
Tanto rumore per nulla? Il ruolo delle fluttuazioni nella dinamica delle reti nervose

Quell'altro organo che chiamiamo cervello, quello con cui veniamo al mondo, quello che trasportiamo nel cranio e che trasporta noi affinché noi trasportiamo lui, non è mai riuscito a produrre altro che intenzioni vaghe, generiche, diffuse, e soprattutto poco variate, riguardo a ciò che le mani e le dita dovranno fare.

————— *“La Caverna”, José Saramago*

Paolo Del Giudice
Maurizio Mattia

*Centro Nazionale per la Protezione dalle Radiazioni e Fisica Computazionale,
Istituto Superiore di Sanità, Roma
Centro Nazionale per la Protezione dalle Radiazioni e Fisica Computazionale,
Istituto Superiore di Sanità, Roma*

L'idea che i sistemi di Intelligenza Artificiale (IA) possano trarre vantaggiosa ispirazione dalle scienze del cervello è molto naturale, ha permeato in varie forme gli sviluppi di IA per decenni e, nella sua formulazione generica, costituisce ormai un luogo comune. Non sorprende quindi che in un numero della rivista dedicato alla IA trovi posto un articolo dedicato ad alcuni aspetti della modellistica teorica del sistema nervoso. Se però si va a scandagliare la storia dei rapporti tra IA e teorie del cervello (che è complessa e non tenteremo di riassumerla) si

scopre che l'ispirazione di cui sopra ha preso forme qualitativamente diverse; ne discuteremo brevemente in una parte introduttiva per fornire il contesto di quanto diremo in seguito. Passeremo poi a descrivere alcuni modelli fisico-matematici di attività nervosa a varie scale, mettendo a fuoco il ruolo delle fluttuazioni casuali come elemento costitutivo della dinamica neuronale.

Pur accettando per semplicità il punto di vista, piuttosto diffuso, secondo il quale il cervello effettua delle *computazioni* sugli input ambientali, e ove opportuno ne traduce il risultato in azio-

ni (trascurando cioè la natura profondamente *relazionale* del loop percezione-azione), in un approccio ad un sistema intelligente ispirato alla conoscenza del cervello si può ad esempio: *i*) emulare, in qualunque forma conveniente, alcune elaborazioni già identificate nel sistema nervoso (imitare l'analisi in frequenza effettuata dalla coclea sul suono; riprodurre parte della gerarchia di filtri implementati nelle aree visive, come quelli suggeriti dalla *independent component analysis*; ricostruire - sulla base di quanto se ne è finora capito in neurofisiologia - la pianificazione e l'esecuzione del movimento per sviluppare robot con capacità motorie complesse, ecc.); *ii*) sviluppare sistemi di apprendimento automatico, i cui elementi siano più o meno vagamente ispirati alle unità di elaborazione del sistema nervoso, e la cui architettura venga determinata da un processo di ottimizzazione rispetto alle prestazioni su specifici compiti (*reti neurali, machine learning*); *iii*) sviluppare, in diretta connessione con i dati sperimentali, modelli fisico-matematici in grado di cogliere elementi fondamentali della dinamica di popolazioni di cellule nervose fortemente interconnesse, per comprendere in che modo le loro proprietà di risposta agli stimoli e di auto-organizzazione possano generare delle *primitive computazionali* per lo svolgimento di compiti cognitivi complessi.

È di quest'ultimo approccio che tratteremo nel seguito.

Alcuni principi dinamici

Un principio molto generale di organizzazione dei sistemi biologici è la ricerca di un compromesso ottimale tra *exploitation* e *exploration*. Da un lato un organismo deve poter fornire risposte riproducibili e funzionalmente adeguate, determinate da adattamenti alle richieste ambientali. Dall'altro è vantaggioso che esso non rimanga fossilizzato nelle soluzioni dettate dalle situazioni contingenti cui è stato esposto, ed elabori una strategia di esplorazione del proprio spazio degli stati, in grado di generare soluzioni non accessibili attraverso un'ottimizzazione 'locale'. Le direzioni di questa esplorazione non dovrebbero essere troppo vincolate dalla configurazione corrente, e questo ci fa già intuire il possibile ruolo costruttivo di una componente di *rumore*.

L'esempio forse più familiare di questo principio generale è costituito dall'evoluzione biologica, prodotta della combinazione di variabilità casuale del patrimonio genetico e di pressione selettiva per la sopravvivenza del più adatto ('caso e necessità'), in un ambiente che in generale è mutevole nel tempo.

Nei classici modelli matematici di evoluzione la pressione selettiva è descritta da un "paesaggio di fitness" (eventualmente dinamico), e l'equilibrio *exploitation/exploration* si esprime nella tendenza a popolare le configurazioni di massima fitness (i picchi locali del paesaggio), come risposta ai vincoli ambientali contingenti, e contemporaneamente nella esplorazione del paesaggio attraverso variazioni casuali del genoma, che rendono accessibili 'soluzioni' evolutive che non sarebbero localmente accessibili.

Cambiando prospettiva dall'evoluzione delle specie allo sviluppo embrionale del singolo individuo, fin dagli anni '40 fu formulata la metafora di un 'paesaggio epigenetico', del quale lo sviluppo cellulare dell'embrione segue le pendenze e le biforcazioni, che rappresentano in modo astratto un sistema di vincoli per lo sviluppo.

Vedremo nel seguito che simili schematizzazioni teoriche, basate su un 'paesaggio' come rappresentazione di un sistema di vincoli, e su fluttuazioni casuali come opportunità di esplorazione dello spazio degli stati, hanno un corrispettivo diretto nella modellizzazione del sistema nervoso.

Il sistema nervoso deve in effetti risolvere un problema di *exploitation/exploration* analogo: deve da un lato assicurare l'affidabilità del circuito percezione-azione nella sue varie forme, e la robustezza delle rappresentazioni che costruisce per classificare la realtà; se così non fosse non sarebbe neanche possibile la concordanza, che ci appare invece naturale, tra ciò che diversi cervelli classificano come 'rosso' o 'cane', ed i movimenti dei nostri simili nel reagire agli stimoli ambientali ci apparirebbero imprevedibili. D'altra parte, non solo diversi cervelli possono costruire sequenze di pensieri ed immagini mentali profondamente diverse a parità di condizioni ambientali, com'è ovvio dall'esperienza comune, ma in determinate condizioni, ed in modo misurabile, l'attività nervosa varia in modo apparentemente casuale in risposta agli stessi stimo-

li, e (ne vedremo un esempio) può alternare in modo stocastico tra le possibili rappresentazioni mentali alternative dello stesso stimolo.

