

POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Ingegneria Biomedica

Tesi di Laurea Magistrale

Algoritmo per la scelta del numero ottimale di sinergie muscolari
(ChoOSyn) e sua validazione



Relatore:
prof.ssa Valentina Agostini

Correlatore:
ing. Marco Ghislieri

Candidato:
Riccardo Ballarini

Anno Accademico 2019/2020

Sommario

1. Introduzione	1
1.1 Le sinergie muscolari	1
1.2 Estrazione delle sinergie muscolari	3
1.3 Obiettivo della tesi	4
2. Materiali e Metodi.....	6
2.1 Dataset reale	6
2.2 Dataset simulato (sintetico)	6
2.3 Pre-processing	9
2.4 Algoritmo di fattorizzazione	11
2.5 Selezione del numero ottimale di sinergie muscolari	14
2.5.1 Correlazione <i>intra-cluster</i> (<i>Intra-Cluster Correlation</i>)	14
2.5.2 Similarità dei pesi muscolari (<i>Weight Similarity</i>).....	15
2.5.3 Diversità dei coefficienti di attivazione (<i>Coefficient Dissimilarity</i>).....	16
2.5.4 Algoritmo ChoOSyn	17
2.6 Valutazione delle performance	20
3. Risultati.....	21
3.1 Ottimizzazione dell'algoritmo di fattorizzazione.....	21
3.2 Parametri.....	22
3.2.1 Intra-Cluster Correlation (ICC).....	22
3.2.2 Weight Similarity (WS)	26
3.2.3 Coefficient Dissimilarity (CD).....	28
3.3 Algoritmo ChoOSyn	29
3.3.1 Dataset simulato	30
3.3.2 Dataset reale	32
4. Discussioni	34
4.1 Conclusioni	34
4.2 Limiti e Prospettive Future	34
5. Informazioni Supplementari.....	36
5.1 Soggetto sano #7	36
5.2 Soggetto sano #12	38
5.3 Soggetto sano #19	40
5.4 Soggetto sano #21	42

1. Introduzione

1.1 Le sinergie muscolari

Le sinergie muscolari sono uno strumento molto importante poiché permettono non solo di comprendere i meccanismi alla base del controllo motorio in maniera quantitativa e non invasiva (tramite elettrodi posizionati sulla pelle in corrispondenza della fascia muscolare), ma anche di monitorare persone affette da malattie neurologiche/neurodegenerative [1], [2] o problemi articolari [3], di effettuare confronti pre/post-trattamento [4], [5], o di controllare componenti robotici o esoscheletri [6], [7].

L'ipotesi delle sinergie muscolari permette di spiegare come il sistema nervoso centrale (SNC) sia in grado di controllare con relativa semplicità un sistema apparentemente complesso, con circa seicento muscoli e duecento articolazioni. Infatti, alla base di questa ipotesi vi è la capacità del SNC di semplificare l'elevato numero di gradi di libertà grazie alla combinazione di pochi elementi discreti [8]. In questo modo è possibile controllare il movimento più semplicemente, mediante un numero ridotto di comandi neurali [Fig. 1A] indirizzati a specifici gruppi muscolari [Fig. 1B, 1C] (chiamati sinergie muscolari). Le sinergie muscolari sono composte da due elementi: i pesi muscolari [Fig. 1B] ed i coefficienti di attivazione [Fig. 1A]. Nel dettaglio, i coefficienti di attivazione sono la componente temporale delle sinergie muscolari, e rappresentano i profili temporali di attivazione che vengono condivisi tra i vari muscoli appartenenti alla stessa sinergia muscolare [9]. Invece, i pesi muscolari sono la componente spaziale delle sinergie muscolari, e definiscono il contributo di ogni muscolo ad una specifica sinergia muscolare [9]. Riassumendo, il SNC tramite un set ridotto di comandi neurali (coefficienti di attivazione) è in grado di comandare un ampio numero di muscoli, poiché vengono raggruppati in base al movimento da eseguire, ed ogni gruppo (i cui muscoli di appartenenza vengono definiti dai pesi muscolari) riceve un determinato segnale di attivazione [Fig. 1].

Nel caso specifico, in questo lavoro verranno presi in esame dodici muscoli che intervengono durante il cammino. Mediamente, durante questo *task* motorio e con questa numerosità di muscoli, un soggetto sano esprime cinque sinergie muscolari [3], [10]–[13] [Fig. 1], ognuna delle quali ha un suo preciso significato fisiologico. Infatti, come si può notare dal grafico [Fig. 1], la sinergia magenta comprende alcuni muscoli della coscia, che flettono la coscia sull'anca portando avanti la gamba durante il cammino e stabilizzano l'anca nella fase di appoggio del piede; la sinergia ciano raggruppa i muscoli del polpaccio, che tramite la flessione plantare e la flessione della gamba generano la propulsione per avanzare durante il cammino; la sinergia rossa rappresenta il solo tibiale anteriore, che dorsiflette la punta del piede prima del contatto iniziale, assicurandosi inoltre che resti adeguatamente discostata dal suolo (*clearance*) durante la fase di oscillazione della gamba; la sinergia verde comprende i muscoli posteriori della coscia, che flettono il ginocchio sollevando la parte inferiore della gamba durante la fase di oscillazione e frenando il pendolamento; la sinergia blu, infine, comprende i due muscoli dorsali, i quali forniscono stabilità ed equilibrio al tronco sul piano frontale in corrispondenza dell'inizio di ciascuna delle due fasi di appoggio di entrambi i piedi (in particolare durante il trasferimento del carico da un singolo arto inferiore ad entrambi gli arti inferiori).

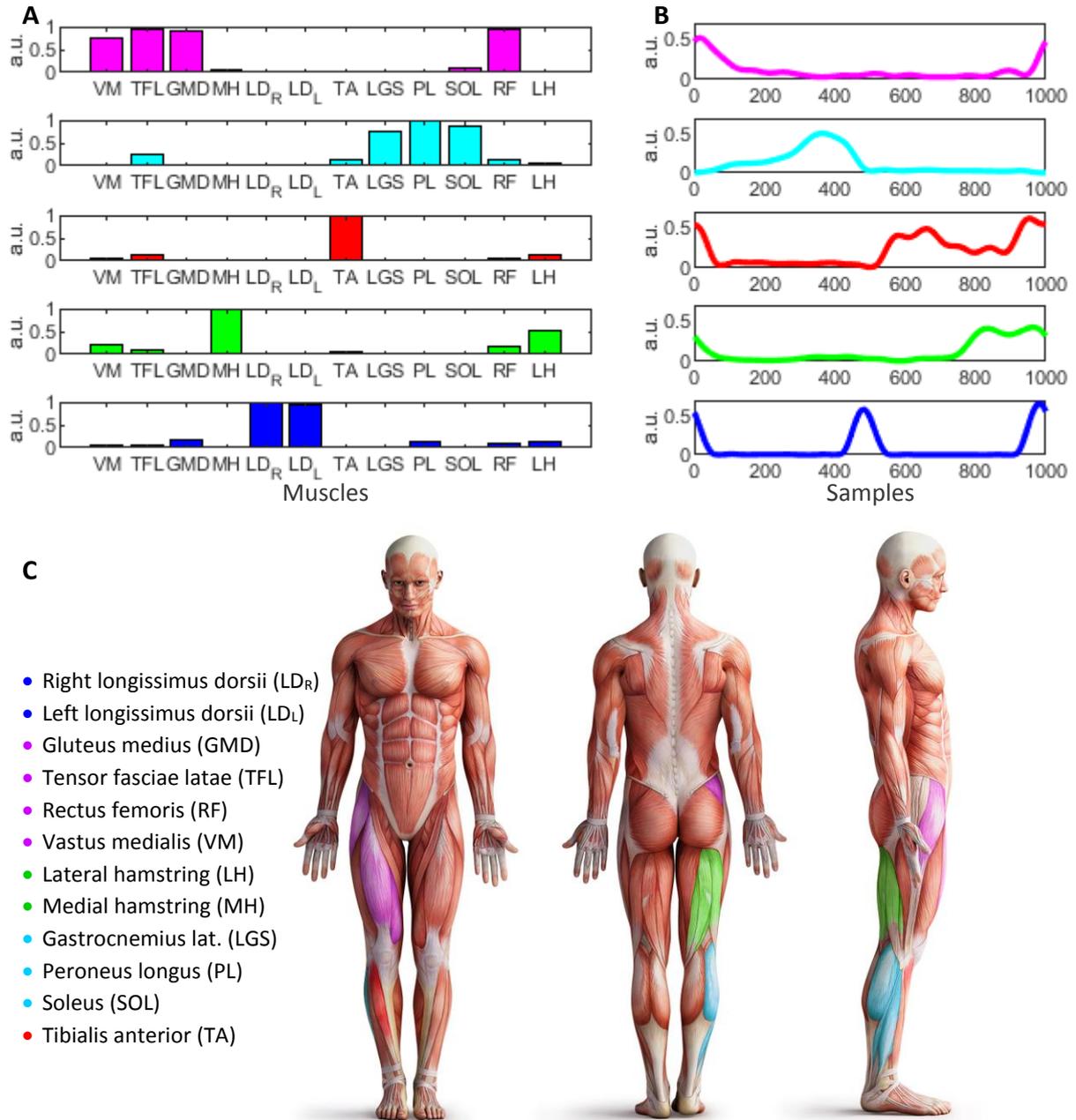


Figura 1: Esempio di cinque sinergie muscolari, composte da **(A)** coefficienti di attivazione e **(B)** pesi muscolari, con **(C)** la rappresentazione anatomica dei dodici muscoli raggruppati nei cinque pesi muscolari (i due muscoli dorsali lunghissimi non sono visibili, ed il prelievo è avvenuto ai lati della vertebra L4).

1.2 Estrazione delle sinergie muscolari

L'estrazione delle sinergie muscolari avviene a partire dai segnali elettromiografici di superficie (sEMG), e si realizza tramite algoritmi di fattorizzazione e criteri di determinazione del numero ottimale di sinergie muscolari. Nel dettaglio, la fattorizzazione permette di ridurre la dimensionalità del set di dati originari, scomponendo la matrice di partenza (contenente i segnali sEMG di tutti i muscoli) in due matrici: i pesi muscolari (W) ed i coefficienti di attivazione (C) [Fig. 2A]. L'algoritmo di fattorizzazione maggiormente utilizzato in letteratura è la *Non-negative Matrix Factorization* (NMF), poiché è l'unica tecnica che non impone vincoli alla matrice di partenza, ad eccezione ovviamente della non-negatività [14]. In seguito alla fattorizzazione, la qualità di ricostruzione [Fig. 2B, C] dei segnali sEMG originari viene generalmente valutata mediante il parametro della *Variance Accounted For* (VAF). Con il termine "ricostruzione" si intende il prodotto tra pesi muscolari e coefficienti di attivazione [Fig. 2A], che genera appunto i dati di partenza al netto dell'errore di ricostruzione. Nell'algoritmo di fattorizzazione la numerosità delle sinergie deve essere scelta a priori, perciò, dopo aver applicato più volte la fattorizzazione con diverse numerosità, è necessario introdurre un criterio per poter scegliere la numerosità corretta tra le varie testate. Per selezionare il numero di sinergie vi sono svariati metodi, ed in letteratura quello utilizzato dalla quasi totalità degli studiosi è l'imposizione di una soglia minima alla curva VAF [1]–[3], [6], [11], [15]–[17], che varia a seconda degli studi tra l'80% ed il 95% [18]. La curva VAF non è altro che la sequenza dei valori del parametro VAF al variare del numero di sinergie muscolari. Visto che tale curva rappresenta l'accuratezza di ricostruzione dell'output della fattorizzazione, imporre una soglia equivale ad indicare il valore massimo di errore che si è disposti ad accettare ed a selezionare la prima numerosità di sinergie muscolari che soddisfa tale condizione. Altri studi invece utilizzano sempre la curva VAF ma vi applicano criteri differenti dalla soglia, come ad esempio la ricerca del punto in cui vi è un significativo cambio di pendenza [14], o la ricerca di un plateau oltre al quale la situazione resta pressoché immutata [19], [20]. Altri studiosi ancora preferiscono utilizzare il coefficiente di determinazione R^2 [21], [22], il quale però non differisce molto dal VAF in quanto è anch'esso un indicatore dell'accuratezza di ricostruzione.

Il principale limite di questi metodi è che la scelta effettuata è spesso scorrelata dai parametri delle sinergie muscolari quali la consistenza e la similarità: infatti non si tiene in considerazione la consistenza delle strategie di controllo motorio durante l'esecuzione del movimento, la similarità tra sinergie muscolari differenti per valutare casi di split indesiderati, e neanche il significato fisiologico delle sinergie muscolari, ma ci si limita a fissare una soglia arbitraria sulla qualità di ricostruzione dell'algoritmo di fattorizzazione. Infatti, ad esempio, imporre una soglia al 90% a priori può escludere un risultato con un VAF dell'89%, nonostante magari dal punto di vista fisiologico possa essere una soluzione migliore. Il metodo della soglia applicata alla curva VAF è perciò un criterio adatto se si deve semplicemente ridurre la dimensionalità di dati generici, ma non è un metodo ottimale nel caso delle sinergie muscolari, in cui lo scopo finale può essere l'osservazione clinica dei comandi motori, e per le quali serve un metodo più elaborato che tenga conto anche del significato fisiologico che l'output finale deve avere.

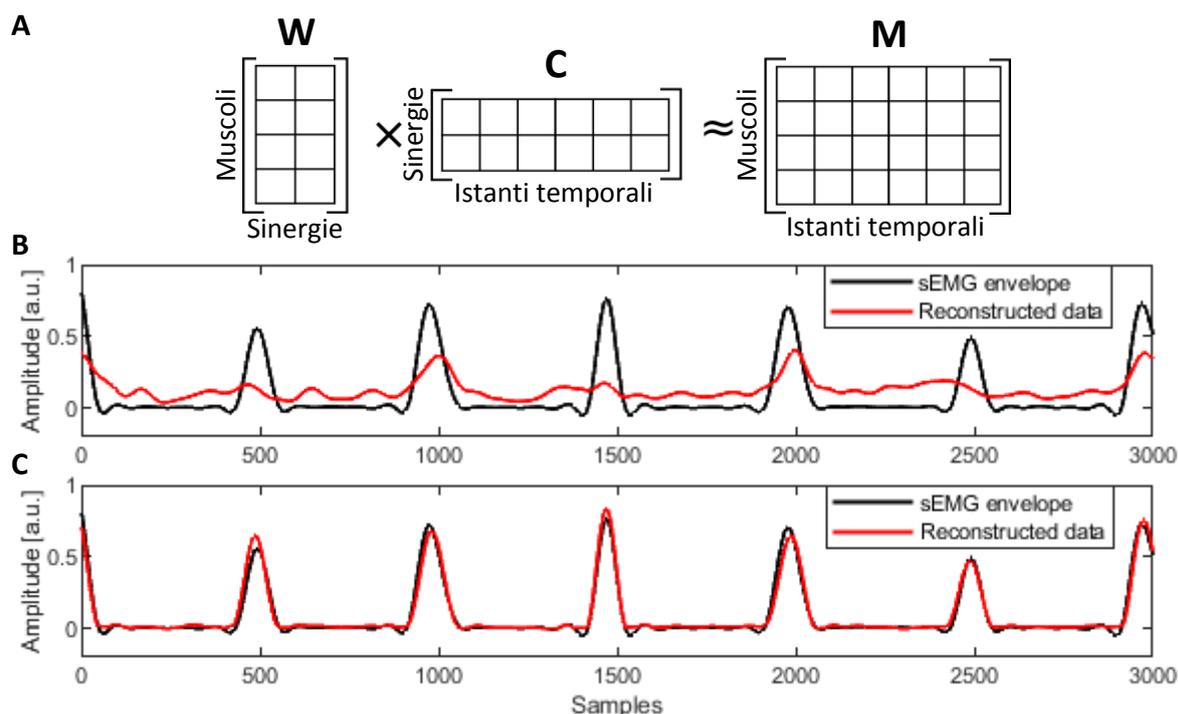


Figura 2: (A) Riduzione della dimensionalità della matrice di dati M , scomposta in pesi muscolari (W) e coefficienti di attivazione (C). (B-C) Confronto tra i dati di partenza (nero) e la ricostruzione dopo la fattorizzazione (rosso) con VAF = 62% (B) e con VAF = 95% (C); in entrambi i casi si può notare una differenza (errore di ricostruzione), tanto più piccola quanto è maggiore il valore del parametro VAF. Nell'esempio sono riportati tre cicli di cammino del muscolo dorsale lunghissimo destro (LD_R).

1.3 Obiettivo della tesi

In questo lavoro presenterò quindi un metodo alternativo per la scelta del numero ottimale di sinergie muscolari, basato su criteri che tengano conto della consistenza e della similarità tra sinergie muscolari. Le sinergie muscolari contengono intrinsecamente informazioni fisiologiche che non possono essere trascurate, perciò sono necessari nuovi indici che osservino direttamente pesi muscolari e coefficienti di attivazione, per trovare la soluzione ottimale. Più semplicemente, l'obiettivo è individuare la situazione in cui né i pesi muscolari né i coefficienti di attivazione si "sovrappongano" tra di loro (ovvero situazioni con valori di correlazione troppo elevati) e la cui evoluzione temporale tra i vari *task* motori sia il più possibile stabile. Entrambe queste situazioni sono infatti sfavorevoli e da evitare, poiché sinergie muscolari molto simili tra loro contengono informazioni ridondanti, mentre sinergie muscolari troppo variabili tra una ripetizione e l'altra di uno stesso *task* motorio non descrivono correttamente il comportamento dei muscoli.

Questo nuovo metodo è stato testato sia su sEMG simulati, opportunamente costruiti per ricreare un comportamento fisiologico, sia su sEMG reali, provenienti sia da soggetti sani che da soggetti patologici affetti da morbo di Parkinson. Nel caso dei segnali simulati, il numero di sinergie muscolari è stato deciso a priori in fase di generazione, mentre nei casi reali è stato utilizzato un *ground truth* fornito da due operatori esperti. L'utilità di utilizzare entrambi i tipi di dataset sta nella possibilità di testare l'algoritmo su due tipi diversi di dati, ognuno con i propri vantaggi e svantaggi: infatti, nel dataset sintetico i segnali sono più semplici da fattorizzare, si ottengono sinergie muscolari più "pulite"

ed è possibile affermare in modo oggettivo quale sia la numerosità di sinergie muscolari corretta, in quanto i dati sono simulati a partire da un numero imposto di sinergie muscolari; invece, nel dataset reale è possibile testare l'algoritmo sui segnali di persone reali, che è effettivamente lo scopo finale dell'algoritmo, ma l'oggettività del *ground truth* non è garantita, in quanto può variare in base agli operatori che effettuano le scelte. Perciò, il dataset simulato ha lo scopo di testare *se* l'algoritmo funziona, mentre il dataset reale fornisce informazioni su *come* funziona in una situazione reale (ad esempio se potrà essere utile per uno studio clinico).

2. Materiali e Metodi

2.1 Dataset reale

Per poter testare il lavoro su dei soggetti reali, mi è stato fornito un dataset sperimentale acquisito durante precedenti studi presso il Biolab del Politecnico di Torino [12], [13], [23]; perciò, per quanto riguarda il dataset reale, in questo lavoro è stata svolta un'analisi retrospettiva dei dati, e in questo paragrafo verranno descritte le caratteristiche del dataset e come è stato ottenuto dagli altri ricercatori.

Il dataset contiene principalmente segnali sEMG che, come detto in precedenza, costituiscono l'input dell'algoritmo di estrazione delle sinergie muscolari. La popolazione comprende 30 soggetti sani (8 maschi, con età: 57.4 ± 9.2 anni, altezza: 1.73 ± 0.09 m, peso: 85.5 ± 22.7 kg; 22 femmine, con età: 37.5 ± 15.5 anni, altezza: 1.64 ± 0.07 m, peso: 60.2 ± 14.9 kg) e 20 soggetti patologici affetti da morbo di Parkinson (12 maschi, con età: 58.6 ± 10.3 anni, altezza: 1.79 ± 0.06 m, peso: 77.3 ± 11.0 kg; 8 femmine, con età: 57.4 ± 6.5 anni, altezza: 1.63 ± 0.09 m, peso: 70.8 ± 16.1 kg). Ho scelto di utilizzare sia soggetti sani che patologici per poter spaziare su più casi reali differenti, in modo da osservare eventuali differenze nel comportamento dell'algoritmo sulle due diverse condizioni del soggetto.

