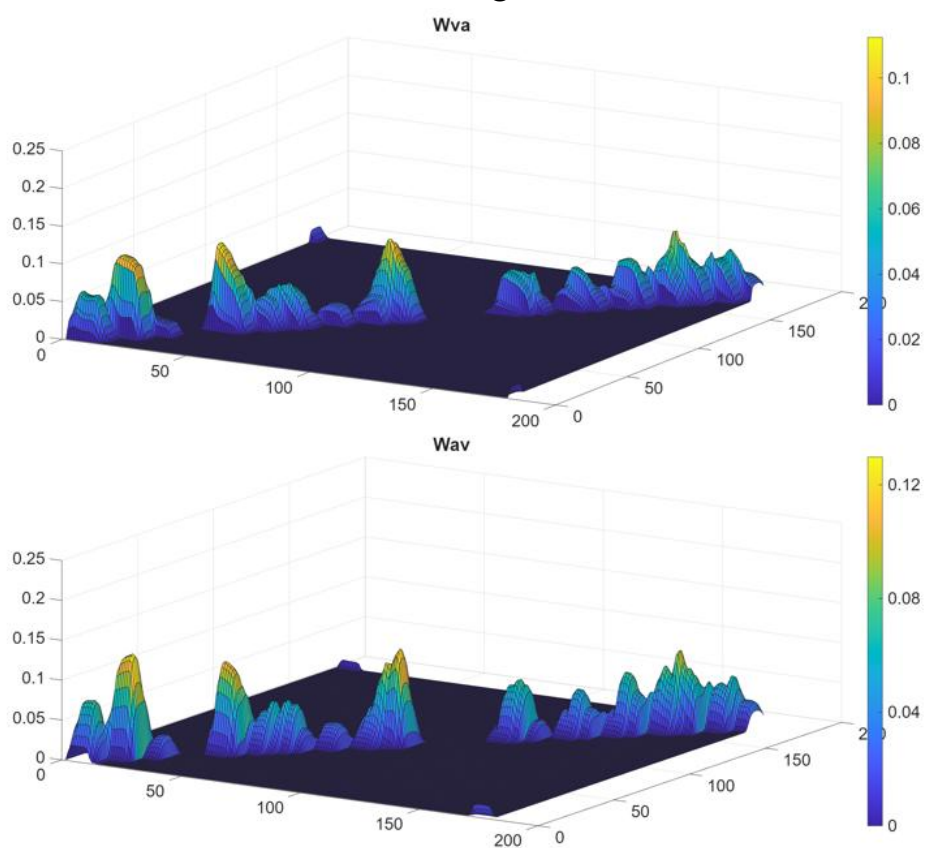
Feedforward con soglia: EPOCA 500



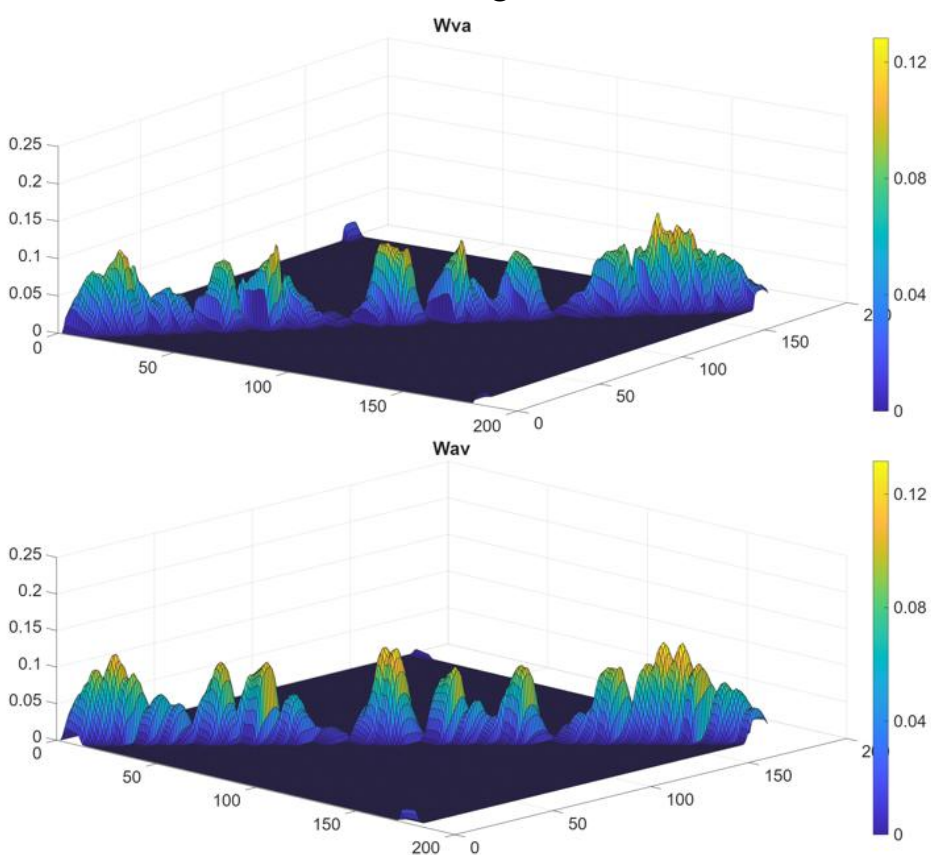
Cross-modali con soglia: EPOCA 0



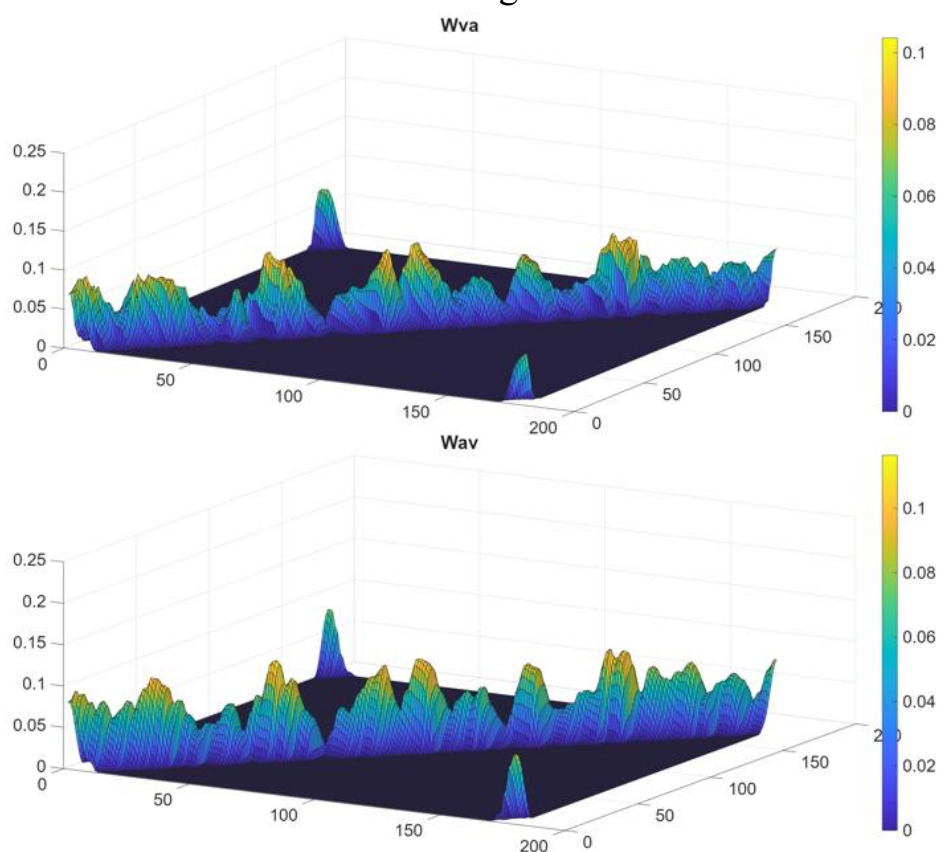
Cross-modal con soglia: EPOCA 125