Se i picchi del paesaggio di fitness dell'evoluzione biologica rappresentano soluzioni evolutive 'localmente ottimali', i picchi di un analogo paesaggio che funga da palcoscenico per la dinamica nervosa possono esprimere ad esempio l'ottimalità locale delle rappresentazioni neurali di scene visive o di atti motori, a seconda del contesto.

In questa metafora, il cervello transisce continuamente tra gli stati che corrispondono a tali rappresentazioni, in parte a causa di input mutevoli, in parte sotto la spinta di elementi dinamici endogeni. Come vedremo, la disposizione a queste transizioni si deve sia alla capacità del sistema di modificare la porzione del paesaggio in cui temporaneamente risiede, sia a diverse sorgenti di rumore.

Notiamo di passaggio che se, al limite, descriviamo la dinamica delle rappresentazioni espresse nel cervello come transizioni quasi-discrete tra configurazioni localmente ottimali, di fatto ci ricollegiamo in parte ad una antica diatriba sulla natura continua o discreta della percezione cosciente, che risale a D. Hume e passa per W. James e H. Bergson. È comunque corretto dire che, a livello sia sperimentale che teorico, il dibattito sulla possibilità di interpretare la dinamica nervosa in termini di transizioni discrete tra stati ben definiti è tuttora aperto.

Assumendo di adottare questo quadro concettuale di riferimento per la descrizione della dinamica del sistema nervoso, dobbiamo definire quali quantità sperimentalmente osservabili definiscano gli stati del sistema e che cosa, e come, determini il 'paesaggio' in cui tali stati evolvono.

A prima vista, l'eterogeneità delle nostre esperienze soggettive in relazione a diversi stati mentali sembra opporsi alla formulazione di un'astrazione di valore generale. In effetti, nella nostra esperienza soggettiva, le diverse situazioni concrete nelle quali si istanzia la tensione tra ottimalità locale ed esplorazione casuale a cui abbiamo accennato ci appaiono qualitativamente diverse: riconoscere un sapore, classificare una forma, memorizzare una informazione, decidere un movimento, corrispondono a rapporti con il nostro corpo, nelle sue funzioni sensoriali, cognitive

o motorie, che avvertiamo come appartenenti a domini diversi. Ma appena non appena ci allontaniamo dall'involucro sensoriale o motorio del sistema nervoso (cioè dalle sue componenti direttamente coinvolte nella trasduzione degli stimoli o nelle istruzioni motorie) queste distinzioni qualitative sfumano in una trama complessa di messaggi stereotipati (potenziali d'azione o 'spike') scambiati tra elementi sostanzialmente molto simili tra loro (neuroni) attraverso giunzioni (sinapsi) che si presentano in pochi tipi largamente distribuiti in tutto il cervello. Se un agente microscopico potesse intercettare gli spike scambiati localmente in qualunque area del cervello non avrebbe modo di qualificarli come 'spike visivi', 'uditivi', 'motori' o altro: solo l'organizzazione spazio-temporale del flusso di spike scambiati definisce la 'computazione'.

Questo fatto (che costituisce una scelta evolutiva peculiare e non ovvia) ci fornisce dunque un livello di descrizione di appropriata generalità e ci permette di definire lo stato del sistema come la collezione degli spike scambiati tra i neuroni di una popolazione di interesse ad un certo istante. Inoltre, per dar conto della componente stocastica di 'exploration', i processi di generazione e scambio degli spike includeranno componenti di 'rumore', di cui esamineremo nel seguito il senso e l'origine. Notiamo subito che questa definizione dello stato di una popolazione neuronale è in certo senso una definizione 'microscopica', analoga a quella che definisce lo stato di un gas ad un certo istante come l'insieme delle posizioni e velocità delle sue molecole a quell'istante. In quell'ambito (in cui il 'rumore' è determinato dalla temperatura), è molto utile adottare una descrizione in termini della distribuzione di probabilità di quantità osservabili (meccanica statistica), e da questa derivare leggi macroscopiche (termodinamica). Vedremo sommariamente che, in qualche misura, lo studio teorico della dinamica neuronale procede attraverso un percorso simile.

Riguardo a cosa e come si determini la forma del 'paesaggio' in cui il sistema neuronale evolve, l'idea chiave, di nuovo in analogia con modelli familiari in fisica, è che essa sia primariamente legata alle interazioni tra i neuroni, cioè alla forza delle sinapsi che mediano lo scambio di spike tra coppie di neuroni. Come abbiamo accennato,

la forma del paesaggio esprime un insieme di vincoli a cui il sistema deve adattarsi in modo ottimale, identificando ad esempio, con i suoi picchi (o valli, secondo le convenzioni adottate) gli stati neuronali che forniscono la risposta ottimale ad una classe di stimoli. La forma del paesaggio deve quindi riflettere il modo in cui, nelle sue relazioni con il mondo esterno e con altre popolazioni neuronali, la popolazione in questione ha strutturato nel tempo le interazioni tra i suoi neuroni in modo da scolpire i profili di picchi e valli per poi guidare la dinamica neuronale in modo appropriato alle richieste ambientali e comportamentali.

Vediamo quindi che siamo di fronte a dinamiche accoppiate su scale di tempo diverse: quella 'veloce' della dinamica neuronale, sulla cui scala temporale il paesaggio è quasi statico (a meno di effetti locali cui abbiamo accennato e che vedremo in seguito), ed una dinamica 'lenta' che scolpisce il paesaggio in funzione dei vincoli esterni ed interni agendo anche sugli accoppiamenti tra i neuroni; quest'ultima viene spesso identificata, in varie forme, come 'apprendimento'; ne accenneremo soltanto, ma è una parte importante dello schema teorico.