I segnali sEMG sono stati raccolti con un dispositivo multicanale specificatamente sviluppato per l'analisi del cammino (STEP32, Medical Technology, Italia), dotato di elettrodi da applicare sulla pelle (opportunamente preparata) in corrispondenza dei muscoli che si desidera monitorare.

Il *task* motorio richiesto ai volontari consiste in una camminata lungo un percorso rettilineo di 10 m a velocità normale (scelta dal soggetto), perciò il prelievo dei segnali sEMG è stato effettuato da dodici muscoli della gamba dominante e del tronco del soggetto, tutti coinvolti nella deambulazione: entrambi i dorsali lunghissimi (destro e sinistro) (LD_R e LD_L), gluteo medio (GMD), tensore della fascia lata (TFL), i due ischiocrurali (bicipite femorale e semitendinoso/semimembranoso) (MH e LH), retto femorale (RF), vasto mediale (VM), gastrocnemio laterale (LGS), soleo (SOL), tibiale anteriore (TA) e peroneo lungo (PL) [Fig. 1].

Oltre al prelievo dei segnali sEMG dai muscoli è importante avere anche informazioni sul ciclo del cammino, tramite sensori basografici, situati sotto la pianta del piede in corrispondenza della prima e della quinta punta metatarsale e del tallone, ed elettrogoniometri, posizionati lateralmente al ginocchio, per valutare gli angoli articolari sul piano frontale.

2.2 Dataset simulato (sintetico)

Per la costruzione e la validazione dell'algoritmo di selezione del numero di sinergie è stato utilizzato un dataset di sEMG simulati generati da una numerosità di sinergie muscolari imposta a priori. Sono stati simulati sEMG generati da 4, 5 e 6 sinergie muscolari, corrispondenti alle numerosità che normalmente si riscontrano su soggetti sani durante il cammino [3], [10]–[13]. Nel dataset sono contenuti segnali sEMG a differenti valori di SNR, compresi tra 15 dB a 30 dB (a passi di 5 dB). Per ogni numerosità di sinergie muscolari sono stati generati 25 gruppi di segnali sEMG simulati senza rumore additivo (che successivamente verrà indicato con la dicitura “no noise”) e 100 gruppi con la presenza di rumore (25 gruppi per ogni valore di SNR, utilizzando 15 dB, 20 dB, 25 dB e 30 dB); perciò in totale il dataset comprende 375 gruppi di segnali sEMG sintetici, 125 per ogni numerosità di sinergie o 75 per ogni valore di SNR. Solo il sottogruppo generato da 5 sinergie muscolari è stato utilizzato per la

costruzione dell'algoritmo per la scelta del numero ottimale di sinergie, gli altri due sottogruppi sono stati utilizzati come test dell'algoritmo.

Per la generazione dei segnali simulati è stata usata come base di partenza una raccolta di pesi muscolari e coefficienti di attivazione [Fig. 3A] presi dai soggetti reali sani citati precedentemente. Successivamente, pesi muscolari e coefficienti di attivazione sono stati moltiplicati tra di loro per ricostruire il set di involucri [Fig. 3C] dei segnali sEMG:

$$envelope_{simulato}(t) = W * C(t) \quad (1)$$

In fase di moltiplicazione, i pesi muscolari e i coefficienti di attivazione sono stati opportunamente accoppiati tramite *clustering* per mantenere intatto il significato fisiologico che ogni sinergia muscolare ha con sé. Inoltre, pesi muscolari e coefficienti di attivazione sono stati accoppiati in modo casuale tra di loro, in modo da ottenere sEMG ibridi, con pesi muscolari di un soggetto e coefficienti di attivazione di un altro soggetto; questo accoppiamento misto è stato fatto per poter generare un dataset più vasto e vario.

Dopo aver ottenuto gli involucri dei segnali dei vari muscoli, essi vengono moltiplicati per una prima distribuzione gaussiana (*media* = 0, *deviazione standard* = 1), per ottenere l'aspetto tipico di un segnale sEMG [Fig. 3C]:

$$sEMG_{simulato} = envelope_{simulato} * randn \quad (2)$$

ed addizionati ad una seconda distribuzione gaussiana (*media* = 0, *deviazione standard* = $1/10^{SNR/20}$) a varianza minore, per applicare rumore con un SNR definito [Fig. 3D]:

$$sEMG_{simulato} = sEMG_{simulato} + \frac{randn}{10^{SNR/20}} \quad (3)$$

Nelle formule riportate, "randn" corrisponde alla funzione MATLAB® con cui si genera una sequenza causale con distribuzione gaussiana a media nulla e varianza unitaria.

Ciascuna simulazione è costituita dagli stessi dodici muscoli utilizzati nei soggetti reali e cinquanta cicli del cammino (ognuno costituito da mille campioni), ovvero matrici di dimensioni 12×50000; ognuno dei 50 cicli del cammino è diverso dagli altri, in modo da simulare correttamente anche la variabilità tra i vari passi che normalmente si presenta durante le acquisizioni sperimentali.

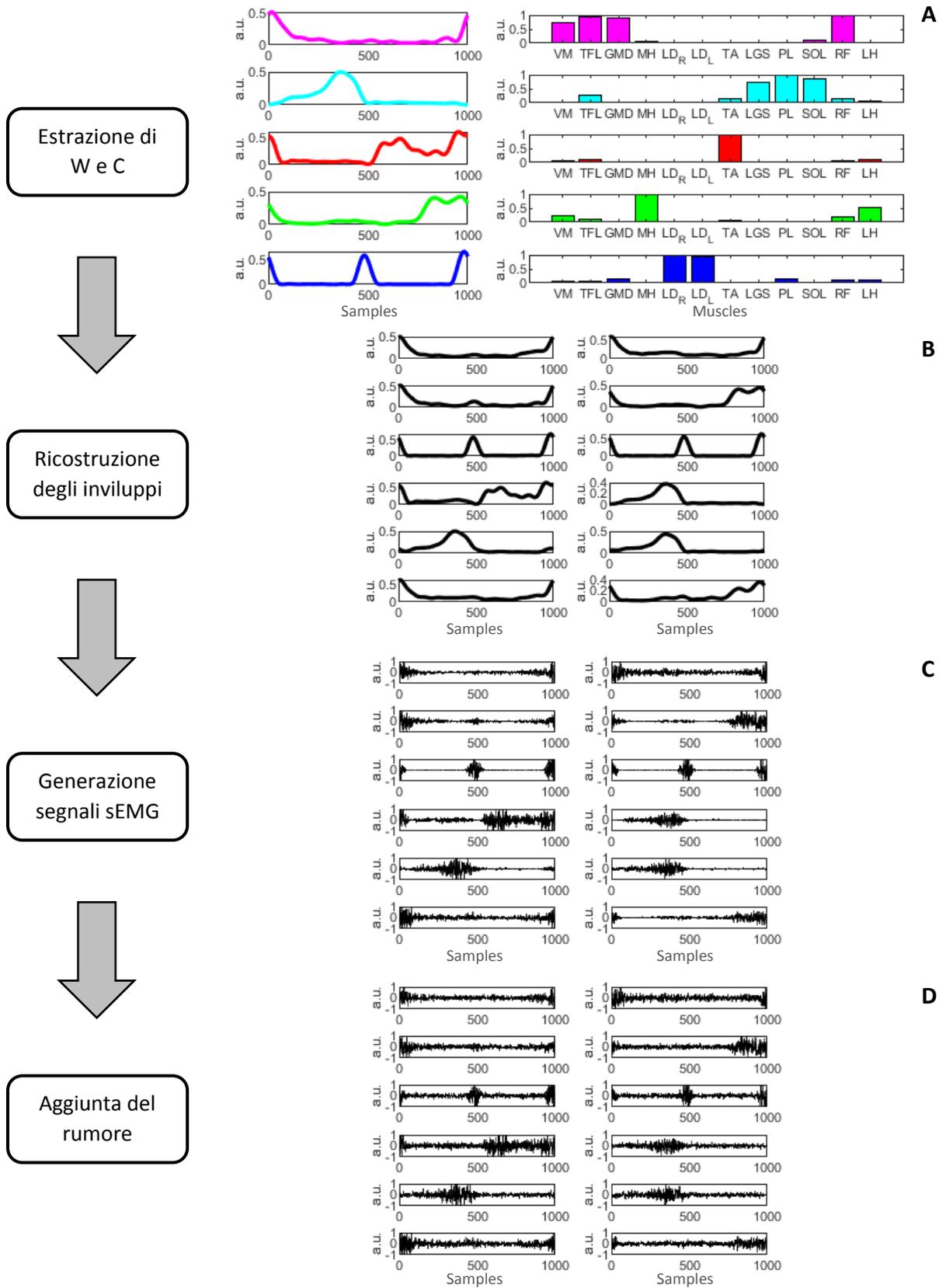


Figura 3: Schema della generazione di segnali sEMG sintetici. In questo esempio sono riportati (A) pesi muscolari e coefficienti di attivazione estratti, (B) involuپی ricostruiti, (C) sEMG sintetici e (D) sEMG con rumore (in questo esempio con SNR = 20 dB).

2.3 Pre-processing

I segnali sEMG raccolti dai dodici muscoli subiscono una fase di elaborazione prima di poter eseguire la fattorizzazione. Per prima cosa, vengono rimosse le parti di segnale corrispondenti alle inversioni ad U in prossimità delle pareti della stanza (presenti solo nel dataset reale), in modo da avere solamente i cicli di cammino corrispondenti ai tratti rettilinei. Successivamente vengono segmentati i vari cicli del cammino utilizzando il segnale basografico, che descrive il susseguirsi delle fasi HFPS (*Heel contact, Flat foot contact, Push off, Swing*) rispetto al tempo. Nel dettaglio, un ciclo di cammino può essere suddiviso in quattro fasi: appoggio del tallone e successiva accettazione del carico, contatto totale della pianta del piede al suolo, generazione della propulsione ed oscillazione della gamba [Fig. 4A]. Ogni ciclo di cammino segmentato, inizialmente campionato a 2000 Hz, viene ricampionato a 1000 campioni, per avere omogeneità dei dati (tutti i passi con lo stesso numero di campioni).

Successivamente, per poter eseguire la fattorizzazione, è necessario estrarre l'involuppo dei segnali sEMG [12], [13], [23], [24], ovvero la curva che ne descrive il profilo [Fig. 4B]. Per estrarre l'involuppo dai segnali, nel dettaglio, viene prima applicato un filtro IIR di tipo Butterworth passa-alto a 35 Hz di ordine 8, per rimuovere i contributi di rumore in bassa frequenza; al filtraggio seguono la rimozione della media e la rettificazione (o valore assoluto), per avere un segnale interamente sul semiasse positivo; infine viene utilizzato un filtro IIR di tipo Butterworth passa-basso a 12 Hz di ordine 5, il quale estrae il profilo del segnale.

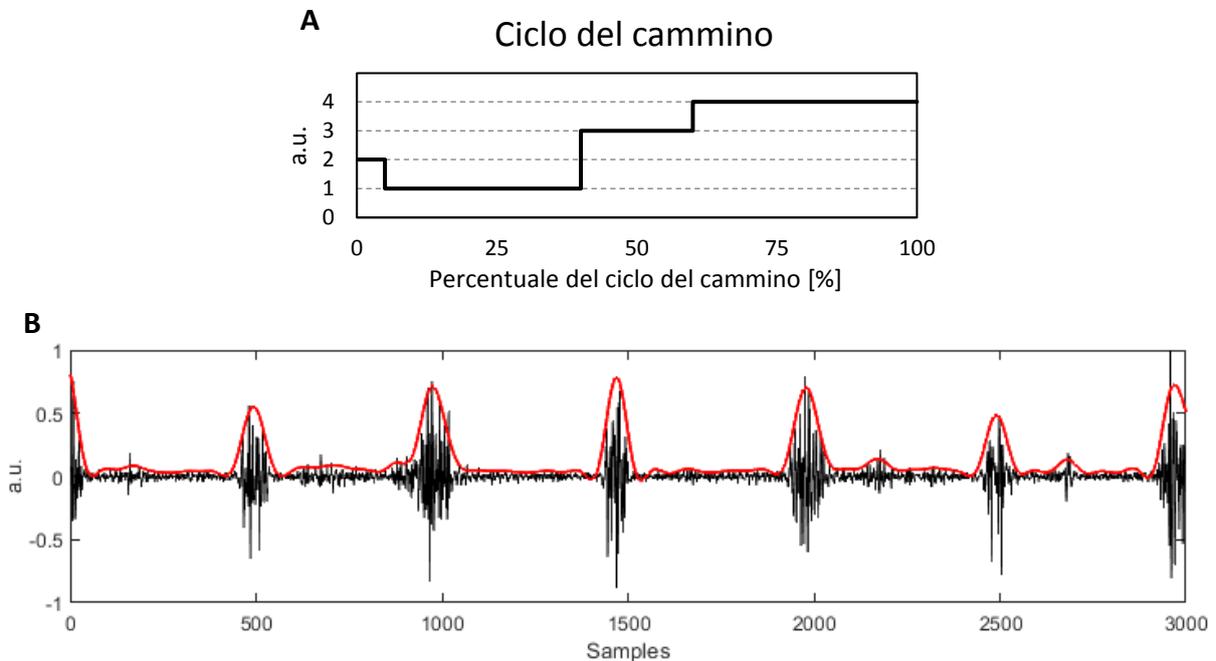


Figura 4: (A) Fasi del cammino normalizzate rispetto alla durata del ciclo di cammino. Partendo da sinistra vi sono le seguenti fasi: appoggio del tallone, contatto totale della pianta del piede al suolo, generazione della propulsione ed oscillazione della gamba. (B) Involuppo (rosso) di tre cicli di cammino provenienti dal segnale sEMG (nero) del muscolo dorsale lunghissimo destro (LD_R).

L'ampiezza di ogni involuppo viene poi normalizzata rispetto al valore massimo di ciascun muscolo, per uniformare i segnali provenienti dai vari muscoli: questo passaggio elimina le differenze in ampiezza tra i segnali provenienti dai vari muscoli, che possono essersi generate ad esempio a causa della differenza di spessore di tessuto adiposo e connettivo tra l'elettrodo ed il muscolo.

Infine, i cicli di cammino vengono concatenati tra di loro e suddivisi in sottogruppi: anziché applicare la fattorizzazione ai singoli cicli di cammino, viene applicata a gruppi di dieci cicli di cammino [12], [13], [23], [24]. La scelta di applicare la fattorizzazione a più cicli concatenati tra di loro permette di osservare al meglio la variabilità propria dell'attività muscolare [25].

Tutti i passi del *pre-processing* sono stati riuniti e schematizzati nella Figura 5.

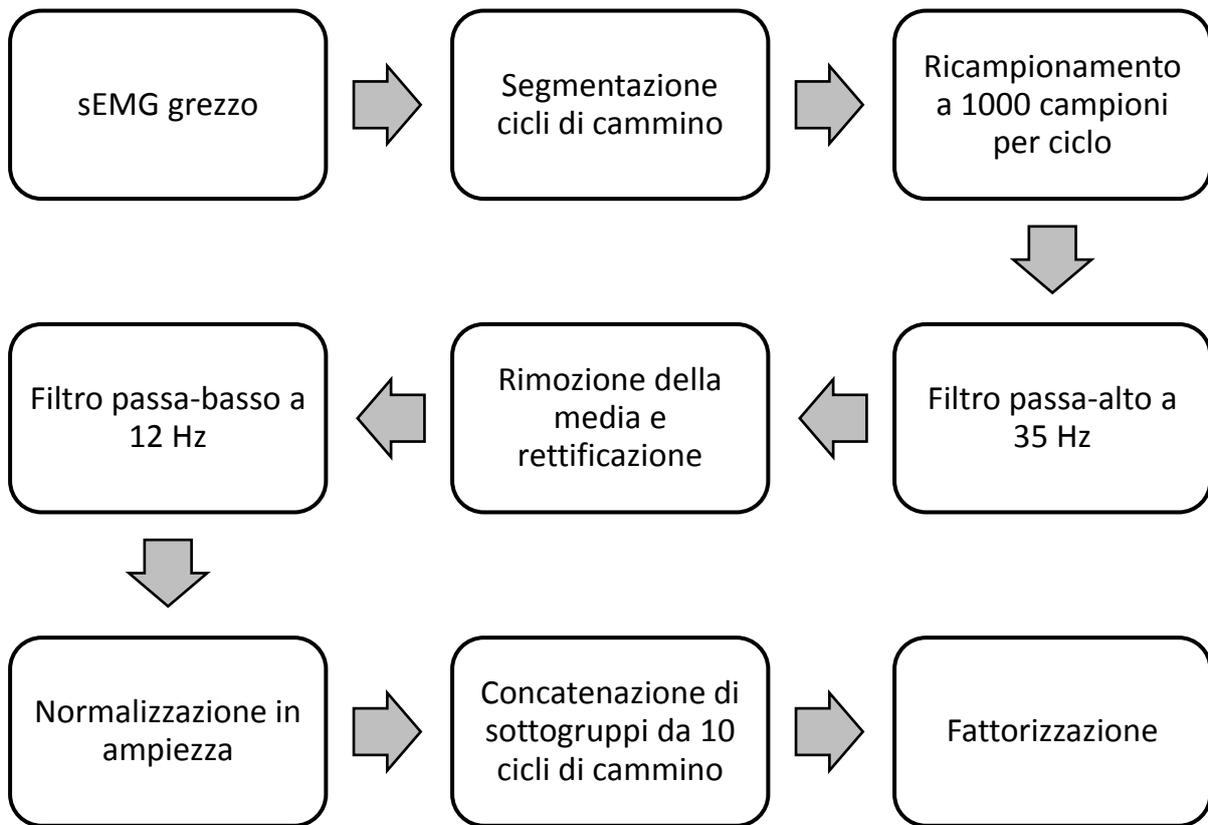


Figura 5: Schema a blocchi del *pre-processing* per preparare il segnale sEMG grezzo alla fattorizzazione.

2.4 Algoritmo di fattorizzazione

Esistono svariati algoritmi che permettono di compiere una fattorizzazione di un dataset iniziale. Tra tutti, quello più comunemente utilizzato per le sinergie muscolari è la *Non-negative Matrix Factorization* (NMF), poiché, oltre ad avere performance robuste e molto buone, è anche l'unica tecnica che non impone vincoli sulla distribuzione dei coefficienti di attivazione (ne impone solamente la non-negatività) [14]. Alla base della NMF vi è la risoluzione del problema:

$$M \cong W * C \quad W, C \geq 0 \quad (4)$$

dove M è la matrice contenente i segnali sEMG originari, W è la matrice tempo-invariante dei pesi muscolari e C è la matrice tempo-variante dei coefficienti di attivazione. Scegliendo un numero di sinergie muscolari pari ad n e avendo una matrice M di dimensioni $m \times s$ (dove m è il numero di muscoli, 12 in questo lavoro, e s è il numero di campioni, 10000 in questo lavoro), si scompone la matrice nelle due componenti W e C , di dimensioni rispettivamente $m \times n$ e $n \times s$ [Fig. 6]. Affinché si abbia un'effettiva riduzione della dimensionalità, n deve essere scelto minore di m .