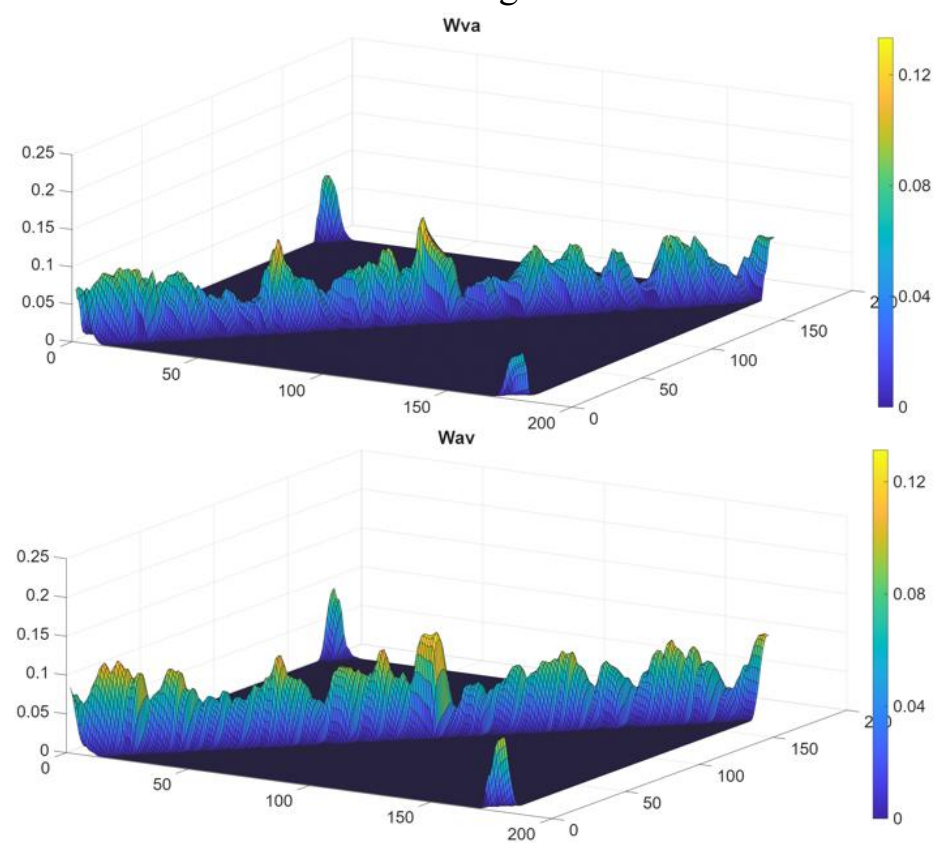
Cross-modal con soglia: EPOCA 250



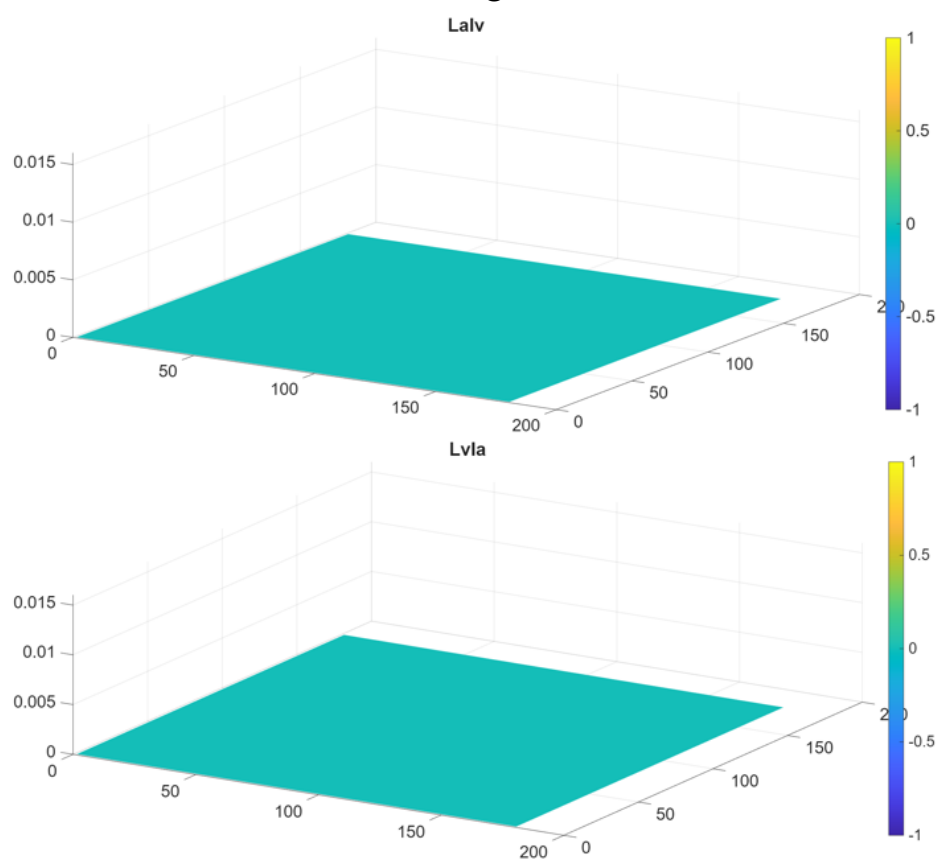
Cross-modali con soglia: EPOCA 375



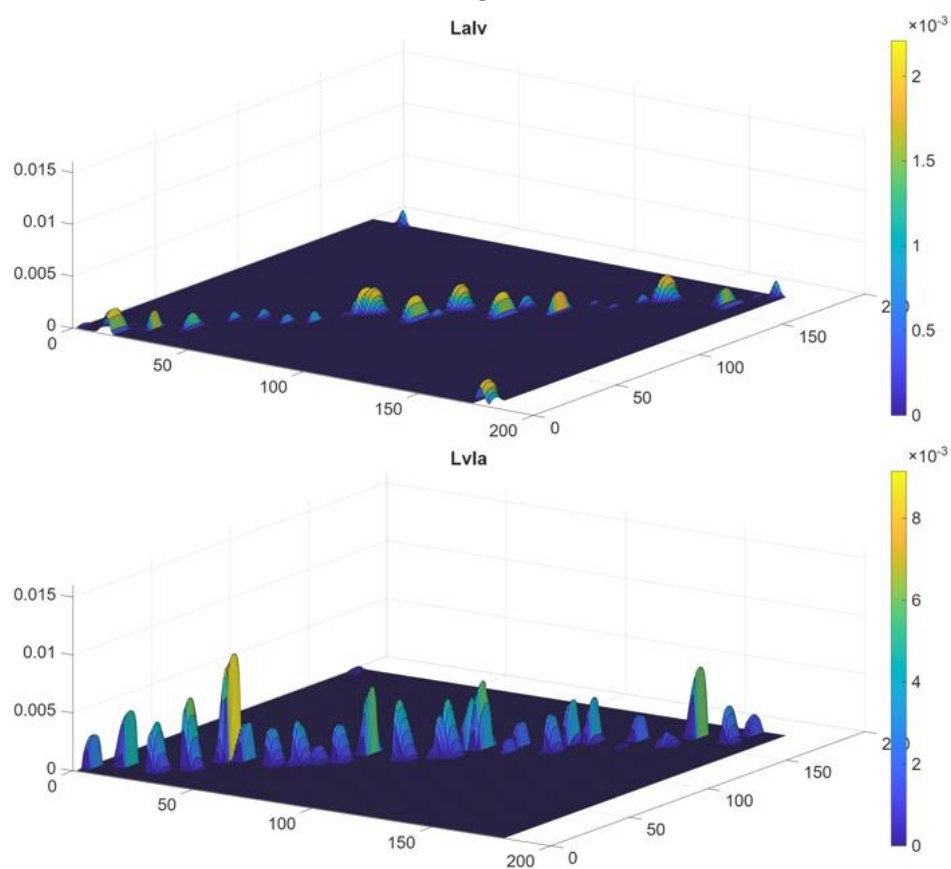