Formalizzare gli elementi di base: neuroni e sinapsi

Nella formulazione di modelli fisico-matematici di sistemi biologici si accetta comunemente un certo livello (spesso elevato) di semplificazione e schematizzazione del dato biologico: questo è il prezzo che si accetta di pagare per poter rendere i modelli trattabili formalmente (o anche solo realisticamente simulabili al calcolatore). Tale semplificazione viene a volte motivata, soprattutto dai fisici, non solo come una inevitabile limitazione, ma come l'espressione del fatto (imparato dalla meccanica statistica) che spesso il comportamento collettivo di grandi insiemi di unità interagenti non dipende troppo dai dettagli delle singole unità. Come in altri domini scientifici, il problema difficile è identificare il livello appropriato di semplificazione in grado di assicurare i vantaggi di cui sopra, senza però allontanare troppo il modello dalla controparte biologica che intende descrivere.

Nel caso dei modelli dinamici di neurone (non è così per quelli sinaptici) si è riusciti a costruire una gerarchia di modelli di semplicità crescente, partendo da modelli 'realistici' in grado di confrontarsi quantitativamente con misure sperimentali dettagliate e passando attraverso una serie di approssimazioni controllate, in modo tale che ad ogni passo di semplificazione è abbastanza chiaro l'impatto dei dettagli che si è scelto di ignorare o accorpare.

La caratteristica comune a tutti i modelli neuronali consiste nel descrivere sistemi *eccitabili*: informalmente, sistemi che rispondono in modo quasi lineare ad input piccoli, ma che per input superiori ad una soglia esibiscono una risposta molto nonlineare, ed essenzialmente stereotipata. L'eccitabilità cellulare non è esclusiva dei neuroni (si pensi alle cellule cardiache). Nel sistema nervoso, le proprietà di risposta nonlineare sono la chiave per comprendere le capacità computazionali dei neuroni.

Non tentiamo di riassumere la gerarchia di modelli neuronali, e le strategie di semplificazione ed analisi che ne sono alla base; adatteremo piuttosto direttamente una classe di modelli che ha avuto un ruolo chiave nella modellistica teorica (e che si colloca verso il fondo della sequenza di semplificazioni a cui abbiamo accennato): il neurone *integrate-and-fire* (IF).

Mentre in modelli più complessi le proprietà eccitabili del neurone sono esplicitamente descritte in termini della dinamica di gradi di libertà interni al neurone, nel modello IF il neurone è descritto da una sola variabile di stato, il potenziale di membrana $V(t)$ (la cui variazione nei modelli dettagliati determina l'apertura o chiusura di canali ionici sulla membrana del neurone, che a loro volta ne determinano il comportamento nonlineare) la cui dinamica lineare guidata dall'input sinaptico $I(t)$ è

$$\frac{dV}{dt} = -\frac{V(t)}{\tau} + I(t).$$

Le proprietà nonlineari necessarie per descrivere la generazione di uno spike (espressione della risposta stereotipata sopra soglia del sistema eccitabile, cui facevamo riferimento) sono condensate matematicamente in una condizione al contorno di questa equazione per il potenziale di membrana: quando V supera una soglia θ ,

lo spike viene emesso, V viene riportato ad un valore di riferimento H , per riprendere dopo un periodo 'refrattario assoluto' τ_0 la sua evoluzione in regime lineare, fino al prossimo eventuale raggiungimento della soglia. In assenza di input, $V(t)$ ha un decadimento esponenziale con tempo caratteristico τ . L'input sinaptico è determinato dalla combinazione lineare degli spike emessi dai K neuroni presinaptici:

$$I(t) = I_{\text{ext}}(t) + \sum_{j=1}^K J_j \sum_k \delta(t - t_{j,k} - d_j).$$

Il k -esimo spike emesso dal neurone j -esimo al tempo t_j , k viene ricevuto con un ritardo assonale d_j ed è rappresentato da una δ di Dirac, e induce un salto del potenziale V pari all'efficacia sinaptica J_j . Il contributo di corrente sinaptica dovuto a neuroni esterni alla rete descritta, è riassunto nel termine I_{ext} .

Dai singoli neuroni alle reti: modularità, teoria di campo medio e descrizione mesoscopica

Come già accennato, le reti nervose mostrano una complessa organizzazione sia da un punto di vista strutturale che funzionale. Diverse componenti, che includono ad un livello macroscopico le aree corticali identificate più di un secolo fa dal neurologo tedesco Korbinian Brodmann, sono tra loro gerarchicamente connesse contribuendo con differenti ruoli all'emergenza delle funzioni cognitive. L'eterogeneità di questi elementi si manifesta principalmente ad una scala spaziale "mesoscopica" più ridotta, dove la diversa citoarchitettura appare evidente nella composizione cellulare dei cosiddetti *strati corticali* (*layers*) (Figura 1a). Ogni strato in queste reti strutturate contribuisce con un ruolo specifico alla elaborazione e al trasferimento dell'informazione ricevuta da altre regioni cerebrali dando luogo ad un circuito "canonico" che si ripete quasi invariato come le piastrelle di un pavimento. In questo circuito l'input sensoriale da strutture subcorticali come il talamo viene acquisito dalle sottoreti del quarto strato, amplificato e modulato da quelle del quinto e ritrasmesso dai neuroni degli strati

più esterni. Questo modulo mesoscopico è riconoscibile come un fascio di fibre perpendicolare alla superficie corticale con una sezione di un centinaio di micron identificata alla fine degli anni '50 del secolo scorso dal neurofisiologo statunitense Vernon B. Mountcastle (Figura 1b). Le reti nervose condividono questo principio costituente di modularità con altri sistemi biologici descrivibili come reti, quali le cosiddette vie metaboliche e la regolazione genica, per citare un paio di esempi. La modularità infatti rende un sistema più versatile e quindi vantaggioso da un punto di vista evolutivo. Rende ad esempio più semplice la riorganizzazione delle connessioni tra unità funzionali, in confronto al costo richiesto nel trasformare la matassa inestricabile di legami associata ad una rete omogenea.