Il metodo base e più usato in letteratura per la soluzione di questo problema è stato proposto da Lee e Seung [26]. Essi infatti introducono una regola di aggiornamento moltiplicativa (*multiplicative update rule*), che, rispetto ad altre regole, fornisce un ottimo compromesso tra velocità e facilità di implementazione. Nel dettaglio, la regola da loro ideata consiste nell'aggiornare iterativamente le due matrici C e W :

$$C \leftarrow C \frac{W^T M}{W^T W C} \quad W \leftarrow W \frac{M H^T}{W H H^T} \quad (5)$$

partendo da due inizializzazioni random, finché non viene soddisfatto il requisito, ovvero che il quadrato della distanza euclidea:

$$\|M - WC\|^2 \quad (6)$$

sia minimizzato. Quindi l'algoritmo iterativo si ferma solo al raggiungimento di una delle seguenti tre condizioni: dopo aver raggiunto la convergenza (la distanza non si riduce ulteriormente), dopo aver raggiunto una soglia imposta sulla distanza minima accettabile, o dopo aver raggiunto un numero massimo di iterazioni impostato a priori.

Purtroppo, il limite principale di questo algoritmo è che, sebbene garantisca il raggiungimento di un minimo, non garantisce che il minimo raggiunto sia il minimo globale: infatti, la regola moltiplicativa permette di ridurre velocemente la distanza euclidea, ma non permette di aumentarla, rischiando di convergere appunto verso un minimo locale. Successivamente a questo algoritmo sono nate moltissime varianti utilizzabili nel campo biomedicale, tutte raccolte in un'unica libreria [27]. In questo lavoro metterò a confronto l'algoritmo base (NMF) [26] e la sua variante con Alternating Non-negative Least Squares (NMF/ANLS), proposta da Kim e Park [28] ed implementata da Li e Ngom [27]. Il codice Matlab® della regola di aggiornamento della variante NMF/ANLS è stata preso dalla libreria appena citata [27].

L'algoritmo utilizzato in questo lavoro è la NMF/ANLS, ma il risultato in termini di sinergie muscolari sarebbe stato il medesimo anche utilizzando la NMF semplice. I criteri utilizzati per la scelta dell'algoritmo sono:

- Risultato finale, ovvero entrambi gli algoritmi devono restituire la medesima fattorizzazione;
- Tempo computazionale, che dipende sia dall'algoritmo che dalle repliche richieste (per raggiungere la condizione di ottimo globale anziché locale);
- Ripetibilità dell'output, ovvero quanto varia il risultato ripetendo la stessa operazione.

Per poter valutare il risultato della fattorizzazione è stata utilizzata la correlazione, utilizzando la metrica *cosine similarity*, applicata ai pesi muscolari e ai coefficienti di attivazione ricavati con i due diversi algoritmi. Quindi, la valutazione consiste nel verificare che i risultati finali degli algoritmi NMF e NMF/ANLS siano uguali (elevata correlazione).

L'esecuzione di più repliche dell'algoritmo è un'operazione necessaria per garantire statisticamente il raggiungimento della condizione di minimo globale, in quanto eseguendo più repliche è possibile conservare la fattorizzazione con l'errore di ricostruzione minore e scartare le altre repliche, in cui potrebbero esserci errori. Per quanto riguarda il numero di repliche richieste, per la NMF è stato utilizzato 50, come suggerito dalla letteratura [12]. Invece, per individuare il numero di repliche necessarie per l'algoritmo NMF/ANLS è stata utilizzata una variante del parametro ICC descritto successivamente [Eq. 7-9], in cui l'operazione di media nell'equazione 9 è stata sostituita con la selezione del valore minimo: questa lieve modifica permette di individuare eventuali *outlier* tra i sottogruppi causati da una fattorizzazione errata (minimo locale e non globale); con l'operazione di media un eventuale *outlier* verrebbe mascherato, per questo è stato preso il valore di correlazione minimo (corrispondente all'eventuale *outlier*).

Infine, per valutare la ripetibilità dell'output, ogni algoritmo è stato ripetuto per 5 volte, e successivamente è stata calcolata la correlazione (*cosine similarity*) tra i cinque output, per ognuno dei due algoritmi. È importante non confondere le repliche con le ripetizioni, in quanto le repliche sono più fattorizzazioni tra cui viene selezionata la migliore da conservare, mentre con ripetizioni si intende l'esecuzione dell'intero codice più volte, per poi osservare le eventuali differenze nei risultati ottenuti.

Per valutare i parametri appena descritti è stata usata una sottopopolazione di 24 soggetti presi dal dataset reale. Non è stato utilizzato l'intero dataset reale poiché sono test che richiedono sia molto tempo computazionale che molta memoria per immagazzinare i dati, perciò si è optato per un dataset ridotto, ma comunque abbastanza grande da garantire una buona validità delle prove effettuate. Il computer utilizzato ha processore Intel® Core™ i5-3570K 3.40 GHz e 8 GB di RAM, perciò i tempi computazionali che verranno riportati sono stati ricavati con queste caratteristiche di sistema.

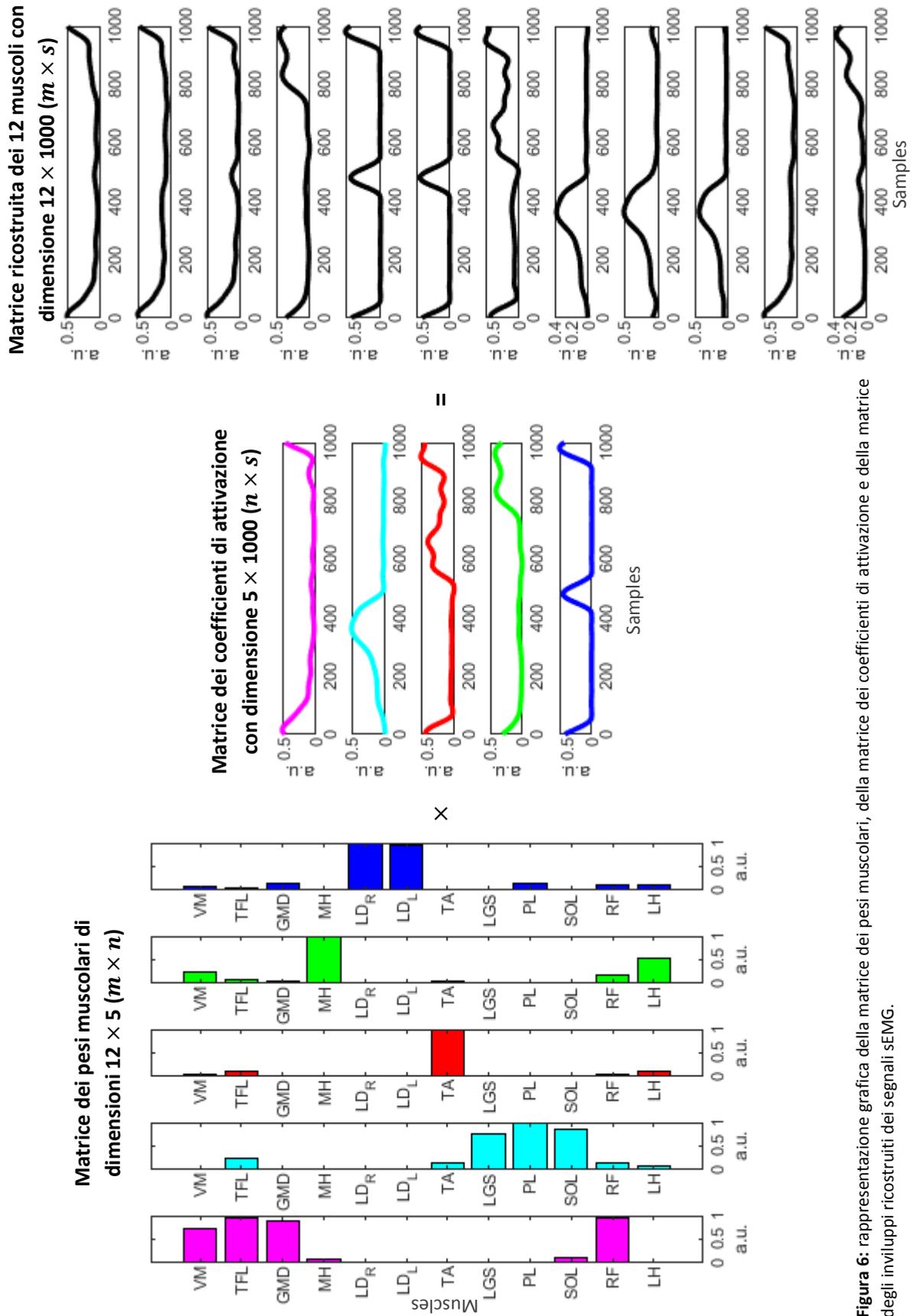


Figura 6: rappresentazione grafica della matrice dei pesi muscolari, della matrice dei coefficienti di attivazione e della matrice degli involuppi ricostruiti dei segnali sEMG.

2.5 Selezione del numero ottimale di sinergie muscolari

L'algoritmo alternativo per la scelta del numero ottimale di sinergie muscolari (algoritmo ChoOSyn, *algorithm for Choosing the Optimal number of Synergies*) si basa su due ipotesi fondamentali:

- Consistenza delle sinergie muscolari tra i vari cicli dello stesso *task* motorio;
- Bassa similarità tra le sinergie muscolari, per evitare informazioni ridondanti e superflue.

A questo scopo sono stati ricavati tre parametri numerici dalle sinergie muscolari, in grado di individuare la situazione ottimale che soddisfa queste condizioni.

Per calcolare i parametri seguenti, poiché la fattorizzazione viene applicata a più cicli concatenati (sottogruppi), i coefficienti di attivazione sono stati suddivisi nei dieci cicli di cammino e mediati tra di loro, per ottenere coefficienti di attivazioni relativi ad un singolo ciclo di cammino.

2.5.1 Correlazione *intra-cluster* (*Intra-Cluster Correlation*)

La correlazione *intra-cluster* (ICC) è il parametro che quantifica la consistenza delle sinergie muscolari [Fig. 7], e assume un valore tanto più prossimo a 1 quanto più la sinergia in esame si ripete simile a se stessa nei vari sottogruppi.

Per calcolare il parametro ICC per i pesi muscolari, contenuti nella matrice W di dimensione $m \times n$, si raccolgono tutti i W_i relativi a tutti gli i -esimi sottogruppi. Successivamente, si raggruppano tra di loro tutti i $w_{i,j}$ appartenenti ad una stessa sinergia j :

$$w_{i,j} = W_i(:,j) \quad 1 \leq j \leq n \quad (7)$$

Più semplicemente, lo stesso peso muscolare appartenente alla stessa sinergia viene preso da tutti i sottogruppi, per poter quindi valutare la sua variabilità tra i vari cicli dello stesso *task*.

Dopo aver quindi raggruppato lo stesso peso muscolare j di ogni sottogruppo i , è possibile ottenere la correlazione tra gli elementi selezionati, calcolandone prima il centroide c_j e successivamente la distanza dal centroide, utilizzando come metrica la *cosine similarity*:

$$ICC_{j,tot} = \frac{w_j^T c_j}{\sqrt{\text{diag}(w_j^T w_j) * c_j^T c_j}} \quad (8)$$

Le correlazioni di ogni elemento rispetto al centroide vengono poi mediate tra loro, per ottenere un unico valore ICC_j relativo ad una specifica sinergia muscolare j . Infine, viene preso il valore minimo tra tutti gli ICC_j :

$$ICC_j = \text{mean}(ICC_{j,tot}) \rightarrow ICC = \min(ICC_j) \quad (9)$$

Prendendo il minimo si tiene conto della situazione peggiore: se ad esempio su cinque sinergie muscolari quattro sono stabili ed una no, il valore finale ICC corrisponderà al valore di ICC_j dell'unica sinergia muscolare con alta variabilità tra i vari sottogruppi.

Per quanto riguarda l'applicazione delle formule ai coefficienti di attivazione, i passaggi sono i medesimi. Per distinguere il parametro ICC , nominerò $ICCW$ quello applicato ai pesi muscolari e $ICCC$ quello applicato ai coefficienti di attivazione.

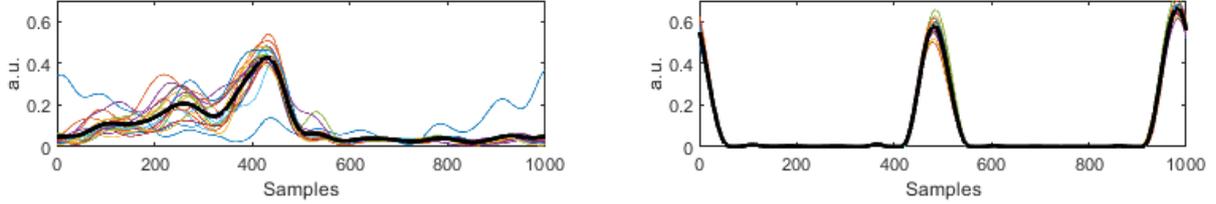


Figura 7: Confronto tra i coefficienti di attivazione di una sinergia muscolare con un valore di ICC prossimo ad 1 (destra) e di una sinergia il cui valore di ICC è molto diverso da 1 (sinistra). In nero è rappresentato l'andamento medio su tutti i sottogruppi, mentre con i colori sono rappresentati i coefficienti di attivazione di ogni singolo sottogruppo.

2.5.2 Similarità dei pesi muscolari (*Weight Similarity*)

La similarità dei pesi muscolari (WS) è il parametro che valuta quanto i pesi muscolari appartenenti a sinergie differenti siano correlati tra loro [Fig. 8], e assume un valore tanto più prossimo a 1 quanto più i pesi muscolari sono uguali tra di loro. Avere due sinergie muscolari con i pesi muscolari identici è una situazione da evitare, poiché dovrebbe formare un'unica sinergia muscolare (fenomeno dello *split*, avviene quando la numerosità delle sinergie muscolari è superiore a quella ottimale). Più in generale, è preferibile una situazione in cui le sinergie muscolari siano ben definite, e si tende a scartare una numerosità in cui vi è elevata similarità tra i pesi muscolari appartenenti a sinergie differenti: un esempio può essere quello riportato nella Figura 8, in cui aumentando eccessivamente la numerosità di sinergie muscolari si crea la situazione appena descritta (immagine a destra, caso a sei sinergie muscolari, muscoli evidenziati in grigio); inoltre, in questo caso si può anche che ridurre la numerosità di sinergie muscolari può nascondere una sinergia muscolare, distribuendola in modo non corretto.

Per poter calcolare il parametro WS , è prima necessario mediare tutti i pesi muscolari W_i dei vari sottogruppi, ottenendo un W medio del *task* motorio di dimensioni $m \times n$ (con m numero di muscoli e n numero di sinergie muscolari). Successivamente, per ogni coppia degli n pesi muscolari w_i :

$$w_i = W(:, i) \quad 1 \leq i \leq n \quad (10)$$

si calcola la correlazione, utilizzando come metrica la *cosine similarity*:

$$WS_{tot} = \frac{w_i w_j^T}{\sqrt{w_i w_i^T * w_j w_j^T}} \quad 1 \leq i \leq n ; 1 \leq j \leq n ; i \neq j \quad (11)$$

Dopo aver calcolato la correlazione per ogni coppia possibile, viene preso il valore massimo:

$$WS = \max(WS_{tot}) \quad (12)$$

in modo da tenere in considerazione la situazione peggiore, ovvero i due pesi muscolari più simili tra di loro all'interno della stessa numerosità di sinergie.

2. Materiali e Metodi

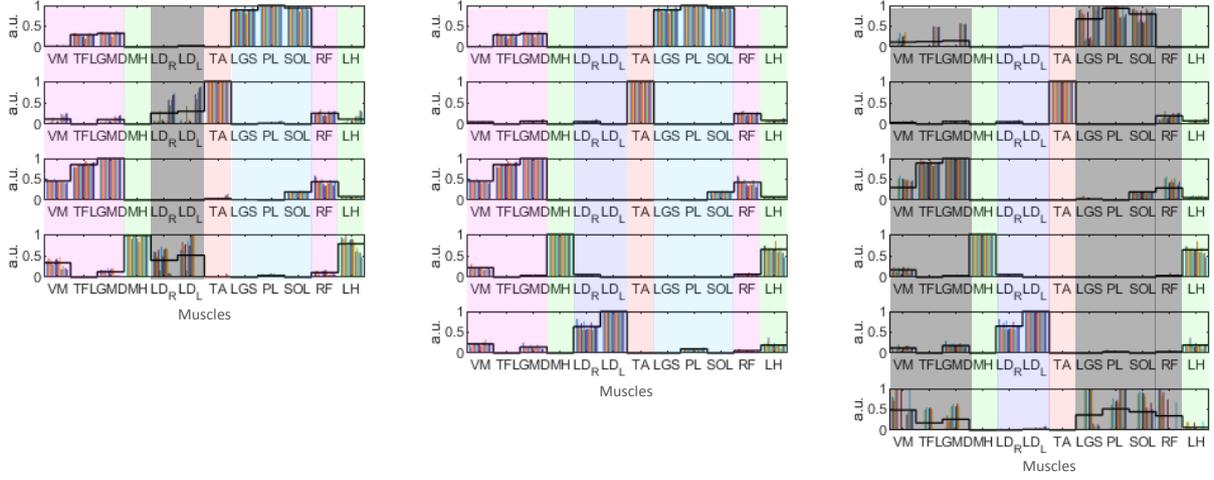


Figura 8: Confronto tra i pesi muscolari a diverse numerosità (4, 5 e 6) di sinergie muscolari. Con i diversi colori sono state evidenziati i muscoli appartenenti ad una stessa sinergia muscolare, mentre in grigio è stata evidenziato il caso di muscoli distribuiti in modo non ottimale.

2.5.3 Diversità dei coefficienti di attivazione (*Coefficient Dissimilarity*)

La diversità dei coefficienti di attivazione (CD) è il parametro che quantifica quanta nuova informazione aggiunge la nuova sinergia [Fig. 9], e assume un valore tanto più prossimo ad 0 quanto più il nuovo coefficiente di attivazione è uguale ad uno degli altri coefficienti di attivazione. Si ipotizza che la situazione in cui due sinergie muscolari con due gruppi di muscoli differenti che però si attivano negli stessi istanti temporali sia da evitare, poiché quelle due sinergie dovrebbero formare un'unica sinergia muscolare, e non due.

Per poter calcolare il parametro CD, è prima necessario mediare tutti i coefficienti di attivazione C_i dei vari sottogruppi, ottenendo un C medio del *task* motorio di dimensioni $n \times s$ (con n numero di sinergie muscolari e s numero di campioni dei segnali). Successivamente, utilizzando come centroidi i coefficienti di attivazione della sinergia muscolare precedente ($n - 1$), si applica l'algoritmo *k-means* per individuare quali sono i due coefficienti di attivazione c_1 e c_2 nati dalla stessa sinergia. Dopo averli individuati, si calcola la correlazione tra di essi, utilizzando come metrica la *cosine similarity*:

$$CD = 1 - \frac{c_1 c_2^T}{\sqrt{c_1 c_1^T * c_2 c_2^T}} \quad (13)$$

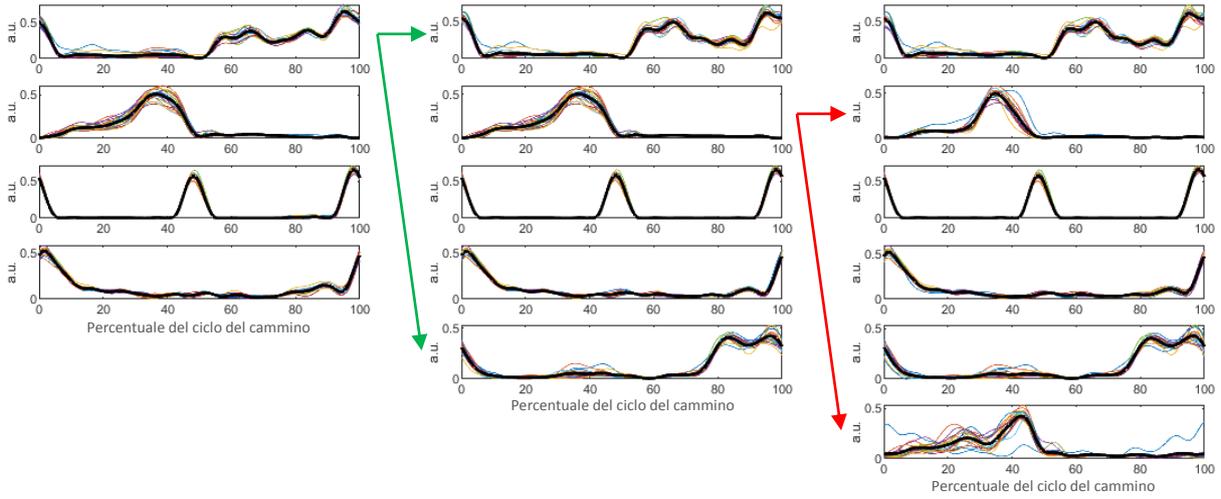


Figura 9: Esempio dei coefficienti di attivazione a numerosità diverse (4, 5 e 6) di sinergie muscolari. Ogni volta che la numerosità viene aumentata, si genera una nuova sinergia muscolare (posizionata al fondo di ogni colonna). Le frecce verdi indicano uno *split* corretto, che genera due coefficienti di attivazione diversi, mentre le frecce rosse indicano uno *split* che ha come risultato informazione ridondante.