Cross-modali con soglia: EPOCA 500



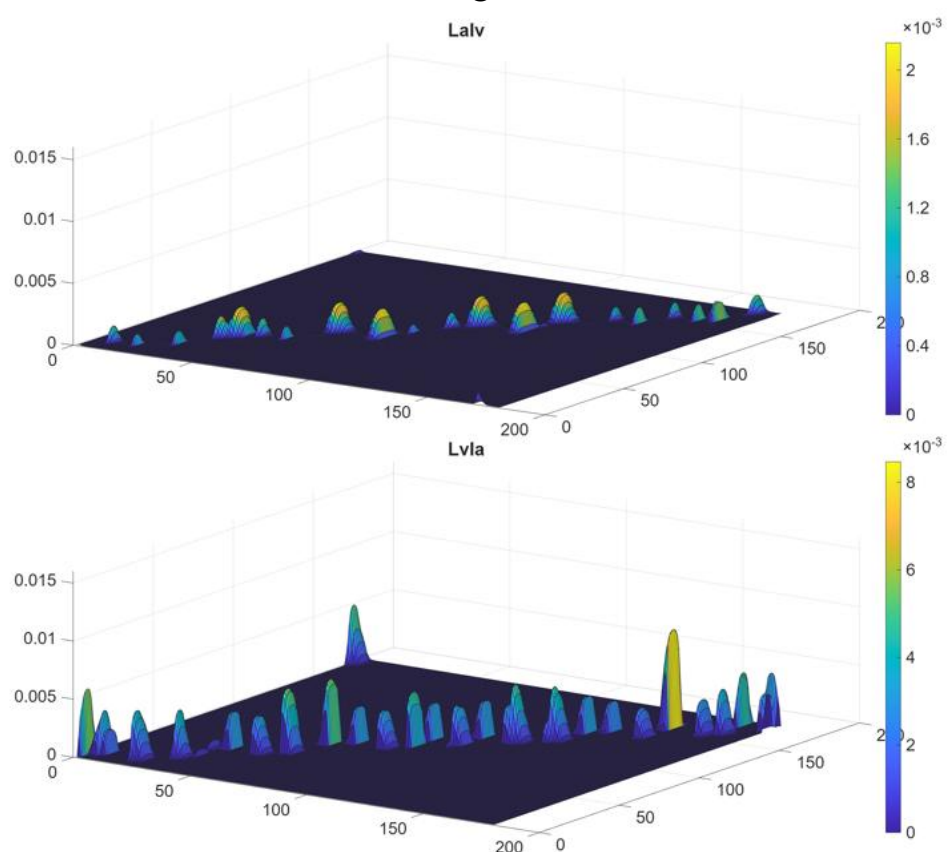
Inibitorie con soglia: EPOCA 0



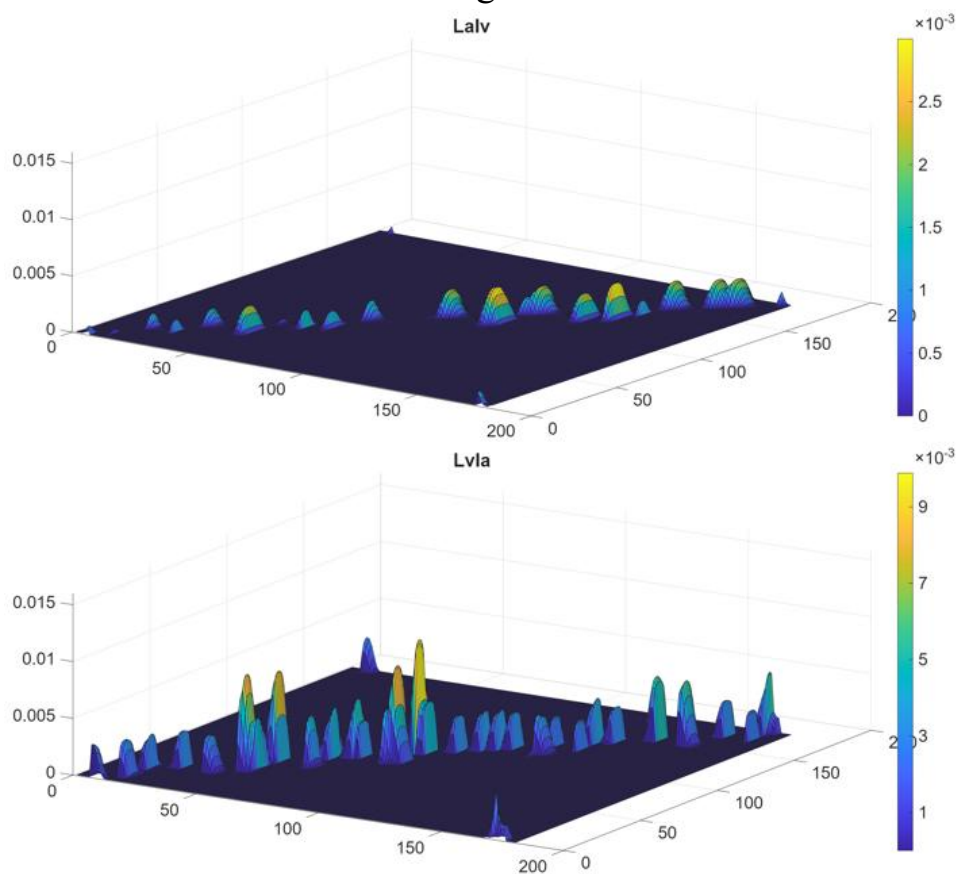
Inibitorie con soglia: EPOCA 125



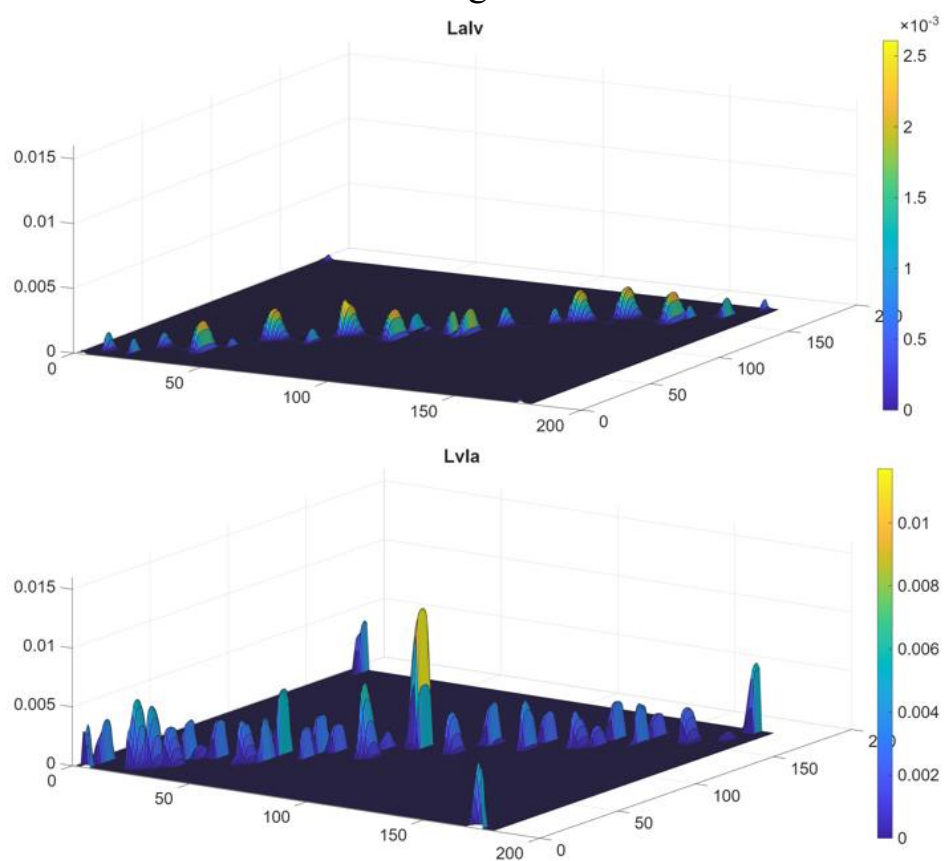