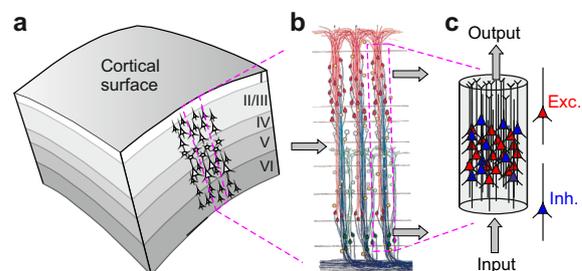


Figura 1: Rappresentazione schematica dell'organizzazione a strati della corteccia cerebrale (a) composta da moduli di fibre nervose segregate in "minicolonne" (b). I diversi strati corticali svolgono ruoli diversi di ricezione e l'invio di messaggi da e verso altre aree cerebrali e colonne (freccie grigie). Nella modellistica delle reti neuronali, una minicolonna è rappresentata da una popolazione di neuroni eccitatori ed inibitori (c). Pannello (b) adattato da [1].

La modularità gerarchica delle reti cerebrali offre d'altro canto un'opportunità di semplificazione nella descrizione teorica delle loro dinamiche collettive. È infatti possibile immaginare di poter trascurare il dettaglio microscopico di queste colonne corticali, quale ad esempio la posizione e l'estensione spaziale delle cellule nervose che le compongono, per concentrarci principalmente sulle caratteristiche di "grana grossa" del loro stato, verificando a posteriori l'accuratezza di questa approssimazione (approccio, questo, ampiamente adottato in settori della fisica teorica ben sedimentati come quello della meccanica statistica).

In questo contesto, consideriamo ragionevolmente adeguato descrivere una colonna corticale come composta da neuroni semplificati del tipo *IF* introdotto precedentemente (Figura 1c). Per mantenere una sufficiente aderenza alla realtà biologica i neuroni di questo modulo corticale devono essere suddivisi in due categorie: eccitatori ed inibitori. Gli spike emessi dai primi inducono attraverso una trasmissione sinaptica *glutammatergica* (cioè mediata dal glutammato) un incremento del potenziale di membrane V nei neuroni post-sinaptici (neuroni bersaglio) favorendone l'avvicinamento alla soglia di emissione θ . Gli inibitori sono invece caratterizzati da una trasmissione sinaptica *GABA-ergica* (cioè mediata dal neurotrasmettitore GABA) che "iperpolarizza" il potenziale V del neurone bersaglio allontanandolo da θ . L'entità di questa deflessione dipende dalla cosiddetta *efficacia sinaptica* J che sarà positiva per i neuroni eccitatori e negativa per gli inibitori.

In corteccia cerebrale, ogni neurone riceve decine di migliaia di contatti sinaptici da neuroni che solo in parte appartengono al modulo corticale in questione. Il contributo maggioritario di corrente sinaptica dovuto agli spike emessi da neuroni "esterni" alla colonna corticale, può in prima approssimazione essere considerato come un *bagno termico* in analogia con la termodinamica, intendendo che questa componente si assume indipendente dall'evoluzione temporale della popolazione di neuroni "interni".

Altra caratteristica chiave dell'attività corticale è l'elevata irregolarità con la quale i neuroni emettono gli impulsi nervosi. A prima vista questi possono apparire come eventi casuali che si manifestano con frequenza variabile nel tempo. L'idea che un elaboratore di informazioni quale è il nostro cervello sia costituito da elementi dalla natura così rumorosa e quindi inaffidabile, può apparire quantomeno bizzarro. In realtà, questa apparente contraddizione svanisce se si considera l'informazione come distribuita tra i neuroni di una rete nervosa. In un elaboratore parallelo di siffatta natura, il singolo elemento darà un contributo molto parziale alla costituzione del messaggio da trasmettere o alla computazione da svolgere. Nelle reti nervose questa strategia è garantita dall'elevato grado di parallelismo implementato, ovverosia dal gran numero di ele-

menti con cui ogni singolo neurone è in contatto attraverso le sinapsi del proprio albero dendritico. I grandi numeri in gioco rendono applicabile il teorema del limite centrale, portando ad avere correnti sinaptiche dominate dalle proprietà medie degli impulsi emessi dall'intera popolazione neuronale (naturalmente l'applicabilità del teorema richiede che le fluttuazioni dei molti input sinaptici siano anche tra loro indipendenti; come discuteremo brevemente appresso, a parte alcune condizioni patologiche la validità di questa ipotesi è in realtà abbastanza ben verificata nelle tipiche condizioni corticali).

Una tale codifica istantanea dell'informazione distribuita nello 'spazio', cioè tra i neuroni della rete, offre notevoli vantaggi come la resilienza ad eventuali malfunzionamenti dei singoli neuroni. Non solo: il singolo neurone potrà essere una unità di elaborazione relativamente lenta e inaffidabile, producendo messaggi/impulsi a frequenza ridotta ed in modo irregolare nel tempo con un conseguente risparmio di energia metabolica. Il prezzo da pagare nella scelta di questa strategia architettonica risiede nella necessità di dover costituire una fitta rete di comunicazione, difficoltà che però le reti cerebrali hanno risolto con successo.

Ma cosa determina l'irregolarità dei treni d'impulsi emessi da un neurone? Una corrente sinaptica $I(t)$ interamente guidata dalla sola media statistica degli impulsi ricevuti istante per istante sull'albero dendritico porterebbe ad una evoluzione del potenziale di membrana somatico $V(t)$ prettamente deterministica. Questo in realtà non accade, e il motivo risiede principalmente nel fatto che l'input sinaptico è una combinazione di contributi sia eccitatori che inibitori. Per capire le ragioni di questa affermazione, occorre prima considerare che i treni d'impulsi dei neuroni presinaptici con buona approssimazione si compongono linearmente sull'albero dendritico. La corrente $I(t)$ può allora essere vista come dovuta ad un treno di impulsi con statistica di Poisson. Combinando infatti un elevato numero di processi puntali (le sequenze di impulsi emessi dai singoli neuroni) con qualsivoglia statistica, il treno di impulsi risultante dalla loro somma è assimilabile ad un processo di Poisson con frequenza pari al prodotto $K \nu(t)$ del numero di neuroni presinaptici e la loro frequenza media di emissione

(Figura 2a). Questa è la conseguenza del teorema del limite centrale di Palm–Khintchin per i processi di ‘renewal’ stazionari generalizzato successivamente da B. Grigelionis al caso in cui la statistica di questi treni non sia omogenea nel tempo. Nell’ipotesi che i tempi di emissione degli spike tra due neuroni possano essere considerati indipendenti, l’intera sequenza di spike presinaptici sarà Poissoniana nel limite di frequenza $K\nu(t)$ infinita. Per quanto accennato in precedenza, ambedue queste condizioni sono ben soddisfatte dai neuroni corticali, sia per l’elevato numero di contatti sinaptici K , che per la ridotta frequenza $\nu(t)$ di ‘sparo’ (firing) registrata *in vivo* nelle cellule nervose. Quest’ultimo aspetto contribuisce a diluire nel tempo l’interazione tra due neuroni sinapticamente connessi, rendendoli nella pratica reciprocamente indipendenti.