2.5.4 Algoritmo ChoOSyn

L'algoritmo ChoOSyn (*Algorithm for Choosing the Optimal number of Synergies*) sfrutta i parametri appena descritti per individuare la numerosità ottimale di sinergie muscolari, leggendo le informazioni contenute nelle curve dei parametri ed osservando gli effetti dell'aumento o della riduzione della numerosità di sinergie muscolari. In questo algoritmo, i parametri vengono elaborati e combinati tra di loro [Fig. 10], generando due curve con le quali verrà scelto il numero ottimale di sinergie muscolari. Le due curve rappresentano rispettivamente lo stato dei pesi muscolari e dei coefficienti di attivazione, e vengono costruite nel seguente modo:

$$Indice_W(n) = WS(n) + \Delta ICCW(n) = WS(n) + (ICCW(1) - ICCW(n)) \quad (14)$$

$$Indice_C(n) = CD(n) - \Delta ICCC(n) = CD(n) - (ICCC(1) - ICCC(n)) \quad (15)$$

dove n è la numerosità di sinergie (che in questo lavoro varia da 1 a 8) e ΔICC rappresenta la differenza di ogni elemento del parametro ICC rispetto al valore che assume con una numerosità di sinergie pari ad 1. Nella figura 11 è riportato l'esempio grafico di come si ottiene il parametro finale nel caso dei coefficienti di attivazione [Eq. 15].

Perciò, i parametri WS e CD sono quelli che attivamente servono per individuare la numerosità ottimale di sinergie muscolari, tramite ad esempio variazioni brusche nel profilo delle curve, mentre il parametro ICC pesa ogni numerosità di sinergie muscolari in modo differente, rendendo meno probabile scegliere sinergie meno consistenti. In questo modo è possibile individuare la situazione ottimale in cui le sinergie muscolari sono sia consistenti e poco variabili tra i vari sottogruppi, sia con pesi muscolari e coefficienti di attivazione poco correlati tra di loro (e quindi senza informazioni ridondanti).

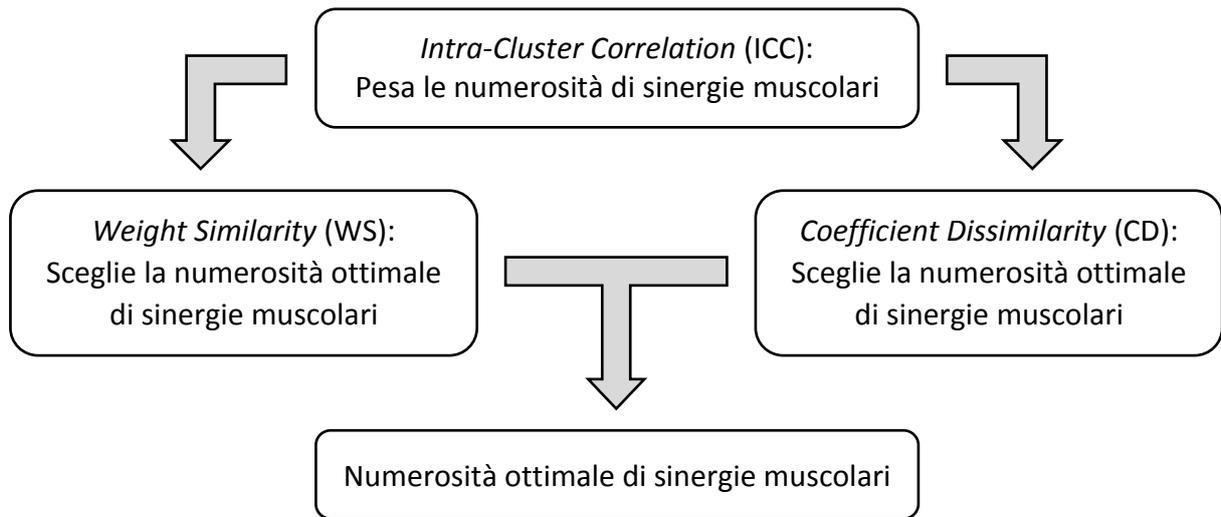


Figura 10: Schema dell'interazione tra i parametri per selezionare la numerosità ottimale di sinergie muscolari.

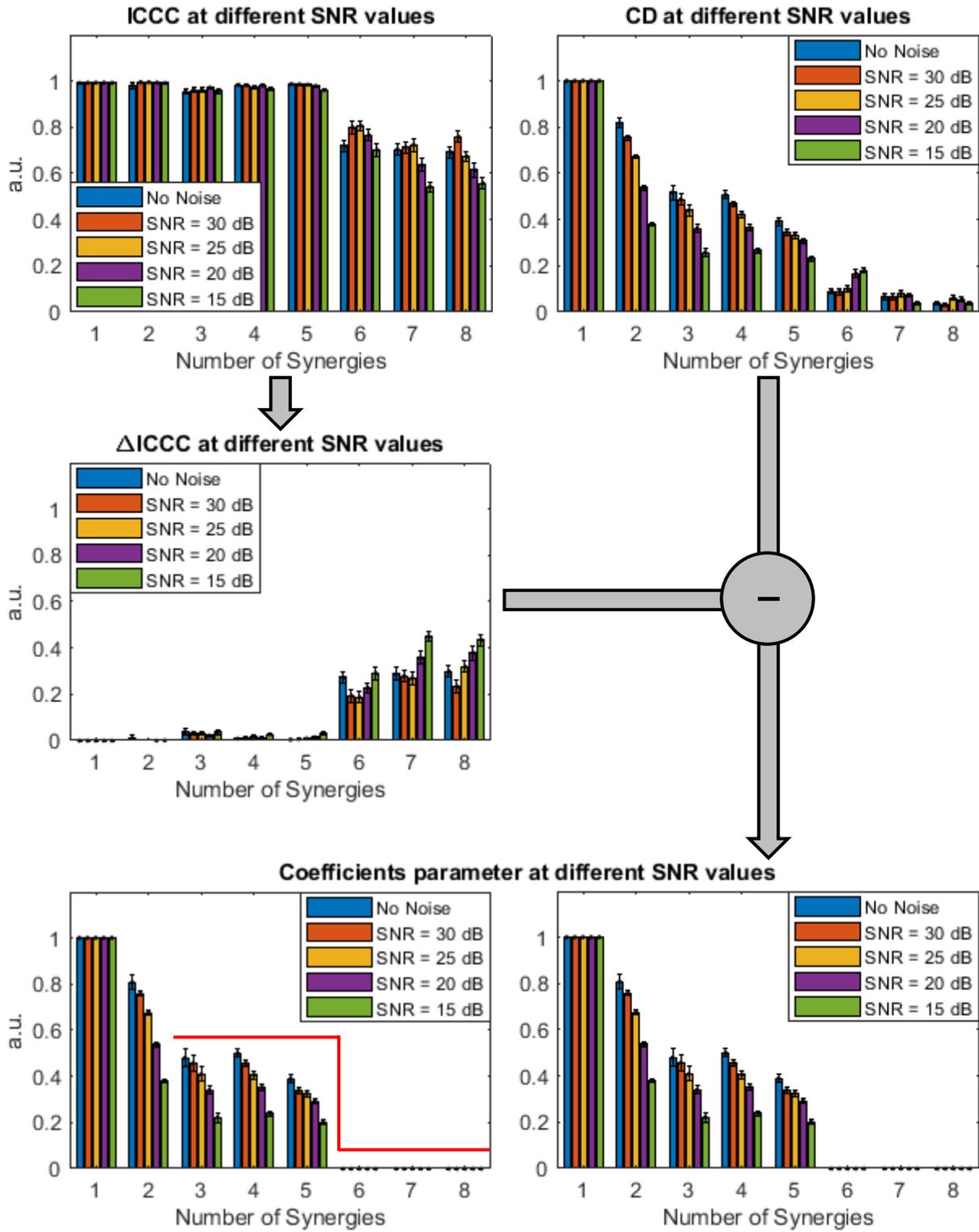


Figura 11: Combinazione della coppia di parametri ICC e CD per ottenere uno dei due parametri finali dell'algoritmo ChoOSyn (parametro dei coefficienti di attivazione) nel caso del dataset simulato da 5 sinergie muscolari. Ogni barra rappresenta *media \pm errore standard* dei valori dell'intera popolazione contenente 25 gruppi di segnali sEMG sintetici.

2.6 Valutazione delle performance

Le performance dell'algoritmo ChoOSyn sono state confrontate con altri criteri utilizzati in letteratura. Infatti, esistono più criteri di scelta basati sulla curva VAF [29], e per il confronto sono stati utilizzati i tre metodi più utilizzati:

- T-VAF (*Threshold VAF*): è il criterio più utilizzato in letteratura, si basa sulla semplice imposizione di una soglia (comunemente posta al 90%) e la scelta avviene selezionando la prima numerosità di sinergie muscolari che supera la suddetta soglia;
- E-VAF (*Elbow VAF*) [14]: sfrutta la curvatura per individuare il "gomito" ("*elbow*") della curva VAF, ovvero il punto in cui la pendenza varia maggiormente, ed è l'unico metodo basato sulla curva VAF a non utilizzare una soglia arbitraria;
- P-VAF (*Plateau VAF*) [19]: metodo basato sulla ricerca del punto oltre il quale la curva VAF può essere considerata una retta, ed utilizza una soglia (10^{-2}) sull'errore quadratico medio (MSE) del *fitting* dei punti della curva con una retta.

Il parametro VAF quantifica l'accuratezza di ricostruzione del risultato della fattorizzazione, ed è stato calcolato nel seguente modo:

$$VAF = \left(1 - \frac{\sum_{i=1}^n (M_i - R_i)^2}{\sum_{i=1}^n M_i^2} \right) * 100 \quad (16)$$

Dove M è la matrice originale contenente i segnali sEMG, R è la matrice ricostruita con i pesi muscolari ed i coefficienti di attivazione e n è il numero di sinergie muscolari.

Al fine poter definire quando i vari criteri (ChoOSyn e VAF) individuano o meno il numero corretto di sinergie muscolari, per il dataset reale è stato creato un *ground truth* utilizzando il parere di due operatori esperti. Per evitare che il giudizio fosse polarizzato, i due operatori erano all'oscuro dei risultati degli algoritmi per la scelta della numerosità ottimale di sinergie muscolari: infatti, le loro selezioni sono state raccolte mostrando loro solamente le immagini dei pesi muscolari e dei coefficienti di attivazione alle diverse numerosità di sinergie muscolari, soggetto dopo soggetto. Alcuni esempi delle selezioni effettuate sono riportati nel materiale supplementare [SI: Fig. 21-24]. Per quanto riguarda invece il dataset simulato, la natura della sua origine garantisce oggettività nella selezione del numero corretto: essendo stati simulati partendo da un numero di sinergie imposto a priori, la numerosità corretta di sinergie muscolari è nota.

3. Risultati

3.1 Ottimizzazione dell’algoritmo di fattorizzazione

Come detto in precedenza, i due algoritmi confrontati sono la NMF semplice [26] e la NMF/ANLS [27], [28]. Confrontando gli output dei due algoritmi si riscontrano valori di correlazione prossimi ad 1 [Tab. 1], perciò entrambi i metodi forniscono il medesimo risultato: quindi, considerando il parametro della correlazione tra i due output, è indifferente utilizzare un algoritmo o l’altro, poiché entrambi forniscono le medesime sinergie muscolari.

Correlazione tra le fattorizzazioni – NMF vs. NMF/ANLS								
	Numero di sinergie muscolari:							
	1	2	3	4	5	6	7	8
Pesi muscolari (W)	1.000 ± 0.000	1.000 ± 0.000	0.999 ± 0.006	0.999 ± 0.001	0.999 ± 0.002	0.994 ± 0.020	0.984 ± 0.063	0.970 ± 0.044
Coefficienti di attivazione (C)	1.000 ± 0.000	1.000 ± 0.000	0.999 ± 0.006	1.000 ± 0.001	1.000 ± 0.001	0.997 ± 0.006	0.993 ± 0.026	0.989 ± 0.017

Tabella 1: Confronto tra gli output dei due algoritmi in esame (NMF vs. NMF/ANLS). Sono mostrati i valori delle correlazioni (con metrica *cosine similarity*) tra i pesi muscolari ed i coefficienti di attivazione ricavati con i due algoritmi. È stato riportato *media ± deviazione standard* del dataset contenente 24 soggetti reali sani.

La fattorizzazione risulta essere più variabile nella NMF semplice, poiché la sua regola di aggiornamento non permette né l’allontanamento dal minimo, rischiando di raggiungere un minimo locale, né garantisce il raggiungimento del minimo, in quanto le variazioni ad ogni iterazione sono piccole. Per assicurarsi di raggiungere sia il minimo assoluto, sia un risultato della fattorizzazione meno variabile, è necessario utilizzare un alto numero di repliche dell’algoritmo (in questo lavoro 50, come suggerito dalla letteratura [12]). Invece, la NMF/ANLS, che garantisce sempre il raggiungimento del minimo [27], [28] (la fattorizzazione non termina mai per il raggiungimento del limite di iterazioni, a differenza della NMF), ha intrinsecamente un output più stabile già ad un numero di repliche più basso: infatti, è stato osservato che già con 3 repliche dell’algoritmo NMF/ANLS le sinergie muscolari erano prive di *outlier*, e perciò in questo lavoro è stato scelto di replicare la fattorizzazione 5 volte.

Come si può osservare dalle due tabelle [Tab. 2, 3], entrambi gli algoritmi possiedono un output altamente ripetibile: infatti, i valori di correlazione sono uguali o prossimi ad 1, e le deviazioni standard sono estremamente basse. Questo risultato è necessario, poiché significa che eseguendo l’intero codice più volte il risultato non varia.

Ripetibilità della fattorizzazione – NMF								
	Numero di sinergie muscolari:							
	1	2	3	4	5	6	7	8
Pesi muscolari (W)	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.002	0.999 ± 0.009	0.997 ± 0.017
Coefficienti di attivazione (C)	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.003	0.998 ± 0.010	0.997 ± 0.017

Tabella 2: Test di ripetibilità della fattorizzazione con NMF. Sono mostrati i valori delle correlazioni (con metrica *cosine similarity*) tra i pesi muscolari (medi) ed i coefficienti di attivazione (medi) rispetto al proprio centroide, ricavati da 5 ripetizioni della fattorizzazione con l’algoritmo NMF. È stato riportato *media ± deviazione standard* del dataset contenente 24 soggetti reali sani, ciascuno ripetuto per 5 volte.

3. Risultati

Ripetibilità della fattorizzazione – NMF/ANLS								
	Numero di sinergie muscolari:							
	1	2	3	4	5	6	7	8
Pesi muscolari (W)	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.006	0.999 ± 0.007	0.999 ± 0.008
Coefficienti di attivazione (C)	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.001	1.000 ± 0.000	1.000 ± 0.005	0.999 ± 0.010	0.999 ± 0.005

Tabella 3: Test di ripetibilità della fattorizzazione con NMF/ANLS. Sono mostrati i valori delle correlazioni (con metrica *cosine similarity*) tra pesi muscolari (medi) ed i coefficienti di attivazione (medi) rispetto al proprio centroide, ricavati da 5 ripetizioni della fattorizzazione con l’algoritmo NMF/ANLS. È stato riportato *media ± deviazione standard* del dataset contenente 24 soggetti reali sani, ciascuno ripetuto per 5 volte.

Il criterio che ha rappresentato un peso maggiore nella scelta è stato il tempo computazionale: l’algoritmo della NMF/ANLS richiede statisticamente meno repliche (5 anziché 50) ed è molto più rapido ad arrivare a convergenza [28], perciò globalmente porta ad un tempo computazionale notevolmente ridotto. Infatti, il tempo medio per fattorizzare un singolo sottogruppo (10 cicli di cammino) è notevolmente inferiore con la NMF/ANLS (con 5 repliche), che impiega 45.9 ± 10.5 s (*media ± deviazione standard*), a differenza dei 342.8 ± 33.9 s (*media ± deviazione standard*) richiesti dalla NMF (con 50 repliche): perciò mediamente la NMF/ANLS impiega l’87% del tempo in meno rispetto alla NMF. È importante notare che l’enorme differenza è frutto di un numero di repliche nettamente superiore nel caso della NMF. Invece, a parità di ripetizione (ovvero eseguendo la NMF con sole 5 repliche) si può notare che la regola di aggiornamento semplice della NMF è leggermente più veloce, impiegando 33.3 ± 3.6 s (*media ± deviazione standard*); ovviamente la NMF ha bisogno di un numero ben maggiore di repliche, perciò quest’ultima considerazione è stata fatta solo per osservare che la maggiore velocità della NMF/ANLS deriva esclusivamente dal ridotto numero di repliche. Infine, è bene tenere a mente che i tempi riportati sono relativi ad un solo sottogruppo di 10 cicli di cammino, mentre nel dataset reale utilizzato per ogni soggetto sono presenti 15-20 sottogruppi, e i soggetti sono 50: perciò il tempo guadagnato in questo caso è di circa 80 ore (considerando il solo dataset reale).

3.2 Parametri

In questo paragrafo sono mostrati i risultati relativi agli indicatori precedentemente descritti, necessari alla determinazione del numero ottimale di sinergie muscolari. Nei grafici a barre successivi, i due dataset di riferimento (simulato e reale) sono stati mantenuti separati, ed ogni barra verticale rappresenta il valore medio di una determinata popolazione interna al dataset (indicato nella legenda).

Nei grafici, il dataset simulato è stato mantenuto diviso nei tre gruppi simulati a 4, 5 e 6 sinergie muscolari, per mostrare al meglio il comportamento dei parametri nelle varie situazioni. Anche per il dataset reale è stato scelto di operare la stessa suddivisione per rappresentare i parametri graficamente: utilizzando il *ground truth* citato precedentemente, è stato possibile dividere il dataset reale in tre gruppi, rispettivamente con soggetti che, secondo gli operatori, esprimono 4, 5 o 6 sinergie muscolari.

3.2.1 Intra-Cluster Correlation (ICC)

La [figura 12](#) mostra l’andamento del parametro ICC applicato a pesi muscolari (ICCW) e coefficienti di attivazione (ICCC) nel dataset simulato. Come si può notare, in tutti i casi la consistenza tra i diversi

cicli di cammino cala bruscamente ($p < 10^{-4}$) dopo aver superato la numerosità corretta di sinergie muscolari; invece, per numerosità uguali o inferiori a quella corretta il parametro assume valori prossimi ad 1, indicando che le sinergie muscolari possiedono un aspetto delineato e ben definito al variare del sottogruppo in analisi. Inoltre, il parametro ICCW subisce una maggiore influenza del rumore, che a parità di numerosità di sinergie muscolari mostra un *trend* crescente all'aumentare del SNR, rendendo il gradino meno evidente. Il valore del *p-value* riportato (ricavato tramite *t-test*) è relativo al caso più rumoroso, ovvero la condizione peggiore, ed è ulteriormente minore negli altri casi a rumorosità inferiori.