Inibitorie con soglia: EPOCA 250



Inibitorie con soglia: EPOCA 375



Inibitorie con soglia: EPOCA 500

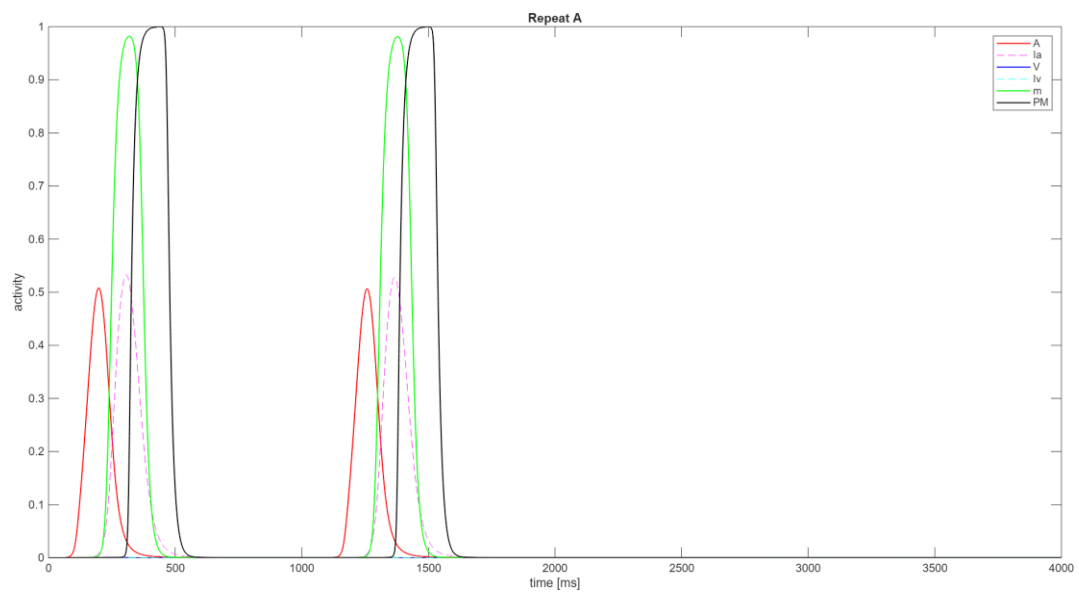
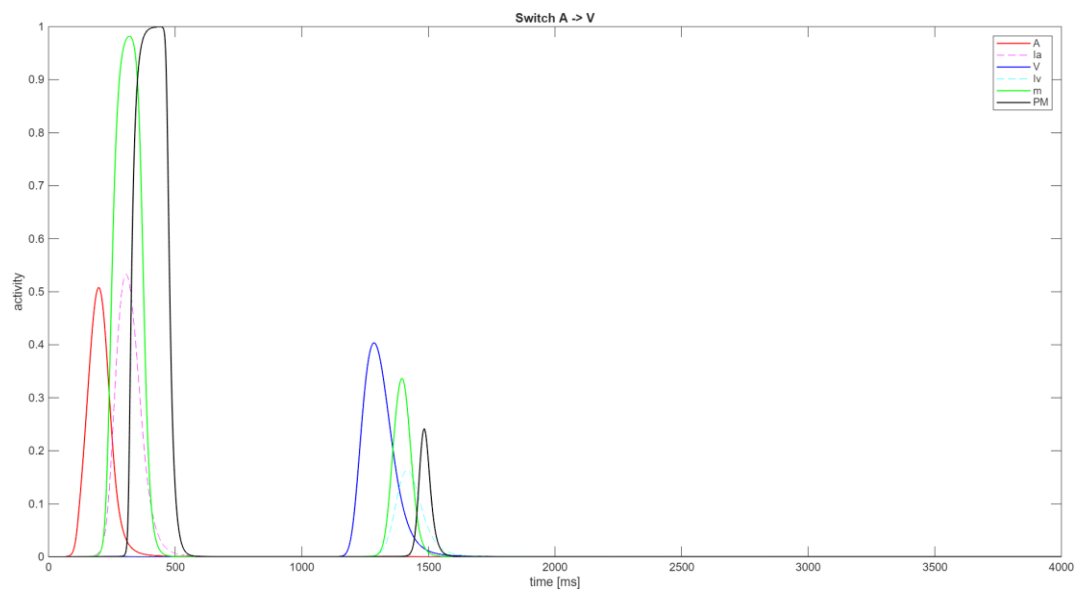