Il continuo bombardamento di spike inibitori ed eccitatori se opportunamente bilanciati può portare ad avere ampie fluttuazioni della corrente presinaptica anche nel limite $K\nu \rightarrow \infty$. Per illustrare questo concetto è sufficiente derivare la media μ e la varianza σ^2 di questa corrente che per la sua natura Poissoniana risultano essere una combinazione lineare delle frequenze totali $K_I\nu_I(t)$ e $K_E\nu_E(t)$ come illustrato da Daniel J. Amit e Nicolas Brunel in [3]:

$$\begin{aligned}\mu(t) &= J_E K_E \nu_E(t) - J_I K_I \nu_I(t) \\ \sigma^2(t) &= J_E^2 K_E \nu_E(t) + J_I^2 K_I \nu_I(t),\end{aligned}$$

in cui J_E e J_I rappresentano le intensità delle efficacie sinaptiche, cioè la variazione assoluta del potenziale V in corrispondenza all’arrivo di uno spike eccitatorio o inibitorio, rispettivamente. Queste due componenti si sottraggono nella media ma si sommano nella varianza e, a patto di riscalare le efficacie sinaptiche J_E e J_I , è sempre possibile ottenere dei valori finiti di μ e σ nel limite suddetto. Se la componente media della corrente non è sufficiente a portare il potenziale di membrana ad attraversare la soglia di emissione, gli spike verranno prodotti solo in presenza di fluttuazioni positive di $V(t)$. Questo porta al comportamento illustrato in Figura 2b, dove il potenziale di membrana del neurone spende la maggior parte del tempo a fluttuare al di sotto della soglia θ , emettendo spike a tempi irregolari. Come illustrato poco più di vent’anni fa da Carl

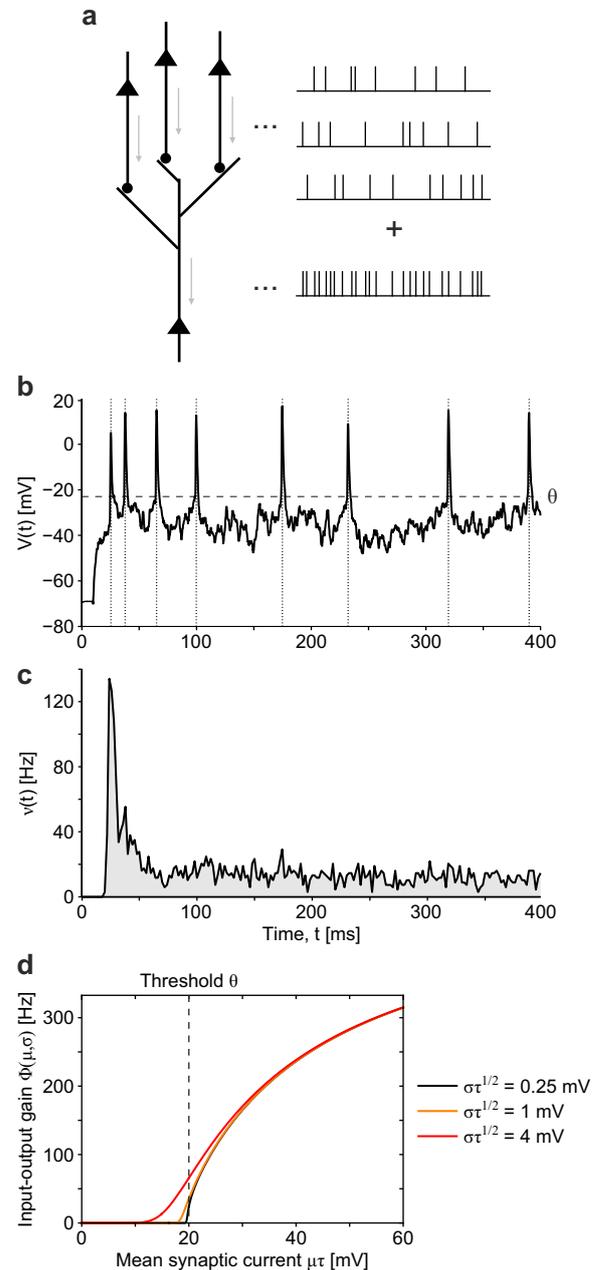


Figura 2: (a) Schema di neurone e della composizione degli spike presinaptici in un unico treno Poissoniano. Sequenze di linee verticali, treni d’impulsi emessi da un neurone. (b) Simulazione di un neurone IF che riceve un treno di impulsi Poissoniani in uno stato bilanciato tra eccitazione e inibizione. θ , soglia di emissione. (c) Frequenza $\nu(t)$ degli spike emessi da una rete di neuroni statisticamente equivalenti a quello mostrato in (b). (d) Funzione di guadagno $\Phi(\mu, \sigma)$ che da la frequenza di sparo di un neurone con corrente d’input di media μ e varianza σ^2 . τ è la costante di rilassamento di $V(t)$ nel neurone IF. Pannello (a) adattato da [2].

van Vreeswijk e Haim Sompolinsky [4], tale regime di emissione guidato principalmente dalle fluttuazioni delle corrente sinaptica può essere

prodotto in modo auto-consistente da una rete neuronale, e va sotto il nome di ‘stato bilanciato’ tra eccitazione e inibizione.