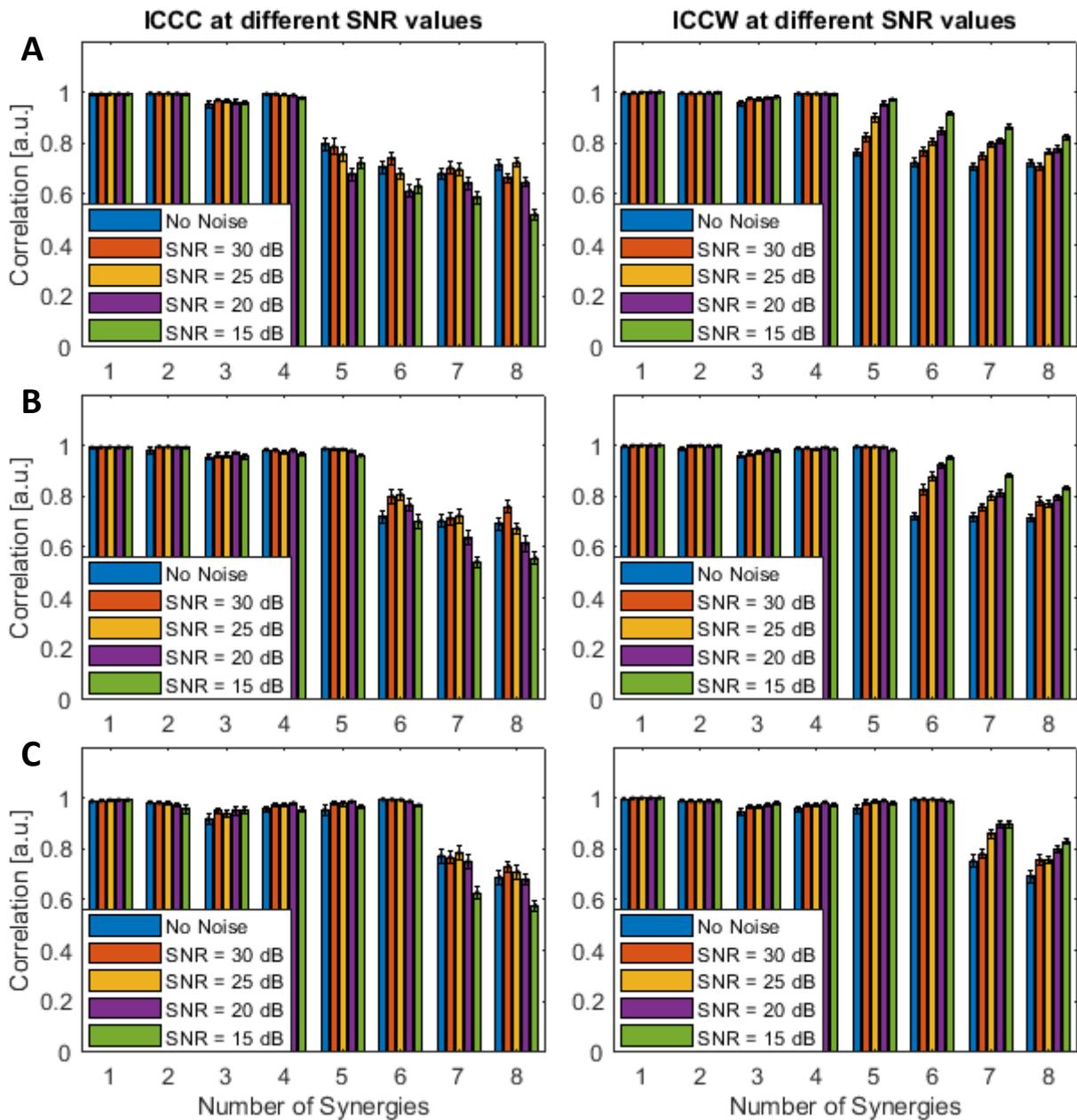


Figura 12: Parametro ICC (ICCC a sinistra e ICCW a destra) applicato al dataset simulato da (A) 4, (B) 5 e (C) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione contenente 25 gruppi di segnali sEMG simulati.

Invece, nel dataset reale, si può notare dalle [figure 13A e 14A](#) che il parametro ICC non presenta un gradino netto, poiché sono presenti soggetti che esprimono numerosità differenti di sinergie muscolari. Inoltre, emerge anche che non vi sono differenze importanti ($p > 0.1$) tra la popolazione sana e quella patologica, perciò si è scelto di considerare un unico gruppo composto sia da soggetti sani che patologici. Quindi l'intero dataset è stato suddiviso in soggetti che esprimono 4, 5 e 6 sinergie muscolari, indipendentemente dalla presenza o meno della patologia di Parkinson. Operando questa suddivisione [[Fig. 13B-D](#), [Fig. 14B-D](#)] è possibile osservare lo stesso comportamento che si è riscontrato nel dataset reale: infatti nelle curve è presente un gradino ($p < 0.005$) in corrispondenza del numero corretto (secondo il giudizio degli operatori esperti) di sinergie muscolari, in entrambi i parametri.

Infine, in tutti i casi nel parametro ICC è presente un *trend* decrescente, indicando che più la numerosità di sinergie muscolari è elevata, maggiore è la sua variabilità tra i vari sottogruppi di cicli di cammino.

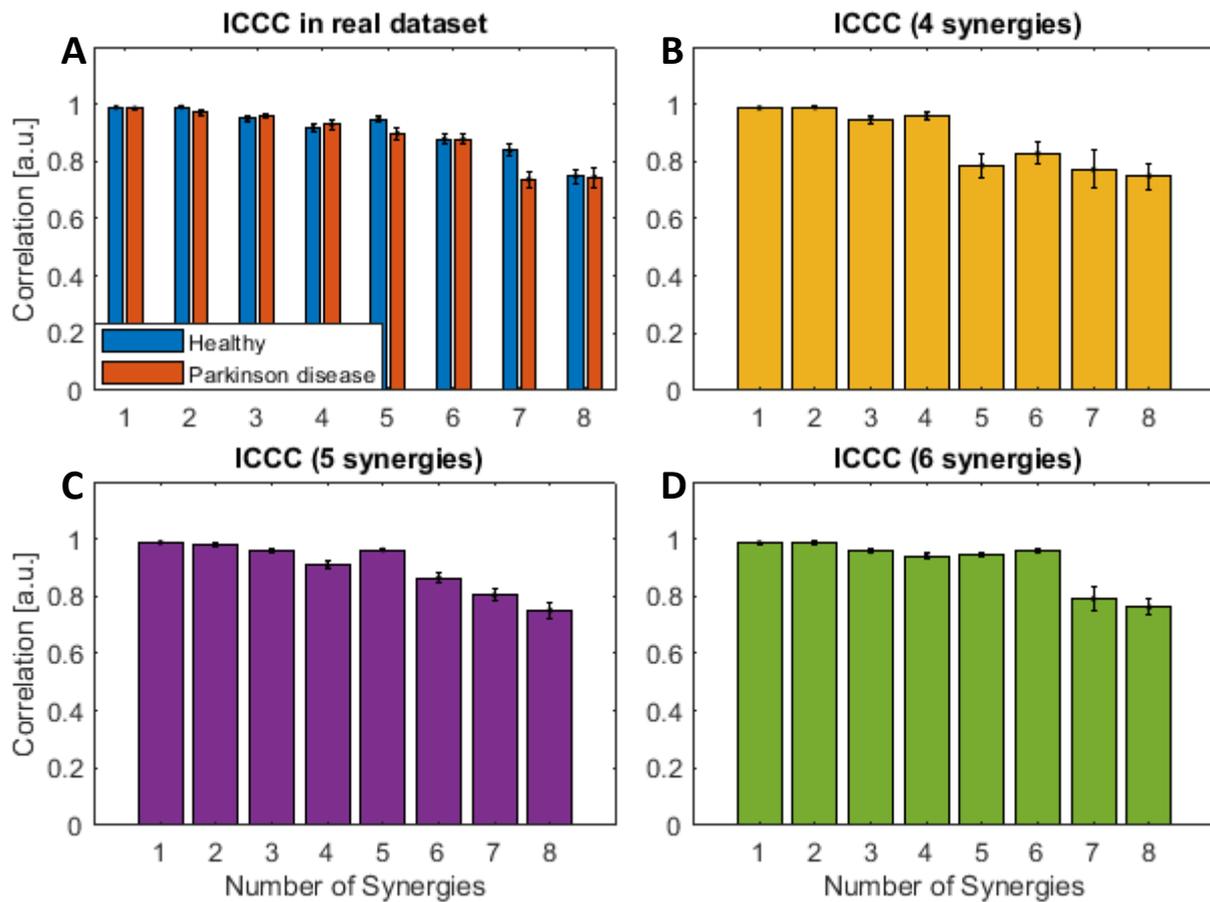


Figura 13: (A) Parametro ICC applicato al dataset reale completo. (B-D) Parametro ICC applicato ai soggetti che esprimono (B) 4, (C) 5 e (D) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione.

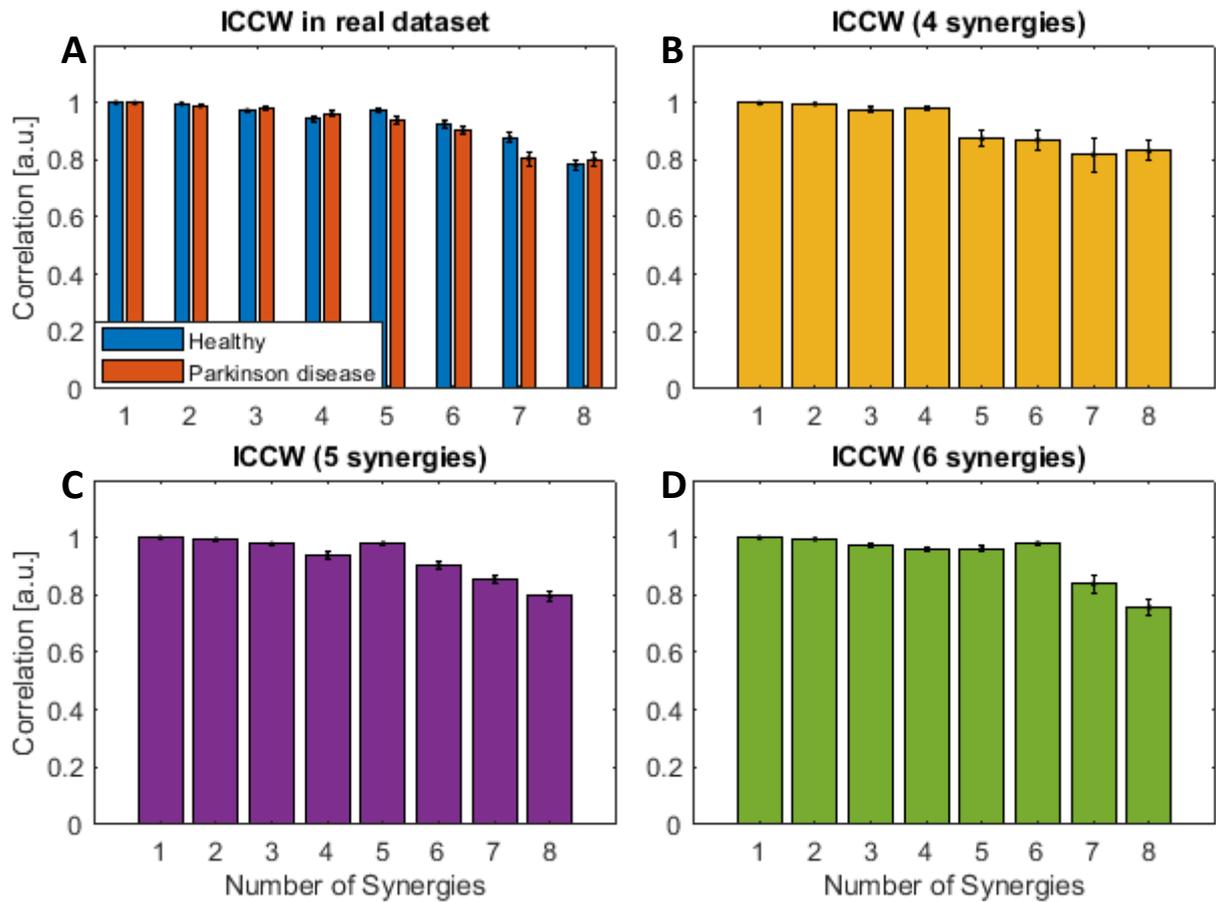


Figura 14: (A) Parametro ICCW applicato al dataset reale completo. (B-D) Parametro ICCW applicato ai soggetti che esprimono (B) 4, (C) 5 e (D) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione.

3.2.2 Weight Similarity (WS)

La figura 15 mostra, nel caso di dataset simulato, come il parametro WS abbia un andamento “a gradino”: a numerosità inferiori o uguali a quella corretta WS assume valori pressoché costanti a parità di SNR, invece per numerosità superiori a quella corretta, dopo un brusco incremento ($p < 10^{-8}$), la curva raggiunge un *plateau*. Anche in questo caso si può notare un *trend* osservando i valori del parametro al variare del SNR e con la numerosità di sinergie muscolari fissata, e anche qui il rumore attenua il gradino presente nel parametro. Il valore del *p-value* riportato (ricavato tramite *t-test*) è relativo al caso più rumoroso, ovvero la condizione peggiore, ed è ulteriormente minore negli altri casi a rumorosità inferiori.

Il valore di WS alla numerosità di sinergie pari ad 1 è stato posto a 0 in quanto un'unica sinergia muscolare non può essere correlata a nessun'altra, essendo l'unica presente.

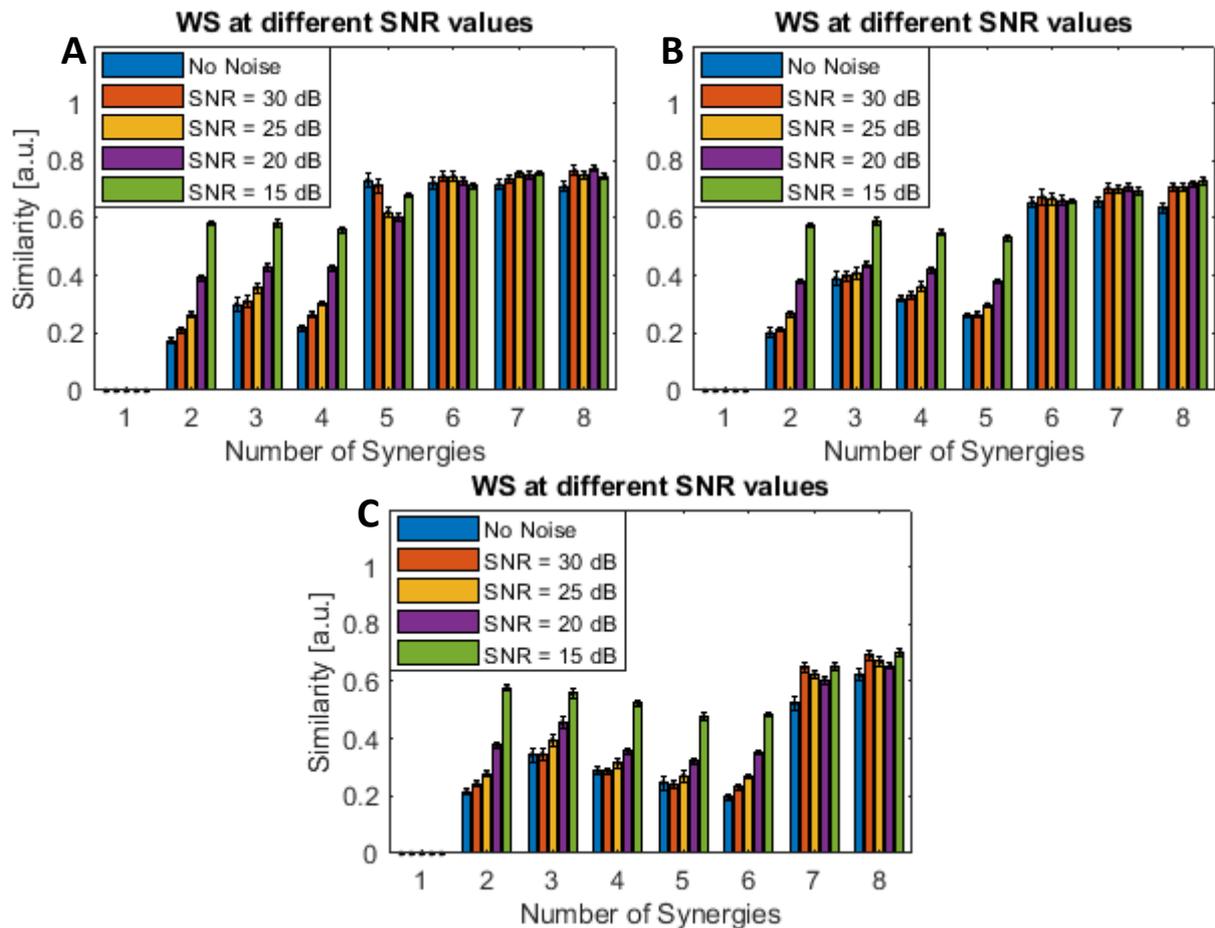


Figura 15: Parametro WS applicato al dataset simulato da (A) 4, (B) 5 e (C) 6 sinergie muscolari. Ogni barra rappresenta *media* \pm *errore standard* dei valori dell'intera popolazione contenente 25 gruppi di segnali sEMG simulati.

Invece, nel caso reale [Fig. 16A] anche WS ha un comportamento che non dipende ($p = 0.7$) dalla condizione del soggetto (sano o patologico), e presenta lo stesso gradino ($p < 0.01$) [Fig. 16B-D] riscontrato nel dataset simulato. Inoltre, è presente anche una lieve “conca” in corrispondenza della numerosità corretta (secondo gli operatori) di sinergie muscolari.

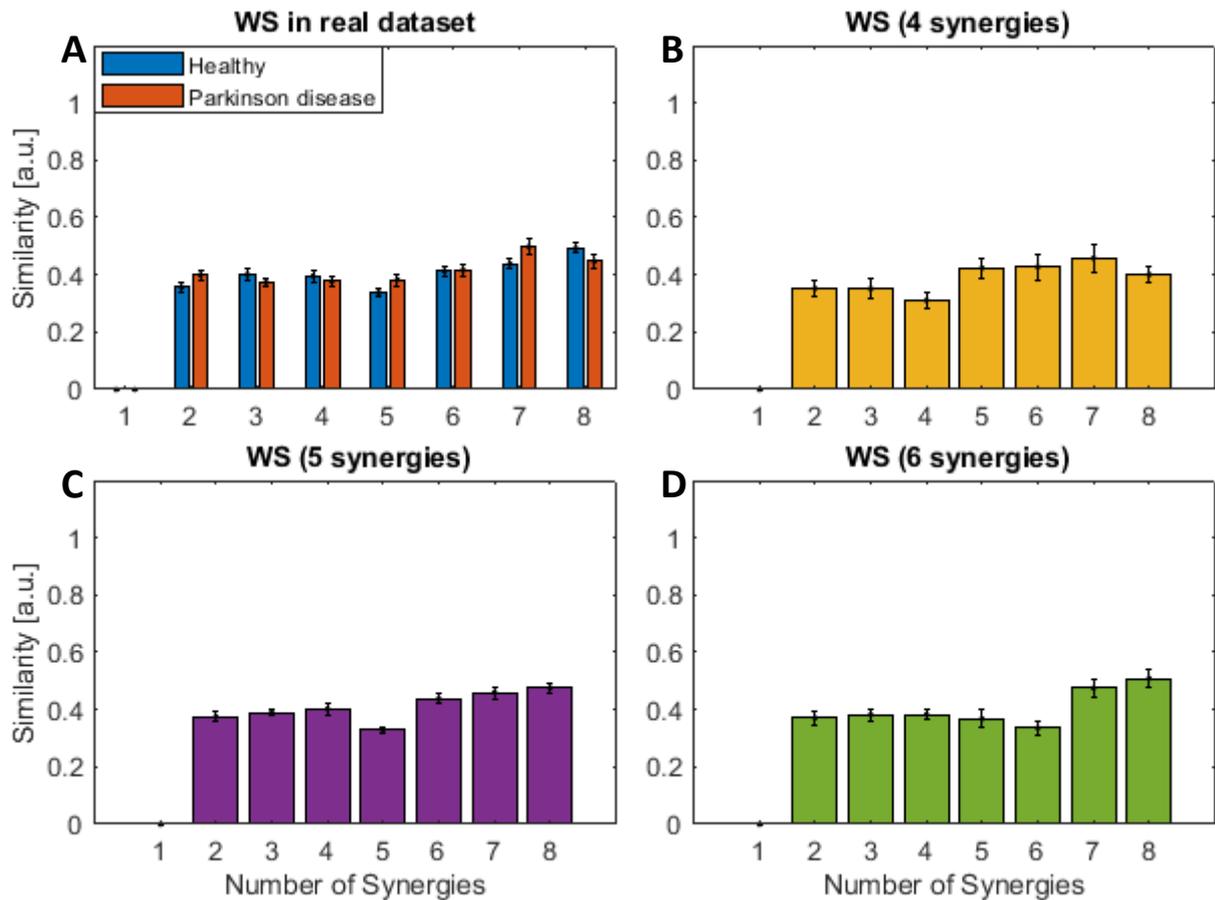


Figura 16: (A) Parametro WS applicato al dataset reale completo. (B-D) Parametro WS applicato ai soggetti che esprimono (B) 4, (C) 5 e (D) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione.