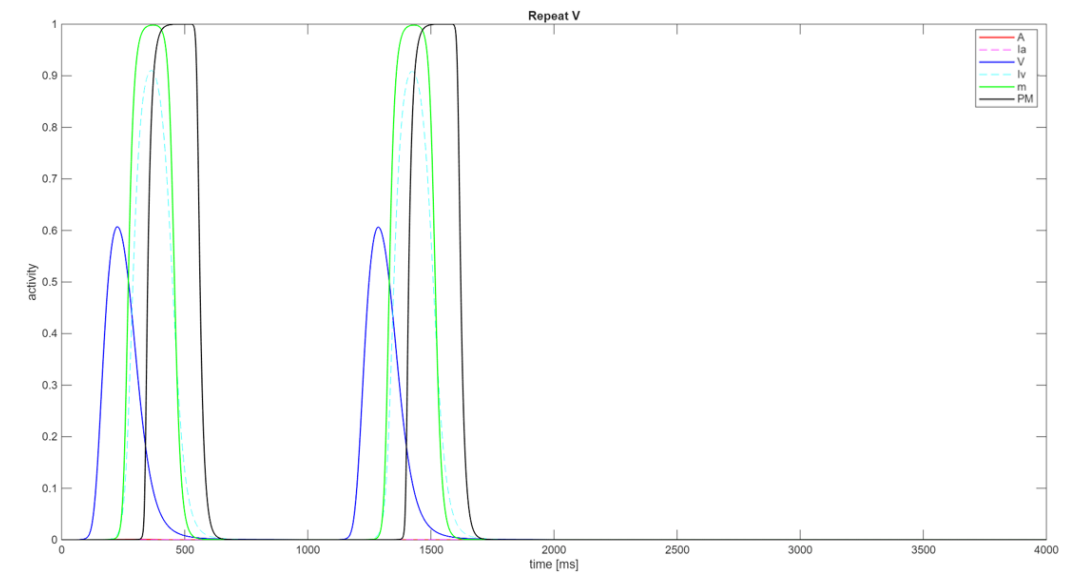


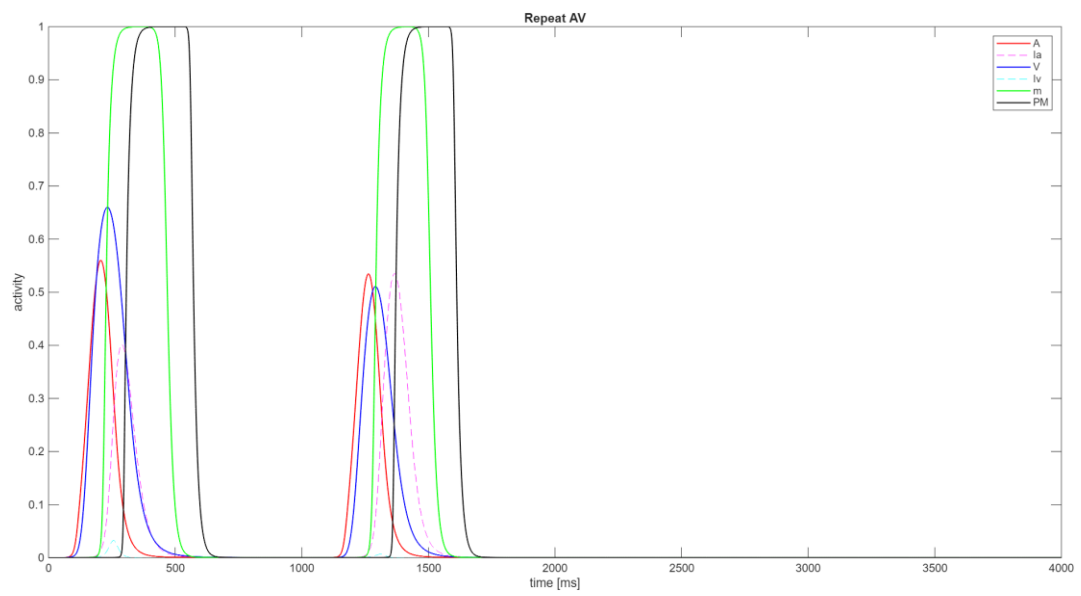
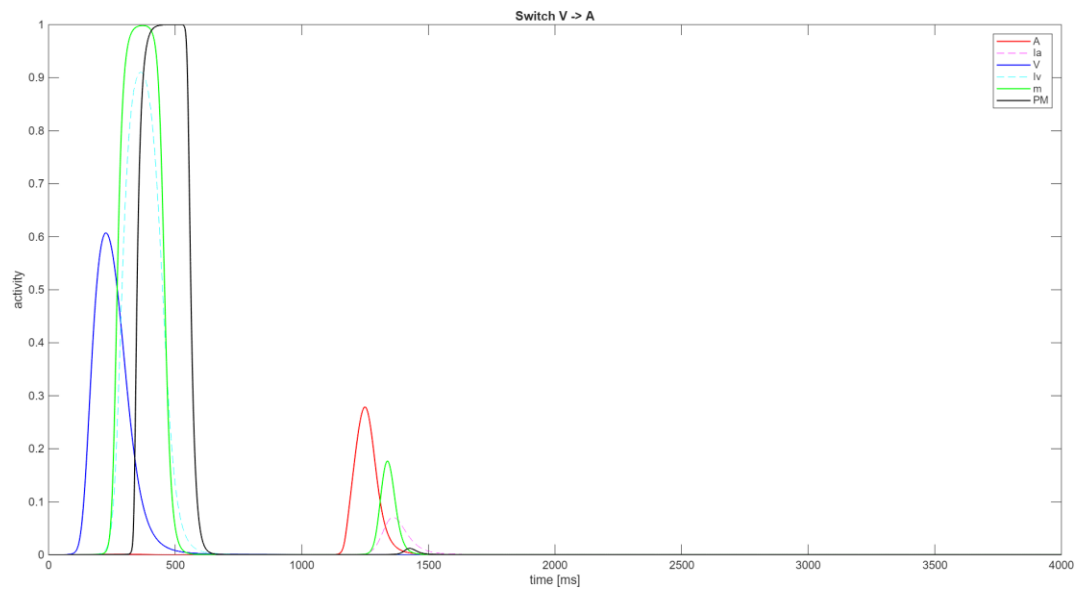
2.5 Differenze nei tempi di reazione tra cambi e ripetizioni in condizioni unisensoriali e multisensoriali