Il limite in cui le efficacie sinaptiche J sono molto più piccole della soglia di emissione θ , oltre ad avvicinarsi a quanto misurato sperimentalmente, permette di descrivere teoricamente il potenziale $V(t)$ come un processo stocastico di diffusione che segue l’equazione di Langevin

$$\frac{dV}{dt} = -\frac{V(t)}{\tau} + \mu(t) + \sigma(t) \xi(t),$$

dove $\xi(t)$ è un rumore Guassiano senza memoria con media nulla e varianza unitaria. In condizioni stazionarie questo è il cosiddetto processo di Ornstein–Uhlenbeck, noto modello della velocità di una particella Browniana in un mezzo viscoso. Per semplicità consideriamo il solo caso in cui la trasmissione sinaptica è istantanea, cioè in cui l’arrivo di uno spike presinaptico induce una variazione di V su una scala di tempo molto breve rispetto a τ , e che i momenti della corrente non siano dipendenti dal potenziale di membrana stesso. Queste complicazioni che rendono la modellistica del neurone più aderente alla realtà biologica non cambiano qualitativamente il comportamento collettivo delle reti neurali, ma in linea di principio possono, e spesso vengono, incluse come una generalizzazione del formalizzazione teorica qui riportata.

Nel descrivere la dinamica collettiva di una rete di neuroni siffatti, si può assumere che, sulla scala di interesse di una colonna corticale, in un dato strato, le proprietà statistiche delle correnti sinaptiche che ricevono siano approssimativamente uguali per tutti i neuroni. La corrente sinaptica in ogni cellula nervosa di una rete omogenea (eccitatoria o inibitoria) di una colonna corticale sarà allora una realizzazione diversa dello stesso processo stocastico. Questa è la cosiddetta approssimazione di ‘campo media estesa’ [3], ed è giustificata dal fatto che essendo il numero di contatti sinaptici K molto grande, lo scostamento dal suo numero medio sarà relativamente piccolo, come ci assicura la legge dei grandi numeri. I neuroni quindi sentiranno l’effetto dell’attività prodotta dalle altre $N - 1$ cellule della rete attraverso la frequenza $\nu(t)$ data dal numero totale degli impulsi emessi al tempo t per unità di neurone (cioè diviso per N). La

$\nu(t)$, di cui è illustrato un esempio in Figura 2c, è una variabile di stato collettiva e l’individualità dei neuroni a questo livello si perde. È da notare che anche se il singolo impulso emesso da un neurone presinaptico ha un effetto trascurabile nella dinamica di $V(t)$, due neuroni hanno comunque un’attività correlata, poiché le correnti (il ‘campo’) che ricevono hanno momenti che dipendono nello stesso modo dall’attività collettiva $\nu(t)$ (la ‘media’, da cui ‘approssimazione di campo medio’). Questo non contraddice l’ipotesi di indipendenza invocata sopra, perché tale indipendenza si intende *condizionata* ai momenti (variabili nel tempo) dell’input sinaptico.

Nell’ambito della teoria dei processi stocastici citati, la dinamica di una popolazione neuronale può essere descritta in modo statisticamente completo seguendo l’evoluzione nel tempo della densità $p(v, t)$ di cellule che hanno potenziale di membrana in un piccolo intorno di v al tempo t [2]. A questo scopo è sufficiente ricorrere ad una equazione di continuità che descriva come varia nel tempo il numero di realizzazioni/neuroni $p(v, t) dv$ in un volumetto dv in funzione della divergenza del flusso $S(v, t)$ di realizzazioni che vi transitano. Nel caso unidimensionale l’equazione assume la forma $\partial_t p(v, t) = -\partial_v S(v, t)$. Per i processi stocastici di diffusione di cui stiamo parlando, questa equazione si esplicita nella seguente equazione di Fokker-Planck:

$$\partial_t p(v, t) = \partial_v \left[\left(\frac{v}{\tau} - \mu(t) \right) p(v, t) \right] + \frac{1}{2} \sigma^2(t) \partial_v^2 p(v, t).$$

Per quanto questo problema possa apparire come archetipico in fisica statistica, lo è solo in apparenza. In primo luogo le condizioni al contorno di questa equazione alle derivate parziali devono tenere conto del fatto che i neuroni che emettono un impulso al tempo t ‘scompaiono’ dal conteggio delle realizzazioni in $p(v, t)$ per poi ‘ricomparsire’ nel potenziale di reset H dopo un periodo di inattività corrispondente al periodo refrattario assoluto. Questo porta ad avere un flusso di realizzazioni $S(v, t)$ diverso da zero anche in condizioni stazionarie, elemento questo tipico dei campi non conservativi. Il flusso alla soglia θ ci dice quindi la frazione di neuroni che emettono uno spike per unità di tempo, cioè $\nu(t) = S(\theta, t)$, che per media μ e varianza σ^2 costanti ci dice quanti spike per unità di tempo un generico neu-

rone della rete emetterà, fornendo una relazione di input-output tra la corrente d'ingresso ai neuroni e la loro attività, spesso indicata come funzione di guadagno $\Phi(\mu, \sigma)$, dal tipico aspetto sigmoidale. Alcuni esempi di questa funzione di guadagno sono illustrati in Figura 2d per diversi valori di σ , ovverosia della taglia delle fluttuazioni della corrente sinaptica. È interessante notare come per correnti quasi deterministiche (correnti con piccoli σ) i neuroni non sparano ($\Phi \simeq 0$) fino a quando la corrente media non supera il valore che permette al potenziale di membrana di attraversare la soglia di emissione ($\mu\tau > \theta$). All'aumentare di σ , e quindi delle fluttuazioni della corrente, le frequenze di output inizia a crescere anche per valori medi $\mu\tau$ del potenziale di membrana inferiori alla soglia: è questa la condizione in cui si manifesta lo stato bilanciato tra eccitazione e inibizione.

Altro elemento che rende l'equazione di Fokker-Planck della densità $p(v, t)$ di difficile soluzione è la sua nonlinearità. Poiché la media e varianza della corrente sinaptica dipendono dalla frequenza di firing $\nu(t)$, e questa a sua volta dipende dalla densità $p(v, t)$ che in un intorno di θ determina il flusso $S(\theta, t)$, μ e σ saranno indirettamente funzione di p . L'operatore di Fokker-Planck definito come $Lp \equiv -\partial_v S$ sarà quindi esso stesso funzione della densità: $L = L(p)$. Nel corso degli anni in aggiunta alle integrazioni numeriche di questa equazione, si sono moltiplicati approcci teorici perturbativi, uno dei quali basato sullo sviluppo spettrale dell'operatore di Fokker-Planck L [6]. Per molti aspetti questo approccio ricorda il metodo di Hartree-Fock usato per stimare grandezze fisiche d'interesse in sistemi quantistici, con la differenza principale che le autofunzioni di L usate nello sviluppo, dipendono in questo caso esplicitamente dall'attività $\nu(t)$ della rete stessa. Il vantaggio di questa rappresentazione è quella di esprimere la dinamica di popolazione come la combinazione di 'modi' con una gerarchia di scale temporali data dall'inverso della parte reale degli autovalori dell'operatore L . Nello stato bilanciato il modo più lento è associato ad un autovalore reale che permette di derivare un'espressione relativamente semplice della dinamica di $\nu(t)$:

$$\tau_\nu \frac{d\nu}{dt} = \Phi(\nu) - \nu.$$

L'equazione della frequenza di emissione qui riportata è un'equazione differenziale ordinaria di primo ordine la cui nonlinearità risiede nella funzione di guadagno Φ e nel fatto che la costante di tempo τ_ν varia con l'attività della rete. Nella sua forma, questa equazione delle frequenze ricorda da vicino quella presentata all'inizio degli anni '70 da Hugh R. Wilson e Jack D. Cowan come dinamica fenomenologica delle reti nervose eccitatorie e inibitorie, dando un contributo determinante negli anni a venire allo sviluppo delle neuroscienze computazionali; notiamo però che nelle equazioni di Wilson-Cowan la scala di tempo per l'evoluzione di ν è un parametro fenomenologico assegnato arbitrariamente.

Le soluzioni stazionarie dell'equazione precedente per la dinamica di ν corrispondono alla condizione di autoconsistenza $\nu_0 = \Phi(\nu_0)$: in approssimazione di campo medio lo stato stazionario identifica la condizione in cui la frequenza media di sparo del generico neurone uguaglia quella dei suoi neuroni presinaptici. Se il punto fisso ν_0 della dinamica di ν identificato dalla condizione di autoconsistenza è stabile, uno scostamento temporaneo da ν_0 (per esempio dovuto ad un piccolo stimolo) è seguito da un rilassamento verso ν_0 , la cui componente lenta è governata dalla dinamica di campo medio sopra illustrata: ν_0 è quindi un *attrattore* della dinamica. In certe condizioni (essenzialmente, forte auto-eccitazione e conseguentemente una funzione di guadagno Φ fortemente nonlineare) l'equazione di autoconsistenza può avere più soluzioni (punti fissi) ν_0 (nel caso unidimensionale tre soluzioni, di cui due stabili ed una instabile nel mezzo). In questo caso, una perturbazione abbastanza forte a partire da un punto fisso stabile può spingere il sistema a 'scavalcare' il punto fisso instabile e rilassare nell'altro punto fisso stabile. Illustriamo nel seguito questo scenario in un contesto specifico.

Primitive computazionali più complesse: prendere decisioni

Prendere una *decisione* è, nell'accezione comune, sinonimo di compiere una scelta, e all'interno di questa definizione generale si può far rientrare gran parte del comportamento volontario (auto-

generato o innescato da stimoli esterni), dalla decisione di allontanare la mano da una fiamma a quella riguardante l'opportunità di un investimento finanziario. Ad un simile livello di generalità, la *percezione* indica qualunque processo fisico-fisiologico che converte segnali provenienti dal mondo esterno in specifici pattern di attività nervosa nel cervello.

La ricerca sulle basi neuronali dei processi di decisione copre un orizzonte molto ampio, e arriva a intersecare questioni psicologiche riguardanti l'attribuzione di valore alle opzioni disponibili, e perfino filosofiche riguardo al libero arbitrio. Qui ci limitiamo qui all'ambito ristretto della *decisione percettiva*, nella cui descrizione teorica molti dei concetti che abbiamo descritto giocano un ruolo importante.

In molte circostanze, la percezione avviene in modo quasi (o del tutto) inconsapevole, e semplici decisioni conseguenti sono affini a dei riflessi e non ingaggiano funzioni cognitive. A volte però l'ambiguità degli stimoli percettivi, e/o la criticità delle scelte conseguenti, richiedono invece l'elaborazione di processi di decisione.

Stiamo guidando nel traffico, di sera e sotto la pioggia; ad un tratto vediamo con la coda dell'occhio qualcosa, che potrebbe segnalare un imminente ostacolo e suggerire di frenare. Non vogliamo rischiare di investire qualcuno o qualcosa, ma non vogliamo rischiare inutilmente di sbandare o farci tamponare con una frenata di emergenza se si tratta di un falso allarme: dobbiamo decidere, e dobbiamo farlo sulla base della convinzione che riusciamo a farci sulla natura del potenziale ostacolo, e di un vincolo temporale stretto. La decisione percettiva è il processo attraverso il quale accumuliamo informazione (inclusa quella sensoriale e quella relativa al passaggio del tempo) per compiere una scelta.

La modellistica dei processi di decisione ha una lunga storia [5], e modelli astratti di questo processo sono stati proposti già negli anni '60. Parte della ricerca in questo ambito, di cui non ci occuperemo, riguarda ad esempio la formalizzazione dei principi di ottimalità delle decisioni in relazione al livello di incertezza sullo stato del mondo esterno (modelli Bayesiani).

In seguito ad una serie di importanti esperimenti negli anni '90, in cui si sono messi in luce alcuni meccanismi neurofisiologici della decisio-

ne percettiva, a partire dagli anni 2000 un ricco filone di ricerca teorica si è occupato della costruzione di modelli miranti a riprodurre, nella dinamica di popolazioni neuronali interagenti, alcune caratteristiche sia dell'attività neuronale osservata che delle misure psicofisiche effettuate nelle stesse condizioni.

I celebri esperimenti citati utilizzano un ingegnoso paradigma basato sui 'cinetogrammi a puntini' (traduzione italiana poco usata di 'random dots kinetograms'): nella versione più semplice, il soggetto guarda uno schermo sul quale sono proiettati molti puntini luminosi in movimento, parte verso sinistra e parte verso destra, e lo sperimentatore controlla la percentuale ('coerenza') di puntini che si muovono nella stessa direzione. Il compito consiste nel comunicare (attraverso un brusco cambiamento della direzione dello sguardo - 'saccade') la decisione percettiva del soggetto sulla direzione verso la quale si muove la maggioranza dei puntini (Figura 3a). È chiaro che per elevata coerenza il compito è facile, per il 50% di coerenza il soggetto può solo tirare a caso, e livelli intermedi di coerenza modulano la difficoltà del compito. In due versioni base del protocollo sperimentale, il soggetto è lasciato libero di prendere e comunicare la decisione appena si sente pronto a farlo, oppure lo sperimentatore fissa un tempo al quale la decisione deve comunque essere presa. Dal punto di vista psicofisico, tipicamente si misurano, per diversi livelli di coerenza, il tasso di errori ed il tempo di decisione (nella prima versione del compito), o il tasso di errori in funzione del tempo fissato per la decisione (nella seconda versione del compito) (Figura 3b).