3.2.3 Coefficient Dissimilarity (CD)

La figura 17 mostra il parametro CD applicato al dataset sintetico. Si può notare che il parametro ha un andamento decrescente, ma per numerosità superiori a quella corretta vi è un gradino netto ($p < 10^{-4}$) oltre al quale la curva raggiunge una sorta di *plateau*. Esattamente come i casi precedenti, il rumore influisce sul valore del parametro, e nel dettaglio all'aumentare del SNR si riduce l'ampiezza del gradino appena citato. Il valore del p -value riportato (ricavato tramite t -test) è relativo al caso più rumoroso, ovvero la condizione peggiore, ed è ulteriormente minore negli altri casi a rumorosità inferiori.

Il valore di CD alla numerosità di sinergie pari ad 1 è stato posto a 1 in quanto la prima sinergia muscolare aggiunge il massimo dell'informazione possibile, essendo la prima.

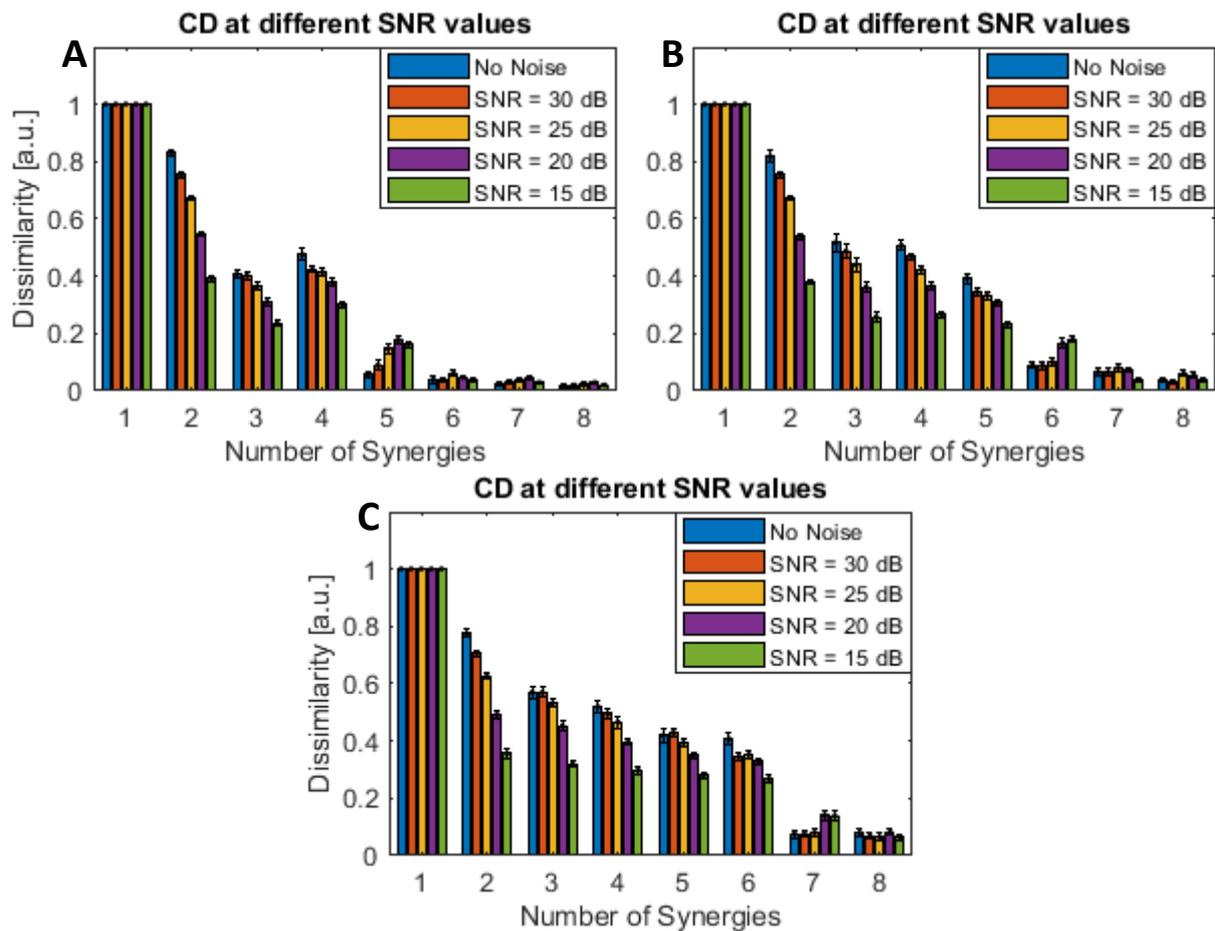


Figura 17: Parametro CD applicato al dataset simulato da (A) 4, (B) 5 e (C) 6 sinergie muscolari. Ogni barra rappresenta *media* \pm *errore standard* dei valori dell'intera popolazione contenente 25 gruppi di segnali sEMG simulati.

Nel caso reale [Fig. 18A] anche CD ha un comportamento che non dipende ($p = 0.7$) dalla condizione del soggetto (sano o patologico), e presenta lo stesso gradino ($p < 0.005$) [Fig. 18B-D] riscontrato nel dataset simulato.

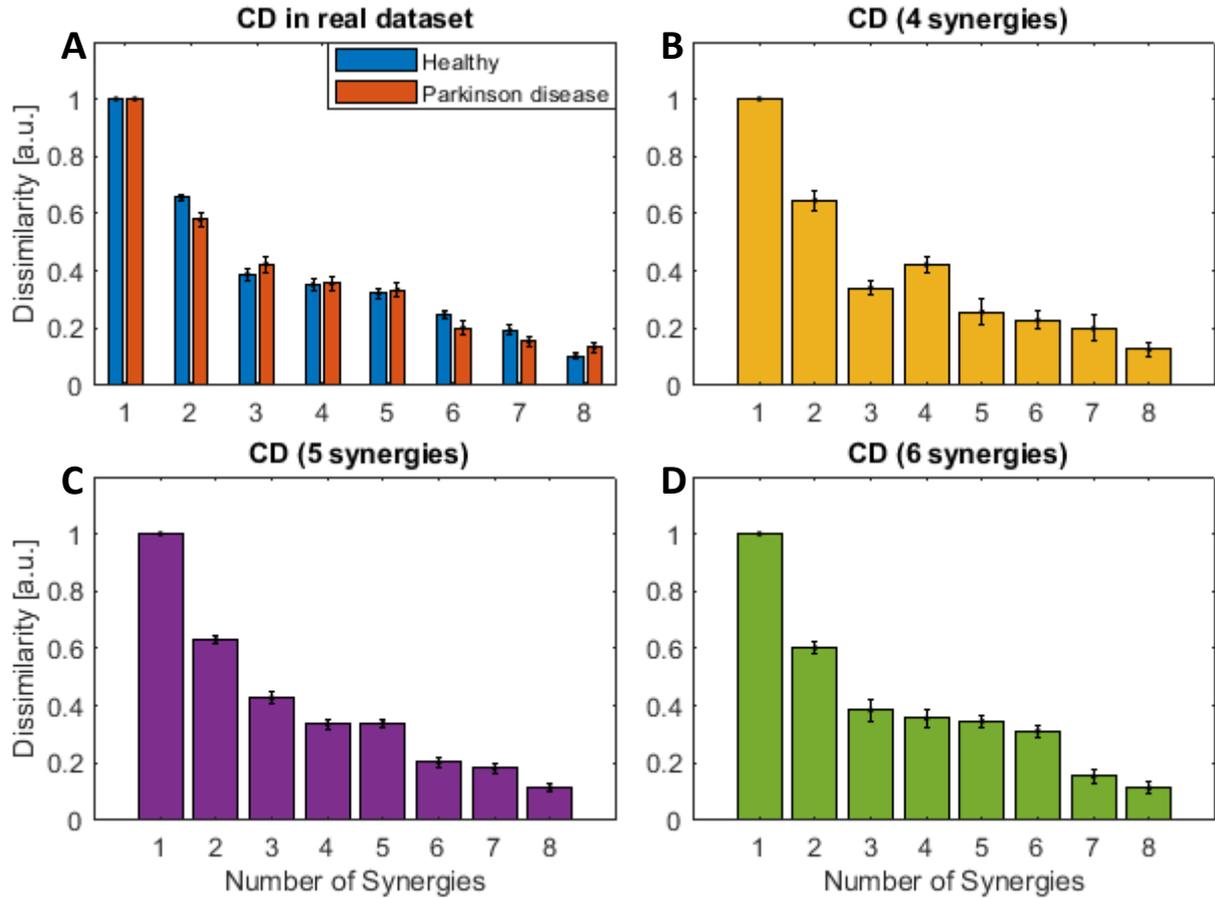


Figura 18: (A) Parametro CD applicato al dataset reale completo. (B-D) Parametro CD applicato ai soggetti che esprimono (B) 4, (C) 5 e (D) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera sottopopolazione.

3.3 Algoritmo ChoOSyn

Come descritto nella sezione dei metodi, l'algoritmo ChoOSyn non utilizza direttamente i parametri ICC, WS e CD, ma una loro combinazione [Fig. 10], in cui i parametri WS e CD contribuiscono alla selezione della numerosità di sinergie muscolari ottimale, mentre il parametro ICC pesa diversamente ogni numerosità per favorire situazioni più stabili e per sfavorirne altre più variabili. In seguito all'unione a coppie dei parametri si ottengono curve con caratteristiche tali da poter individuare al meglio la numerosità di sinergie muscolari ottimale; in questo paragrafo verranno quindi mostrate le curve dei due parametri appena citati e le performance dell'algoritmo che li sfrutta.

Nel dettaglio, l'algoritmo ChoOSyn ricerca nel parametro dei coefficienti di attivazione un gradino, mentre nel parametro dei pesi muscolari un gradino e/o una conca. Sono state scelte queste "maschere" poiché l'algoritmo è stato costruito sulla popolazione del dataset simulato da 5 sinergie muscolari, e queste sembrano essere le caratteristiche più discriminanti per quel determinato gruppo di segnali; successivamente, testando l'algoritmo anche sugli altri dataset, si è visto che il comportamento è il medesimo, indipendentemente da segnale reale o simulato, da sano o patologico.

3.3.1 Dataset simulato

Nella [tabella 4](#) sono riportate le performance dell'algoritmo ChoOSyn, che lavora sui parametri riportati nella [figura 19](#), e dei tre criteri basati sulla curva VAF utilizzati in letteratura. Tutti i metodi, ad eccezione del T-VAF, riescono ad individuare la numerosità corretta in quasi tutti i casi, e i pochissimi errori appartengono ai casi più rumorosi (SNR = 15 dB). Andando ad osservare l'aspetto dei pesi muscolari e dei coefficienti di attivazione di questi casi errati, si è notata la presenza di una sinergia muscolare aggiuntiva con coefficiente di attivazione piatto, come se la fattorizzazione avesse riconosciuto il rumore elevato (presente in tutti i muscoli) come una sinergia muscolare a sé.

Invece, per quanto riguarda il metodo con soglia al 90% del VAF (T-VAF), esso non è in grado di individuare la numerosità corretta in praticamente nessun caso poiché, essendo segnali sEMG simulati, la NMF riesce a fattorizzare il segnale con una buona accuratezza già a numerosità più basse, ottenendo una curva VAF traslata verso l'alto: con la soglia al 90% vengono selezionate in tutti i casi numerosità di sinergie muscolari inferiori a quella corretta. Per questo motivo, è stata testata anche la soglia al 95% della curva VAF (poiché in letteratura 90% è la più gettonata, ma come già detto varia da lavoro a lavoro tra 80% e 95%), ottenendo un lieve miglioramento delle performance nel caso a 4 sinergie muscolari.

Confronto tra i metodi di scelta del numero ottimale di sinergie muscolari – Dataset sintetico

		Metodo				
		T-VAF (90%)	T-VAF (95%)	E-VAF	P-VAF	ChoOSyn
4 sinergie	Nessun Rumore	0/25	24/25	25/25	25/25	25/25
	SNR = 30 dB	0/25	22/25	25/25	25/25	25/25
	SNR = 25 dB	0/25	18/25	25/25	25/25	25/25
	SNR = 20 dB	0/25	9/25	25/25	25/25	25/25
	SNR = 15 dB	2/25	0/25	24/25	25/25	24/25
	Totale	2/125	73/125	124/125	125/125	124/125
5 sinergie	Nessun Rumore	0/25	11/25	25/25	25/25	25/25
	SNR = 30 dB	0/25	2/25	25/25	25/25	25/25
	SNR = 25 dB	0/25	0/25	25/25	25/25	25/25
	SNR = 20 dB	0/25	0/25	24/25	22/25	25/25
	SNR = 15 dB	0/25	0/25	18/25	23/25	19/25
	Totale	0/125	13/125	117/125	120/125	119/125
6 sinergie	Nessun Rumore	0/25	1/25	25/25	25/25	25/25
	SNR = 30 dB	0/25	0/25	25/25	25/25	24/25
	SNR = 25 dB	0/25	0/25	25/25	25/25	22/25
	SNR = 20 dB	0/25	0/25	25/25	25/25	25/25
	SNR = 15 dB	0/25	0/25	23/25	25/25	24/25
	Totale	0/125	1/25	123/125	125/125	120/125
Totale dataset simulato		2/375	87/375	364/375	370/375	363/375

Tabella 4: Performance dei vari metodi di selezione del numero di sinergie muscolari nel dataset simulato. È riportato per ogni criterio il numero di volte in cui individua la corretta numerosità di sinergie muscolari per ognuno dei tre gruppi, generati rispettivamente da 4, 5 e 6 sinergie muscolari.

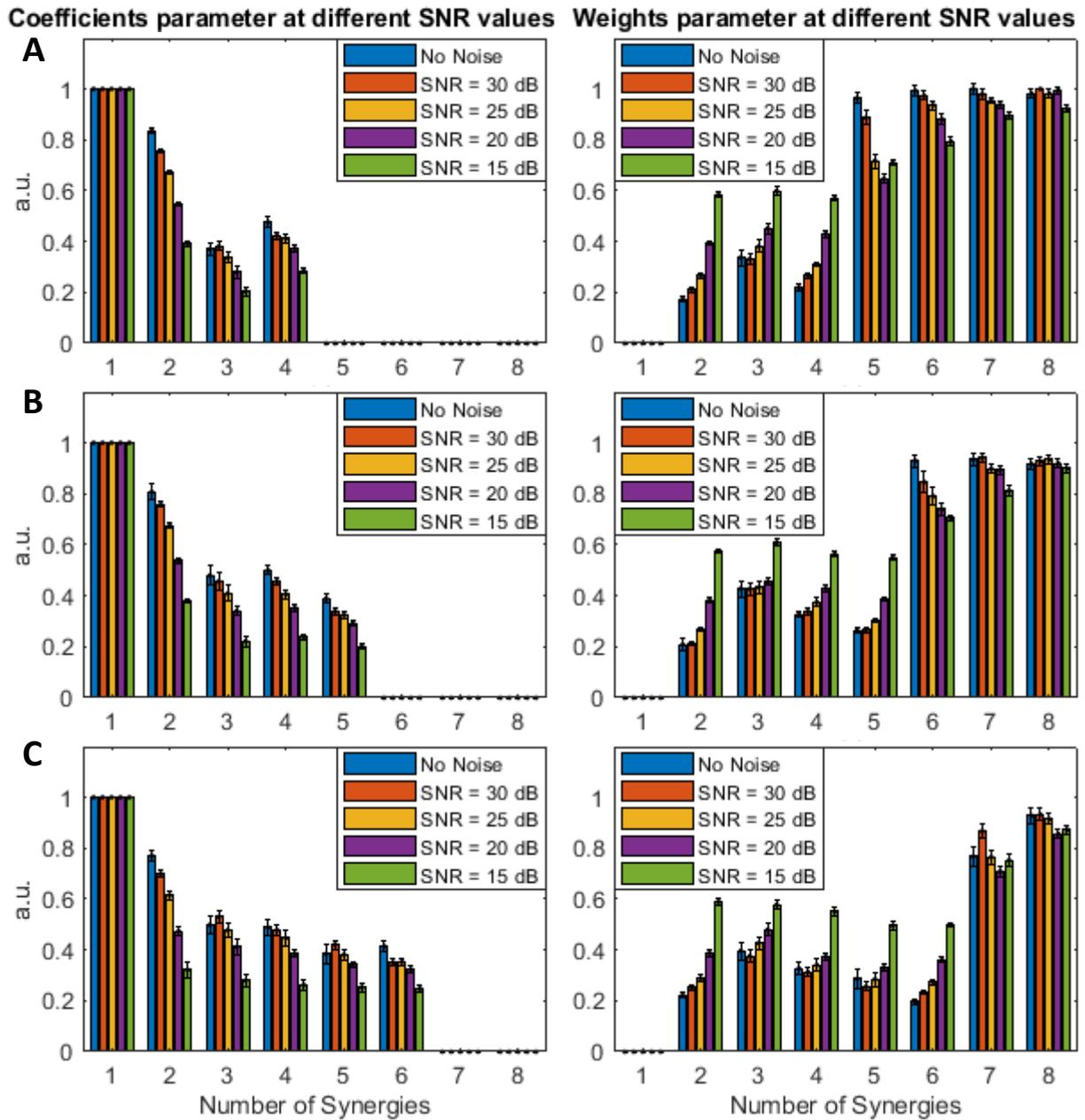


Figura 19: Parametri finali dell'algorithm ChoOSyn (parametro dei coefficienti di attivazione a sinistra e parametro dei pesi muscolari a destra) applicato al dataset simulato da (A) 4, (B) 5 e (C) 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione contenente 25 gruppi di segnali sEMG simulati.

3.3.2 Dataset reale

Nella [tabella 5](#) sono riportate le selezioni delle numerosità di sinergie muscolari da parte dei due operatori esperti, utilizzate come *ground truth* per la valutazione delle performance con il dataset reale. Si può notare che, esattamente come ci si aspetta da letteratura per il *task* motorio in analisi, la maggior parte dei soggetti esprime 5 sinergie muscolari.

<i>Ground truth</i> – Dataset reale				
	4 sinergie	5 sinergie	6 sinergie	Media
Soggetti sani	2/30	20/30	8/30	5.20
Soggetti patologici	5/20	12/20	3/20	4.90
Totale	7/50	32/50	11/50	5.08

Tabella 5: Selezione delle numerosità di sinergie muscolari corrette nel dataset reale da parte dei due operatori esperti. Sono riportati i numeri dei soggetti che esprimono (da sinistra verso destra) 4, 5 e 6 sinergie muscolari.