Nei grafici riportati di seguito si osserva che, in presenza di stimoli unisensoriali (solo A o solo V), i tempi di reazione risultano più lenti nei cambi (Sw) rispetto alle ripetizioni (Rp). Con stimoli multisensoriali (AV), invece, i tempi di reazione sono più rapidi, poiché l'effetto dell'inibizione cross-sensoriale è attenuato sia dal contributo eccitatorio cross-modale sia dalla maggiore attivazione dell'area multisensoriale. Quindi, con stimoli AV non c'è quasi differenza tra Sw e Rp e l'effetto del "cambio di modalità" non si osserva.

Nel caso del cambio uditivo, l'attività nella regione uditiva evocata dal secondo stimolo risulta inferiore rispetto a quella osservata nelle ripetizioni. La ridotta attivazione è prodotta dall'interneurone inibitorio eccitato dallo stimolo visivo precedente. Nel modello, l'efficacia di questa inibizione cross-modale per un lungo periodo è simulata per mezzo della dinamica lenta utilizzata per implementare l'effetto delle sinapsi inibitorie. La minore attivazione della regione uditiva produce un'attivazione ritardata della regione multisensoriale/motoria, con un tempo di reazione di conseguenza più lungo.

Risultati simili sono stati riscontrati nel caso del cambio visivo: un input visivo, preceduto da uno stimolo uditivo, susciterebbe un tempo di reazione più lungo rispetto alla condizione di ripetizione visiva. L'input uditivo precedente ha abbassato l'attività complessiva suscitata nella regione visiva dalla stimolazione esterna, determinando così un'attivazione più lenta dell'area multisensoriale e un tempo di reazione più lungo in risposta all'input visivo. Anche in questo caso, tale effetto e il relativo costo dello Sw visivo sono prodotti dall'inibizione di lunga durata guidata dall'interneurone uditivo nella regione visiva di ingresso.



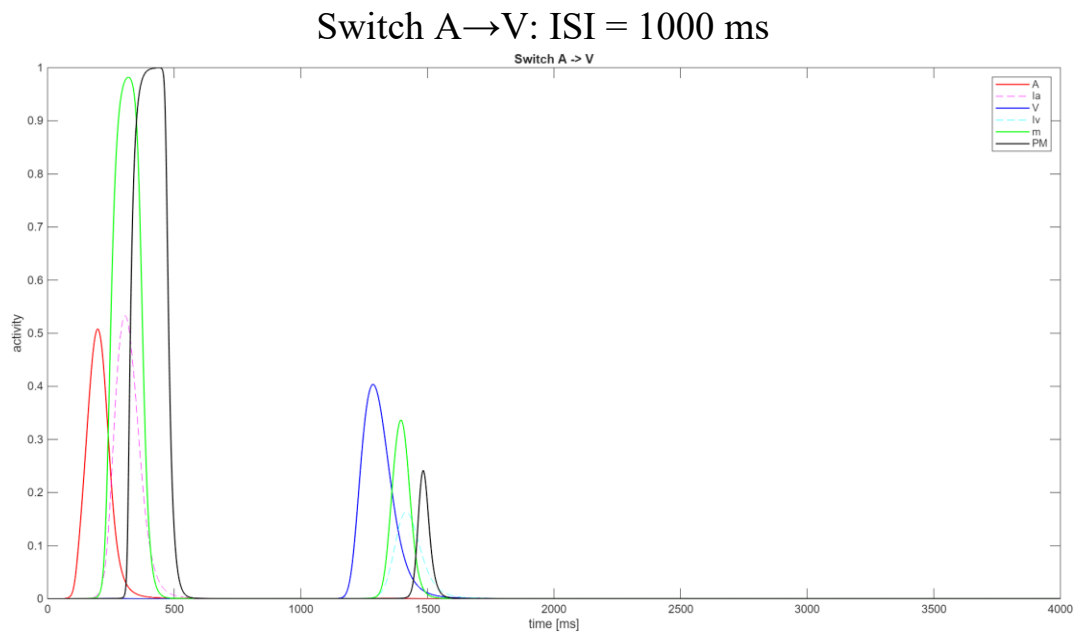


2.6 Influenza degli intervalli interstimolo sui tempi di reazione

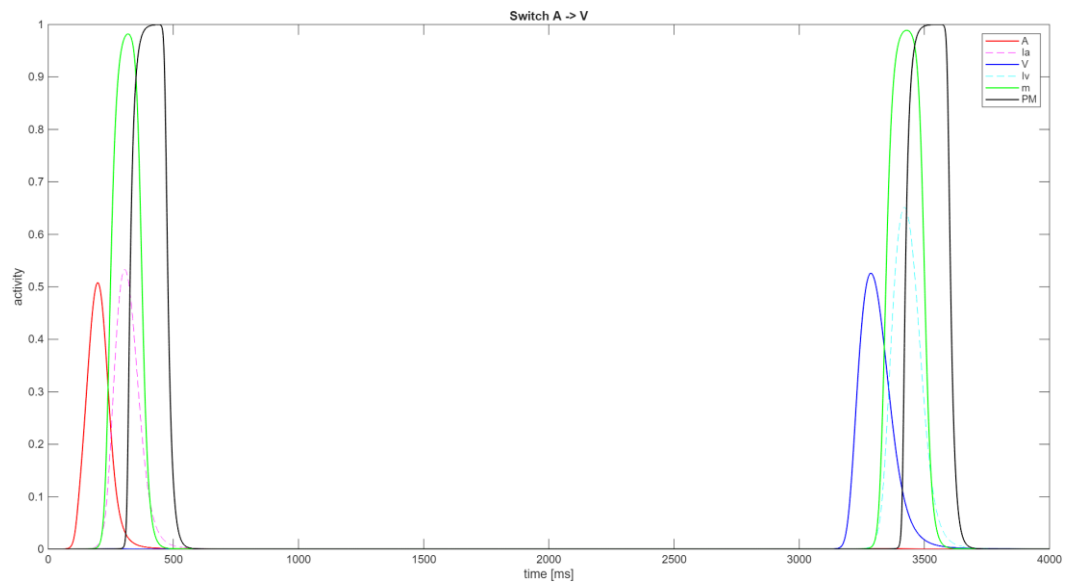
Sono state fatte simulazioni per capire come il tempo tra due stimoli (ISI) influenza i tempi di reazione (RTs) in prove di tipo Switch (Sw). L'aspettativa era che gli ISI brevi producessero un costo di Switch maggiore rispetto agli ISI lunghi, assumendo l'inibizione come meccanismo predominante. Questa ipotesi è stata confermata stimolando il modello con input uditivi e visivi unisensoriali intervallati da diversi ISI.