Utilizzando il *ground truth* ricavato, si ottengono le performance riportate nella [tabella 6](#). Mentre l'algoritmo ChoOSyn individua correttamente la numerosità in molti dei soggetti, nei tre metodi basati sulla curva VAF le performance sono basse. Nel dettaglio, nel caso del criterio T-VAF tutti gli errori sono imputabili ad una ricostruzione accurata dei segnali sEMG già con numerosità di sinergie muscolari minori, infatti in tutti i casi errati il T-VAF ha individuato come corretta una numerosità di sinergie muscolari inferiore di 1 o 2 rispetto a quella selezionata dagli operatori. Nel caso dei criteri E-VAF e P-VAF si riscontra una situazione opposta: il gomito ed il *plateau* vengono spesso raggiunti a numerosità elevate, e perciò questi metodi indicano nei casi errati una numerosità di sinergie muscolari superiore di 1 o 2 rispetto a quella selezionata dagli operatori. Infine, anche per l'algoritmo ChoOSyn tutti gli errori sono in eccesso rispetto alla numerosità selezionata dagli operatori, ed i grafici relativi ai parametri sono riportati nella [figura 20](#).

Confronto tra i metodi di scelta del numero di sinergie muscolari – Dataset reale				
	Metodo			
	T-VAF	E-VAF	P-VAF	ChoOSyn
Soggetti sani	11/30	19/30	10/30	25/30
Soggetti patologici	6/20	12/20	7/20	17/20
Totale	17/50	31/50	17/50	42/50

Tabella 6: Performance dei vari metodi di selezione del numero di sinergie muscolari nel dataset reale. È riportato per ogni criterio il numero di volte in cui individua la corretta numerosità di sinergie muscolari, utilizzando come riferimento il *ground truth* dei due operatori esperti.

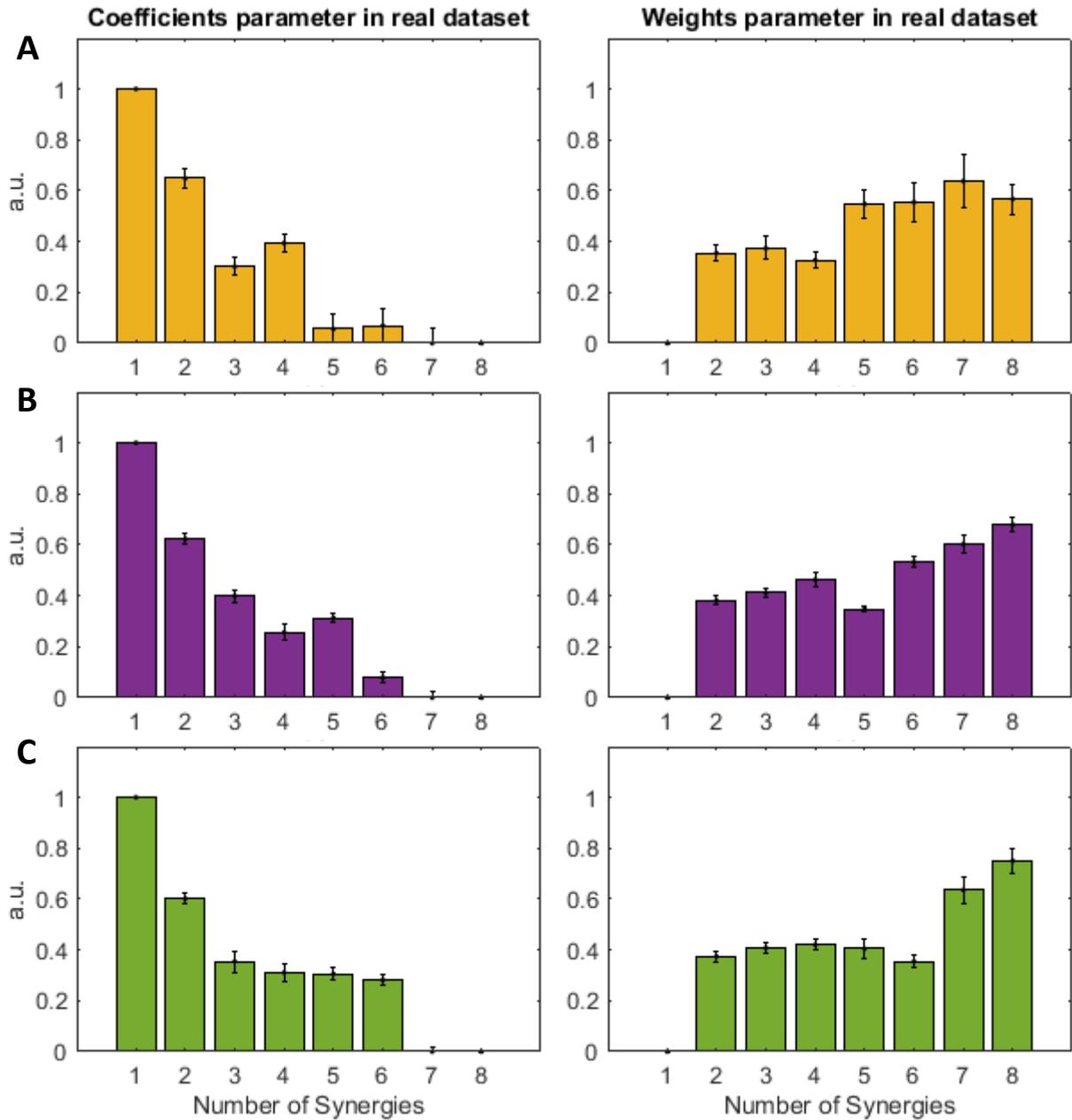


Figura 20: Parametri finali dell'algorithm ChoOSyn (parametro dei coefficienti di attivazione a sinistra e parametro dei pesi muscolari a destra) applicato al dataset reale suddiviso in soggetti che esprimono **(A)** 4, **(B)** 5 e **(C)** 6 sinergie muscolari. Ogni barra rappresenta *media ± errore standard* dei valori dell'intera popolazione.

4. Discussioni

4.1 Conclusioni

La selezione automatica del numero ottimale di sinergie muscolari è tutt'ora un problema irrisolto. Il metodo attualmente più utilizzato in letteratura consiste nell'utilizzo di una soglia arbitraria, comunemente posta al 90% [1]–[3], [6], [11], [15]–[17], ma variabile dall'80% al 95% [18] della curva del parametro VAF (*Variance Accounted For*) (metodo T-VAF, *Threshold VAF*), il quale quantifica la qualità di ricostruzione delle sinergie muscolari. Il problema principale è che il VAF non fornisce informazioni sul significato delle sinergie muscolari, ottenendo molte volte sinergie muscolari non ottimali. Infatti, in vari studi, si utilizza il VAF solo in modo indicativo, per poi selezionare manualmente la numerosità che si ritiene corretta [30] o utilizzando una numerosità media per tutto il dataset [11], [31]. Perciò, l'obiettivo di questo lavoro di tesi è il superamento dei metodi che utilizzano la curva VAF, tramite un algoritmo che effettui scelte solide, basate su consistenza e similarità dei pesi muscolari e dei coefficienti di attivazione. L'algoritmo presentato in questo lavoro, ChoOSyn, soddisfa pienamente questo requisito, in quanto è completamente slegato dal VAF, ed utilizza una combinazione [Eq. 14, 15] di altri tre parametri: *Intra-Cluster Correlation* (ICC) [Fig. 12-14] [Eq. 7-9], *Weight Similarity* (WS) [Fig. 15, 16] [Eq. 10-12] e *Coefficient Dissimilarity* (CD) [Fig. 17, 18] [Eq. 13]. A differenza del VAF, le curve che si ottengono con questi tre parametri utilizzati presentano un netto cambio di pendenza in corrispondenza della numerosità di sinergie muscolari ottimale [Fig. 19, 20], evitando quindi l'utilizzo di soglie arbitrarie. Nonostante abbia questi vantaggi, un metodo viene giudicato principalmente per le sue performance, e l'algoritmo ChoOSyn riesce ad individuare la numerosità corretta in 42 soggetti reali su 50 e in 363 casi simulati su 375. Come visto precedentemente nei risultati [Tab. 4, 6], nei casi simulati anche i due criteri P-VAF e E-VAF riescono ad ottenere performance vicine al 100%, ma nei soggetti reali l'algoritmo ChoOSyn performa nettamente meglio rispetto ai metodi con VAF, con risultati corretti superiori di 25/50 rispetto a P-VAF e T-VAF e di 11/50 rispetto a E-VAF.

4.2 Limiti e Prospettive Future

Il principale limite dell'algoritmo ChoOSyn è che, nonostante le performance siano migliori rispetto al VAF, non raggiunge il 100% dei casi individuati correttamente. Osservando i grafici medi dei parametri [Fig. 19, 20] si nota che tutti presentano il cambio di pendenza in corrispondenza della numerosità corretta, suggerendo che forse il problema non sia tanto nei parametri utilizzati, ma nel metodo con cui vengono analizzate le curve. Infatti, nonostante i parametri medi abbiano tutti un andamento quasi ideale, se presi singolarmente in alcuni casi sono presenti situazioni dubbie [SI: Fig. 22, 24] che l'attuale algoritmo ChoOSyn non è ancora in grado di giudicare correttamente. Un futuro miglioramento può quindi essere l'ottimizzazione del codice dell'algoritmo ChoOSyn, in modo da interpretare al meglio le curve dei due parametri.

Un secondo limite può risiedere nella natura dei parametri utilizzati: infatti, nonostante siano stati ideati sull'osservazione delle sinergie muscolari e sulla conoscenza del loro significato fisiologico, essi restano comunque parametri matematici, frutto di formule applicate ai valori numerici di pesi muscolari e coefficienti di attivazione. L'algoritmo ChoOSyn possiede ottime performance, ma uno studio futuro potrebbe affiancargli un secondo metodo che indaghi le sinergie muscolari conoscendo ad esempio quali muscoli operano insieme, o, più in generale, conoscendo la fisiologia dietro al *task*

motorio in esame. Questo ipotetico supporto permetterebbe di operare scelte ancor più solide, poiché giustificate da conoscenze fisiologiche dei muscoli, e perciò di coniugare l'ottimo dal punto di vista matematico con l'ottimo secondo ciò che l'operatore si aspetta di vedere dalle sinergie muscolari. Nonostante questa modifica possa portare ad un vantaggio, renderebbe la selezione mirata e specifica per il *task* motorio e i muscoli in esame.

Infine, è bene evidenziare che, oltre ad essere indipendente dal VAF e ad avere performance migliori, l'algoritmo ChoOSyn è ipoteticamente utilizzabile con ogni *task* motorio, e perciò la prospettiva futura più interessante è sicuramente l'applicazione di questo metodo ad altri dataset con *task* motori e muscoli differenti.

5. Informazioni Supplementari

5.1 Soggetto sano #7

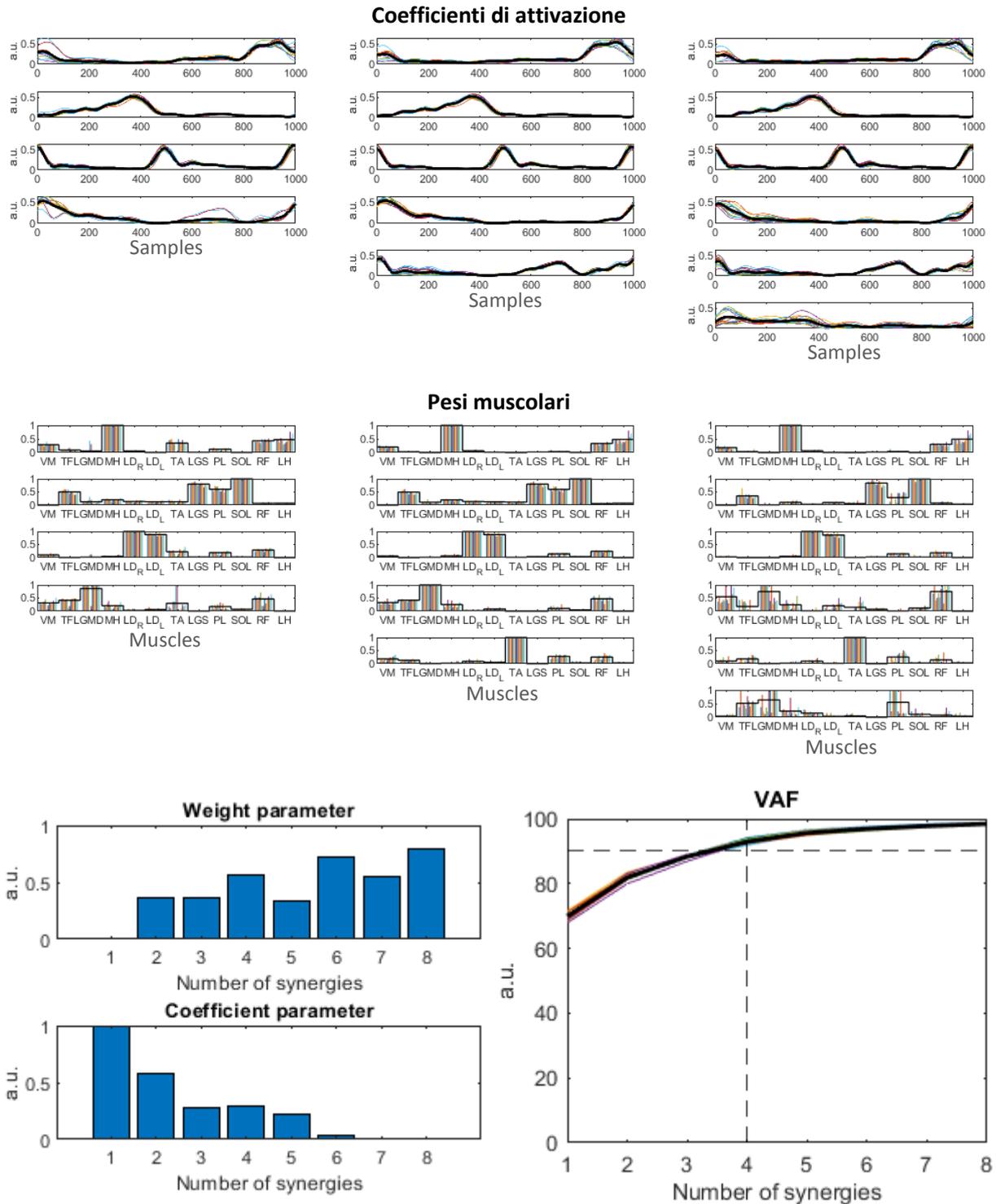


Figura 21: Coefficienti di attivazione, pesi muscolari, parametri dell'algorithm ChoOSyn e curva VAF relativa al soggetto reale sano numero 7

Il soggetto reale sano numero 7 [Fig. 21] è un esempio preso tra i soggetti a cui gli operatori esperti hanno attribuito come numerosità di sinergie muscolari corretta 5. Come si può notare, le motivazioni che hanno portato a questa scelta sono:

- il passaggio dalla numerosità 4 alla 5 permette di ottenere la quinta sinergia muscolare con il tibiale anteriore (TA) (prima assente), che come già visto precedentemente quasi sempre questo muscolo ha una sinergia a sé;
- il passaggio dalla numerosità 5 alla 6 non aggiunge informazioni, poiché si genera una sinergia muscolare con pesi correlati agli altri cinque, poco consistente tra i vari cicli di cammino, e con un coefficiente di attivazione praticamente piatto.

La numerosità selezionata come corretta (5) è stata individuata correttamente sia da ChoOSyn, che dai metodi E-VAF e P-VAF, mentre T-VAF ha indicato 4 (come si può vedere anche dalla figura). Come si può notare, l'algoritmo ChoOSyn è stato in grado di individuare la numerosità corretta grazie alla presenza di una conca nel parametro dei pesi muscolari e al cambio di pendenza con plateau nel parametro dei coefficienti di attivazione.

5.2 Soggetto sano #12

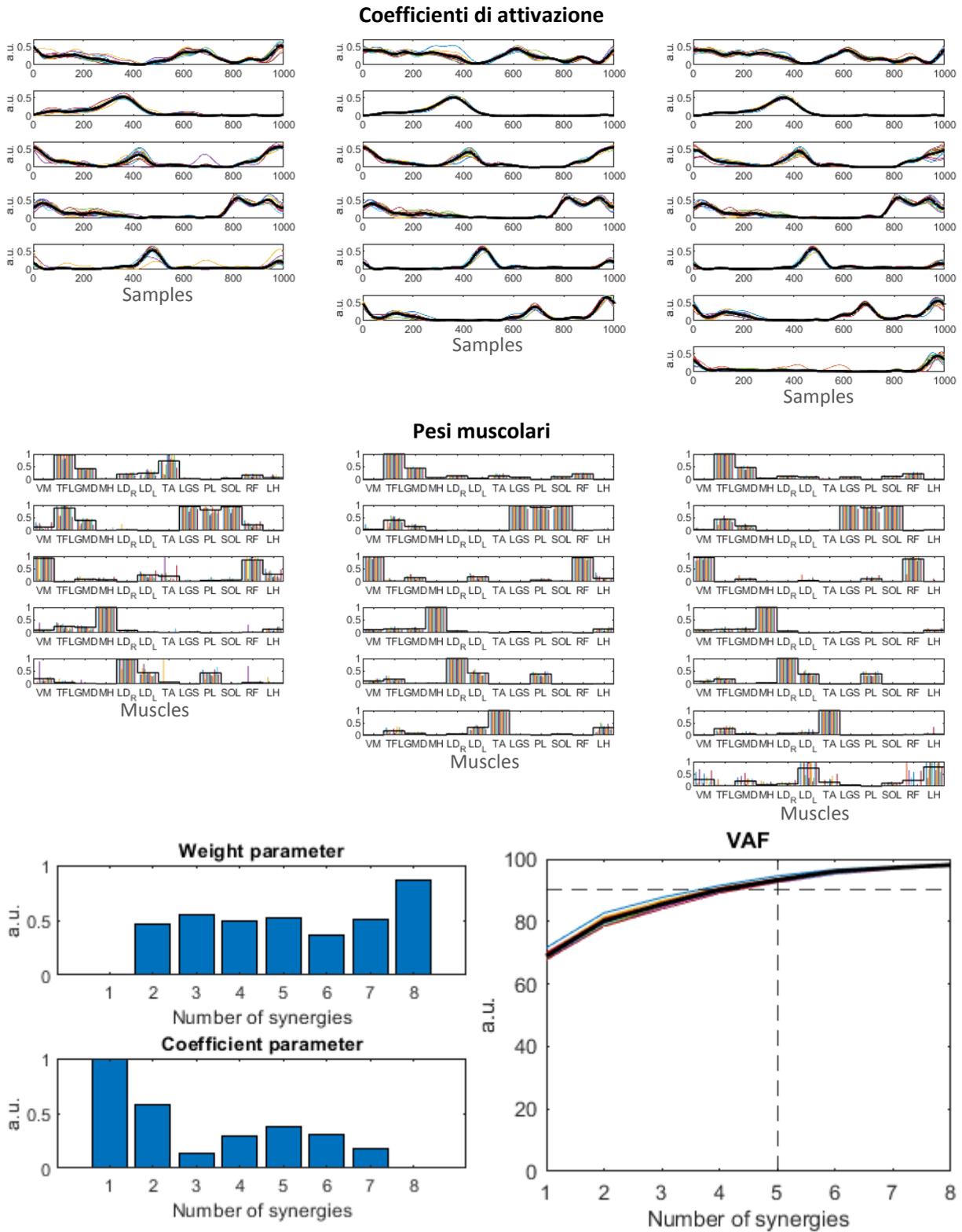


Figura 22: Coefficienti di attivazione, pesi muscolari, parametri dell’algoritmo ChoOSyn e curva VAF relativa al soggetto reale sano numero 12

Il soggetto reale sano numero 12 [Fig. 22] è un esempio preso tra i soggetti a cui gli operatori esperti hanno attribuito come numerosità di sinergie muscolari corretta 6. Come si può notare, le motivazioni che hanno portato a questa scelta sono:

- esattamente come nel caso precedente, il passaggio dalla numerosità 5 alla 6 permette di ottenere la sesta sinergia muscolare con il tibiale anteriore (TA) (prima assente), che come già visto precedentemente quasi sempre questo muscolo ha una sinergia a sé;
- il passaggio dalla numerosità 6 alla 7 rovina l'aspetto delle sinergie muscolari, poiché separa la sinergia dei due muscoli dorsali lunghissimi (LD_R e LD_L), generando inoltre una settima sinergia muscolare poco consistente.