In presenza di un cambio uditivo con ISI breve, l'inibizione esercitata dall'input visivo precedente sull'input uditivo risulta più marcata rispetto a quanto accade con ISI lungo, determinando una minore attività nell'area uditiva e una risposta più lenta nella regione multisensoriale. Risultati analoghi sono stati osservati nella condizione di cambio visivo.

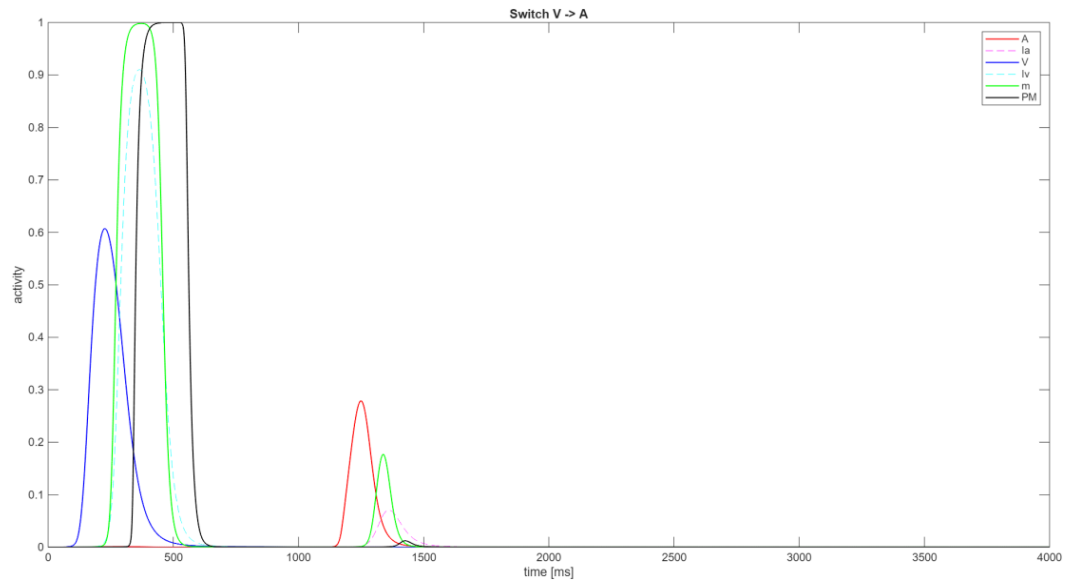
I grafici confermano quanto riportato sopra.



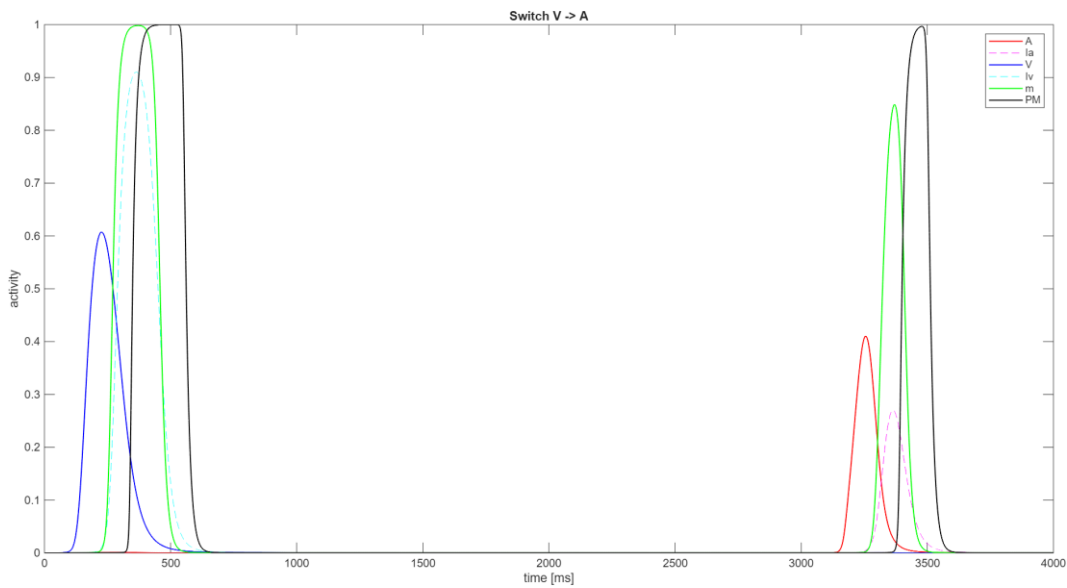
Switch A→V: ISI = 3000 ms



Switch V→A: ISI = 1000 ms



Switch V→A: ISI = 3000 ms



3. Discussione

I risultati ottenuti dalle diverse analisi forniscono una visione coerente e approfondita del ruolo del filtraggio tramite soglia e della regolazione dei parametri di crescita nell'apprendimento sinaptico. L'applicazione della soglia a tutte le sinapsi si dimostra cruciale per discriminare tra attività neuronale significativa e rumore di fondo, evitando che fluttuazioni minime o segnali casuali influenzino la formazione delle connessioni. Questo meccanismo ha un impatto marcato sulle sinapsi feedforward, che, grazie al filtraggio, raggiungono rapidamente la stabilizzazione dei pesi e assumono un ruolo trainante nella dinamica complessiva della rete. Le sinapsi cross-modali, pur risentendo in misura minore dell'effetto della soglia, mostrano una riduzione della variabilità e un'evoluzione più uniforme, mentre le sinapsi inibitorie, caratterizzate da un apprendimento più lento e graduale, contribuiscono a modulare e bilanciare l'attività delle aree sensoriali, esercitando un effetto stabilizzante sulla rete nel tempo.

L'analisi dei parametri di crescita conferma che la velocità di aggiornamento dei pesi è determinante per distinguere i diversi momenti di stabilizzazione: variando α e α_M è stato possibile evidenziare come il filtraggio alteri la sequenza temporale di consolidamento tra feedforward e cross-modali, enfatizzando il ruolo trainante delle prime.

Nei grafici ottenuti con stimoli singoli, la differenza tra presenza e assenza di soglia è evidente: senza filtraggio le matrici di connettività risultano più piatte e con variazioni minori, indicando un apprendimento poco selettivo, mentre con la soglia le connessioni si rafforzano solo in presenza di attività post-sinaptica significativa, generando mappe di attività specializzate e coerenti con principi di apprendimento biologico.

L'analisi dei grafici con stimoli sequenziali ha consentito di simulare un addestramento delle sinapsi $LvIa$ e $LaIv$ coerente con i principi biologici. Anche in questo caso, le sinapsi feedforward e cross-modali rafforzano rapidamente le correlazioni dirette e incrociate tra le modalità sensoriali, mentre le sinapsi inibitorie modulano gradualmente le risposte, riflettendo la necessità di integrazione temporale a lungo termine per garantire stabilità e bilanciamento. Questo comportamento evidenzia come diverse tipologie di sinapsi possano contribuire in modo complementare alla formazione di rappresentazioni sensoriali coerenti, con feedforward e cross-modali specializzate nell'apprendere rapidamente le correlazioni significative e l'inibizione capace di modulare e stabilizzare l'attività complessiva.

I risultati relativi ai tempi di reazione offrono ulteriori approfondimenti sul funzionamento dinamico del modello. Con stimoli unisensoriali (solo A o solo V), i tempi di reazione sono più lunghi nei trial di cambio (Sw) rispetto alle ripetizioni (Rp), suggerendo che l'adattamento alla nuova modalità sensoriale richiede più tempo. L'introduzione di stimoli multisensoriali (AV), invece, riduce significativamente il costo dello switch, indicando che il contributo eccitatorio cross-modale facilita l'attivazione delle aree multisensoriali e accelera la risposta. L'inibizione cross-modale svolge qui un ruolo fondamentale: agendo come meccanismo temporale dipendente, ritarda l'attivazione delle aree multisensoriali in presenza di stimoli unisensoriali, mentre con stimoli AV questa modulazione è ottimizzata per ridurre i tempi di reazione. Inoltre, gli intervalli interstimolo (ISI) influenzano in maniera evidente la dinamica della risposta: ISI brevi accentuano il costo dello switch, producendo una riduzione più marcata dell'attività nelle aree sensoriali e un ritardo nella risposta multisensoriale, mentre ISI più lunghi permettono un recupero dell'attività e un'integrazione più efficiente, confermando che l'inibizione è strettamente legata alla temporalità della stimolazione. Questi dati suggeriscono che il modello non solo replica le caratteristiche biologiche dell'apprendimento sinaptico, ma è anche capace di simulare dinamiche comportamentali complesse, evidenziando come feedforward, cross-modali e inibitorie cooperino per ottimizzare la risposta temporale e l'integrazione multisensoriale.

Conclusioni

In conclusione, i risultati ottenuti confermano che il filtraggio tramite soglia, unitamente a una regolazione accurata dei parametri di crescita, costituisce un meccanismo fondamentale per garantire robustezza, stabilità e specializzazione funzionale del sistema sinaptico. Questo approccio permette al modello di riprodurre in modo realistico i processi di apprendimento biologico, favorendo la formazione di mappe di attività coerenti, la modulazione temporale dei segnali in ingresso e un'integrazione multisensoriale efficiente. Inoltre, il modello consente di osservare l'evoluzione delle sinapsi nel tempo, offrendo una preziosa prospettiva su come il cervello modifichi le proprie connessioni e adatti le risposte sensoriali in funzione dell'esperienza. In questo senso, il modello non solo ricostruisce le dinamiche sinaptiche, ma permette anche di comprendere i meccanismi attraverso cui l'esperienza plasma la plasticità cerebrale, fornendo un quadro completo dei processi di apprendimento e adattamento neuronale.

Questi risultati rappresentano un riferimento significativo per la progettazione di compiti sperimentali più robusti, capaci di tenere conto delle differenze individuali e della natura dinamica dell'interazione multisensoriale, offrendo spunti per future indagini sul funzionamento del sistema nervoso e sulle strategie di integrazione sensoriale.

Bibliografia

- Ben-Yishai, R., Bar-Or, R. L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(9), 3844–3848.
- Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, 29(6), 578–584.
- Bolognini, N., Leo, F., Passamonti, C., Stein, B. E., & Làdavas, E. (2007). Multisensory-mediated auditory localization. *Perception*, 36(10), 1477–1485.
- Crosse, M. J., Foxe, J. J., Tarrit, K., Freedman, E. G., & Molholm, S. (2022). Resolution of impaired multisensory processing in autism and the cost of switching sensory modality. *Communications Biology*, 5(1), 601.
- Cuppini C., Di Rosa, E., Astolfi, L., et al. (2025). Unraveling audiovisual perception across space and time: A neuro-inspired computational architecture. *Authorea*, February 02.
- Cuppini C., Magosso, E., Bolognini, N., Vallar, G., & Ursino, M. (2014). A neurocomputational analysis of the sound-induced flash illusion. *NeuroImage*, 92, 248–266.
- Cuppini C., Shams, L., Magosso, E., & Ursino, M. (2017). A biologically inspired neurocomputational model for audiovisual integration and causal inference. *The European Journal of Neuroscience*, 46(9), 2481–2498.
- Cuppini C., Ursino, M., Magosso, E., et al. (2020). Cross-sensory inhibition or unisensory facilitation: A potential neural architecture of modality switch effects. *Journal of Mathematical Psychology*, 99, 102438.
- Diederich, A., & Colonius, H. (2004a). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, 66(8), 1388–1404.
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16(5).
- Jansen, B., & Rit, V. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4), 357–366.
- Magosso, E., Cuppini, C., & Ursino, M. (2012). A neural network model of ventriloquism effect and aftereffect. *Plos One*, 7(8), e42503.
- Musacchia, G., & Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hearing Research*, 258(1–2), 72–79.

- Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research*, 258(1–2), 89–99.
- Recanzone, G. H., Guard, D. C., & Phan, M. L. (2000). Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *Journal of Neurophysiology*, 83(4), 2315–2331.
- Rohe, T., & Noppeney, U. (2015). Sensory reliability shapes perceptual inference via two mechanisms. *Journal of Vision*, 15(5), 22.
- Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., & Woldorff, M. G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*, 45(3), 561–571.
- Shaw, L. H., Freedman, E. G., Crosse, M. J., Nicholas, E., Chen, A. M., Braiman, M. S., Molholm, S., & Foxe, J. J. (2020). Operating in a Multisensory Context: Assessing the Interplay Between Multisensory Reaction Time Facilitation and Inter-sensory Task-switching Effects. *Neuroscience*, 436, 122–135.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. The MIT Press.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), 255–266.
- Treves, A. (1993). Mean-field analysis of neuronal spike dynamics. *Network*, 4, 259–284.
- Ursino, M., Cuppini, C., & Magosso, E. (2014). Neurocomputational approaches to modelling multisensory integration in the brain: A review. *Neural Networks*, 60, 141–165.
- Von Saldern, S., & Noppeney, U. (2013). Sensory and striatal areas integrate auditory and visual signals into behavioral benefits during motion discrimination. *The Journal of Neuroscience*, 33(20), 8841–8849.
- Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research*, 158(2).
- Wendling, F., et al. (2002). Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *European Journal of Neuroscience*, 15(9), 1499–1508.