La numerosità selezionata come corretta (6) è stata individuata correttamente sia da ChoOSyn, che dai metodi E-VAF e P-VAF, mentre T-VAF ha indicato 5 (come si può vedere anche dalla figura). Come si può notare, l'algoritmo ChoOSyn è stato comunque in grado di individuare la numerosità corretta grazie alla presenza di una conca nel parametro dei pesi muscolari, nonostante il parametro dei coefficienti di attivazione non fosse completamente chiaro.

5.3 Soggetto sano #19

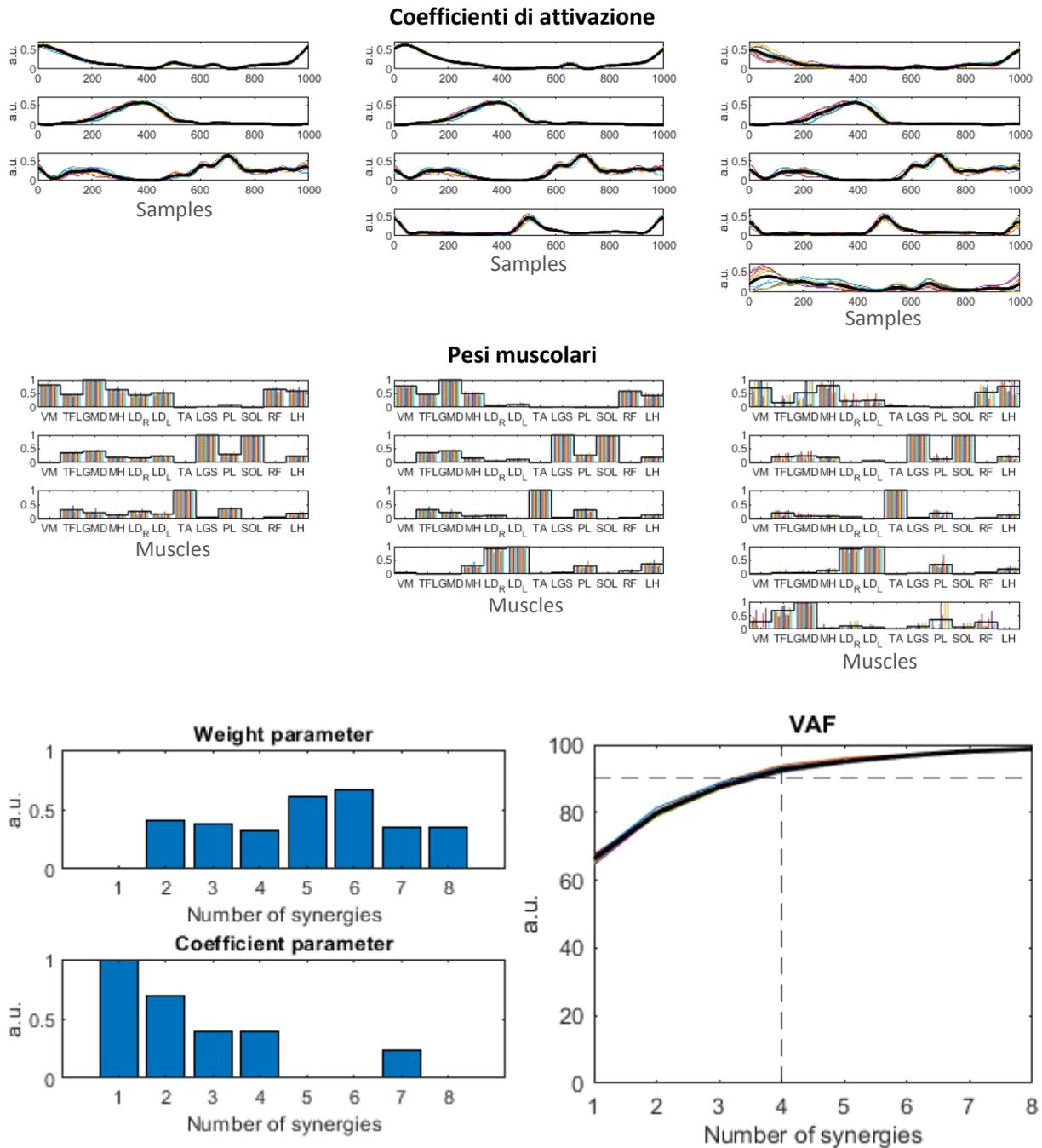


Figura 23: Coefficienti di attivazione, pesi muscolari, parametri dell’algoritmo ChoOSyn e curva VAF relativa al soggetto reale sano numero 19

Il soggetto reale sano numero 19 [Fig. 23] è un esempio preso tra i soggetti a cui gli operatori esperti hanno attribuito come numerosità di sinergie muscolari corretta 4. Come si può notare, le motivazioni che hanno portato a questa scelta sono:

- il passaggio dalla numerosità 3 alla 4 permette di ottenere la quarta sinergia muscolare con i due muscoli dorsali lunghissimi (LD_R e LD_L) (prima assente);
- il passaggio dalla numerosità 4 alla 5 non porta benefici, poiché separa malamente i muscoli della coscia, generando due sinergie muscolari (la prima e l'ultima) correlate tra di loro e molto poco consistenti.

La numerosità selezionata come corretta (4) è stata individuata correttamente sia da ChoOSyn, che dal metodo T-VAF, mentre E-VAF e P-VAF hanno indicato 7. Come si può notare, l'algoritmo ChoOSyn è stato in grado di individuare la numerosità corretta grazie al cambio di pendenza sia nel parametro dei pesi muscolari che in quello dei coefficienti di attivazione.

5.4 Soggetto sano #21

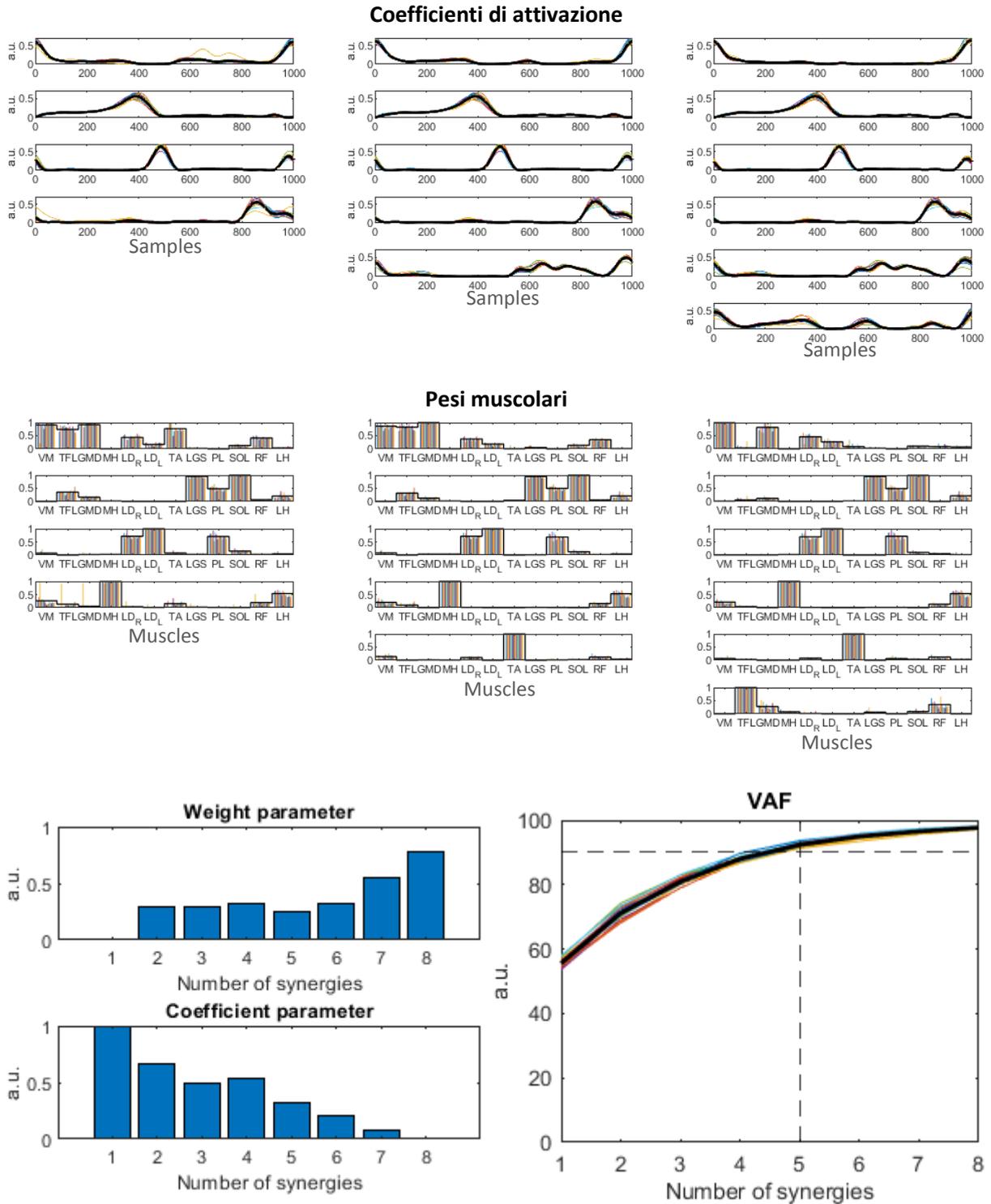


Figura 24: Coefficienti di attivazione, pesi muscolari, parametri dell’algoritmo ChoOSyn e curva VAF relativa al soggetto reale sano numero 21

Il soggetto reale sano numero 21 [Fig. 24] è un esempio preso tra i soggetti a cui gli operatori esperti hanno attribuito come numerosità di sinergie muscolari corretta 5. Come si può notare, le motivazioni che hanno portato a questa scelta sono:

- il passaggio dalla numerosità 4 alla 5 permette di ottenere la quinta sinergia muscolare con il solo tibiale anteriore (TA) (prima assente), che come già visto precedentemente quasi sempre questo muscolo ha una sinergia a sé;
- l'informazione aggiunta passando dalla numerosità 5 alla 6 è stata ritenuta poco rilevante, nonostante anche alla numerosità 6 le sinergie muscolari siano consistenti e poco correlate tra loro.

La numerosità selezionata come corretta (5) è stata individuata correttamente dal metodo T-VAF, mentre ChoOSyn, E-VAF e P-VAF hanno indicato 6. Come si può notare, l'algoritmo ChoOSyn si è basato principalmente sul cambio di pendenza nel parametro dei pesi muscolari, poiché anche in questo caso il parametro dei coefficienti di attivazione è difficilmente interpretabile.

- [1] F. O. Barroso *et al.*, "Combining muscle synergies and biomechanical analysis to assess gait in stroke patients," *J. Biomech.*, vol. 63, pp. 98–103, 2017, doi: 10.1016/j.jbiomech.2017.08.006.
- [2] C. L. Banks, M. M. Pai, T. E. McGuirk, B. J. Fregly, and C. Patten, "Methodological choices in muscle synergy analysis impact differentiation of physiological characteristics following stroke," *Front. Comput. Neurosci.*, vol. 11, no. August, pp. 1–12, 2017, doi: 10.3389/fncom.2017.00078.
- [3] D. F. Feeney, R. A. Capobianco, J. R. Montgomery, J. Morreale, A. M. Grabowski, and R. M. Enoka, "Individuals with sacroiliac joint dysfunction display asymmetrical gait and a depressed synergy between muscles providing sacroiliac joint force closure when walking," *J. Electromyogr. Kinesiol.*, vol. 43, pp. 95–103, Dec. 2018, doi: 10.1016/j.jelekin.2018.09.009.
- [4] I. Mileti *et al.*, "Parkinson's disease and Levodopa effects on muscle synergies in postural perturbation," *Med. Meas. Appl. MeMeA 2019 - Symp. Proc.*, pp. 1–6, 2019, doi: 10.1109/MeMeA.2019.8802229.
- [5] A. Falaki, X. Huang, M. M. Lewis, and M. L. Latash, "Dopaminergic modulation of multi-muscle synergies in postural tasks performed by patients with Parkinson's disease," *J. Electromyogr. Kinesiol.*, vol. 33, pp. 20–26, 2017, doi: 10.1016/j.jelekin.2017.01.002.
- [6] S. Grazioso *et al.*, "Assessment of upper limb muscle synergies for industrial overhead tasks: a preliminary study," pp. 89–92, 2019, doi: 10.1109/metroi4.2019.8792842.
- [7] J. Taborri *et al.*, "Feasibility of muscle synergy outcomes in clinics, robotics, and sports: A systematic review," *Appl. Bionics Biomech.*, vol. 2018, 2018, doi: 10.1155/2018/3934698.
- [8] A. D'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nat. Neurosci.*, vol. 6, no. 3, pp. 300–308, 2003, doi: 10.1038/nn1010.
- [9] I. Delis, P. M. Hilt, T. Pozzo, S. Panzeri, and B. Berret, "Deciphering the functional role of spatial and temporal muscle synergies in whole-body movements," *Sci. Rep.*, vol. 8, no. 1, pp. 1–17, 2018, doi: 10.1038/s41598-018-26780-z.
- [10] M. S. Shourijeh, T. E. Flaxman, and D. L. Benoit, "An approach for improving repeatability and reliability of non-negative matrix factorization for muscle synergy analysis," *J. Electromyogr. Kinesiol.*, vol. 26, pp. 36–43, 2016, doi: 10.1016/j.jelekin.2015.12.001.
- [11] K. Gui and D. Zhang, "Influence of locomotion speed on biomechanical subtask and muscle synergy," *J. Electromyogr. Kinesiol.*, vol. 30, pp. 209–215, 2016, doi: 10.1016/j.jelekin.2016.07.010.
- [12] M. Ghislieri, V. Agostini, and M. Knaflitz, "Muscle Synergies Extracted Using Principal Activations: Improvement of Robustness and Interpretability," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 453–460, 2020, doi: 10.1109/TNSRE.2020.2965179.
- [13] D. Rimini, V. Agostini, and M. Knaflitz, "Intra-subject consistency during locomotion: Similarity in shared and subject-specific muscle synergies," *Front. Hum. Neurosci.*, vol. 11, no. December, pp. 1–10, 2017, doi: 10.3389/fnhum.2017.00586.
- [14] M. C. Tresch, V. C. K. Cheung, and A. D'Avella, "Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets," *J. Neurophysiol.*, vol. 95, no. 4, pp. 2199–2212, 2006, doi: 10.1152/jn.00222.2005.

- [15] H. Hanawa *et al.*, “Muscle synergies underlying sit-to-stand tasks in elderly people and their relationship with kinetic characteristics,” *J. Electromyogr. Kinesiol.*, vol. 37, no. August, pp. 15–20, 2017, doi: 10.1016/j.jelekin.2017.08.004.
- [16] H. V. N. Sy, I. Nambu, and Y. Wada, “The adjustment of muscle synergy recruitment by controlling muscle contraction during the reaching movement,” *2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc.*, pp. 756–761, 2017, doi: 10.1109/SMC.2016.7844332.
- [17] K. M. Steele, M. C. Tresch, and E. J. Perreault, “The number and choice of muscles impact the results of muscle synergy analyses,” *Front. Comput. Neurosci.*, vol. 7, no. JUL, pp. 1–9, 2013, doi: 10.3389/fncom.2013.00105.
- [18] Y. Kim, T. C. Bulea, and D. L. Damiano, “Novel methods to enhance precision and reliability in muscle synergy identification during walking,” *Front. Hum. Neurosci.*, vol. 10, no. SEP2016, pp. 1–12, 2016, doi: 10.3389/fnhum.2016.00455.
- [19] V. C. K. Cheung, A. D’Avella, M. C. Tresch, and E. Bizzi, “Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviors,” *J. Neurosci.*, vol. 25, no. 27, pp. 6419–6434, 2005, doi: 10.1523/JNEUROSCI.4904-04.2005.
- [20] V. C. K. Cheung, L. Piron, M. Agostini, S. Silvoni, A. Turolla, and E. Bizzi, “Stability of muscle synergies for voluntary actions after cortical stroke in humans,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 46, pp. 19563–19568, 2009, doi: 10.1073/pnas.0910114106.
- [21] M. Hirashima and T. Oya, “How does the brain solve muscle redundancy? Filling the gap between optimization and muscle synergy hypotheses,” *Neurosci. Res.*, vol. 104, pp. 80–87, 2016, doi: 10.1016/j.neures.2015.12.008.
- [22] H. Yokoyama, N. Kaneko, T. Ogawa, N. Kawashima, K. Watanabe, and K. Nakazawa, “Cortical Correlates of Locomotor Muscle Synergy Activation in Humans: An Electroencephalographic Decoding Study,” *iScience*, vol. 15, pp. 623–639, 2019, doi: 10.1016/j.isci.2019.04.008.
- [23] D. Rimini, V. Agostini, and M. Knaflitz, “Evaluation of Muscle Synergies Stability in Human Locomotion: A comparison between normal and fast walking speed,” pp. 1–5, 2017.
- [24] S. A. Chvatal and L. H. Ting, “Voluntary and reactive recruitment of locomotor muscle synergies during perturbed walking,” *J. Neurosci.*, vol. 32, no. 35, pp. 12237–12250, 2012, doi: 10.1523/JNEUROSCI.6344-11.2012.
- [25] A. S. Oliveira, L. Gizzi, D. Farina, and U. G. Kersting, “Motor modules of human locomotion: Influence of EMG averaging, concatenation, and number of step cycles,” *Front. Hum. Neurosci.*, vol. 8, no. MAY, pp. 1–9, 2014, doi: 10.3389/fnhum.2014.00335.
- [26] D. Seung and L. Lee, “Algorithms for non-negative matrix factorization,” *Adv. Neural Inf. Process. Syst.*, vol. 13, no. February, pp. 556–562, 2001.
- [27] Y. Li and A. Ngom, “The non-negative matrix factorization toolbox for biological data mining,” *Source Code Biol. Med.*, vol. 8, 2013, doi: 10.1186/1751-0473-8-10.
- [28] H. Kim and H. Park, “NONNEGATIVE MATRIX FACTORIZATION BASED ON ALTERNATING NONNEGATIVITY CONSTRAINED LEAST SQUARES AND ACTIVE SET METHOD,” vol. 30, no. 2, pp. 713–730, 2008.
- [29] I. Delis, B. Berret, T. Pozzo, and S. Panzeri, “Quantitative evaluation of muscle synergy models:

- A single-trial task decoding approach," *Front. Comput. Neurosci.*, vol. 7, no. FEB, pp. 1–21, 2013, doi: 10.3389/fncom.2013.00008.
- [30] N. Mehrabi, M. H. Schwartz, and K. M. Steele, "Can altered muscle synergies control unimpaired gait?," *J. Biomech.*, vol. 90, pp. 84–91, 2019, doi: 10.1016/j.jbiomech.2019.04.038.
- [31] C. S. Zandvoort, J. H. van Dieën, N. Dominici, and A. Daffertshofer, "The human sensorimotor cortex fosters muscle synergies through cortico-synergy coherence," *Neuroimage*, vol. 199, no. February, pp. 30–37, 2019, doi: 10.1016/j.neuroimage.2019.05.041.