Gli esperimenti identificarono nell'area 'intraparietale laterale' (LIP) popolazioni di neuroni la cui attivazione si correla in modo non ambiguo alla presa di decisione (nel senso che è possibile dissociare tale attività sia dalla percezione della direzione di movimento che dalla codifica dell'atto motorio attraverso il quale la decisione viene comunicata). Questa attività, mediata su molte ripetizioni del protocollo sperimentale, appariva crescere in modo più o meno costante dalla comparsa dello stimolo alla presa di decisione, e raggiungeva un valore fissato al momento della decisione (il che implica una crescita più ripida dell'attività per decisioni semplici e rapide, meno

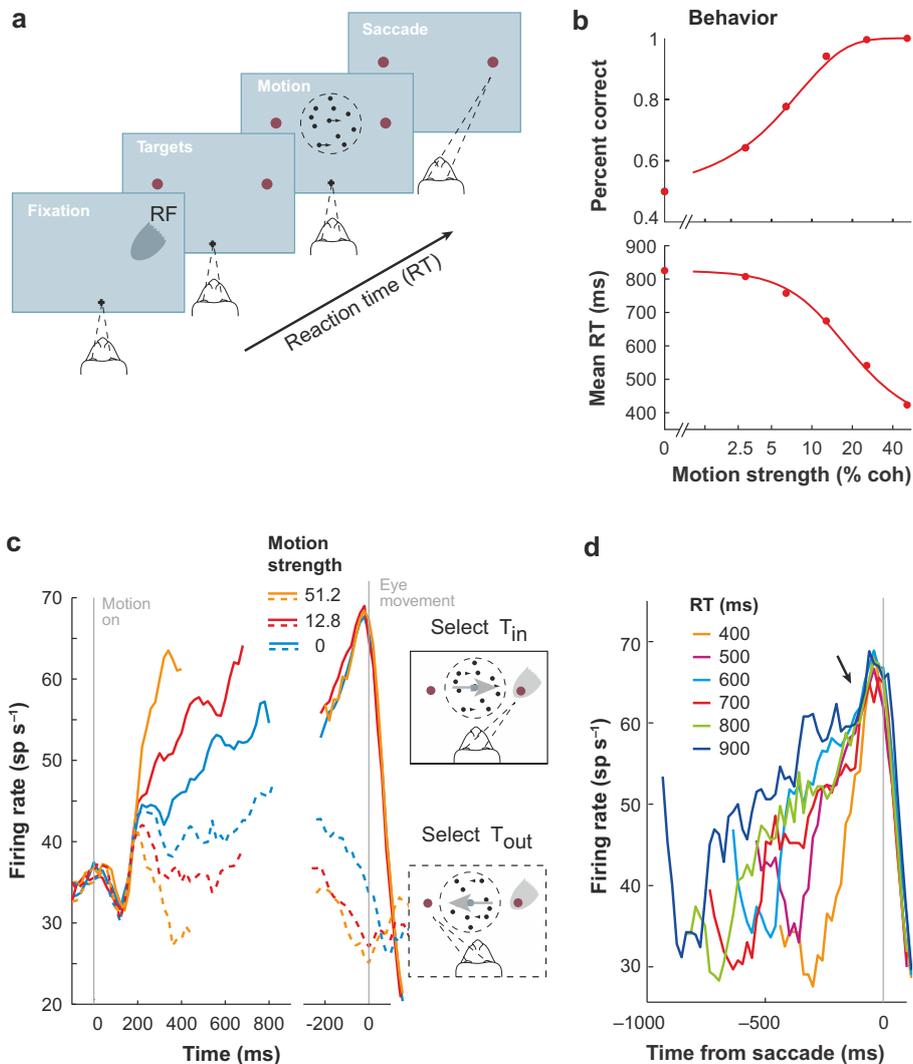


Figura 3: (a) Compito di discriminazione della direzione del moto dei puntini. La difficoltà del compito è determinata dalla frazione di puntini che si muovono coerentemente sullo schermo ('% coherence'). Il tempo di reazione (RT) è il tempo che la scimmia impiega a prendere la decisione di muovere gli occhi nella direzione del moto prevalente dei puntini. (b) Effetto della difficoltà dello stimolo sulla accuratezza (percentuale di trial corretti) e sul tempo di decisione (RT medio). (c) Risposta dei neuroni registrati dalla corteccia intraparietale laterale (LIP). La frequenza di sparo media da 54 neuroni è illustrata per diversi livelli di difficoltà (coerenza del moto), e prendendo come tempo di riferimento l'inizio del moto dei puntini (sinistra) e l'inizio dei movimenti oculari (destra). L'attività cresce se il moto dei puntini è quello preferito dai neuroni registrati (linea continua), mentre decresce nell'altro caso (linee tratteggiate). (d) Stesso grafico di (c) ma con le medie raggruppate per RT. Pannelli adattati da [7].

ripida per decisioni difficile e lente) (Figura 3c).

Coerentemente con l'intuizione e con modelli astratti precedenti, il quadro suggeriva un processo di accumulo di informazione (tanto più lento e lungo quanto più lo stimolo è ambiguo), fino a che il raggiungimento di una soglia rende possibile la decisione (Figura 3d).

Nel modello proposto da X.-J. Wang nel 2002 [8], ciascuna delle due decisioni disponibili ('destra' o 'sinistra') è rappresentata dall'attrattore dinamico di una popolazione eccitatoria ricorrente (nel senso spiegato alla fine della sezione precedente), e il sistema è composto da due popolazioni ricorrenti eccitatorie i cui attrattori codificano le due scelte 'selettive', una popolazione eccitatoria non direttamente coinvolta nella rappresentazione delle due scelte e da una popolazione di neuroni inibitori (Figura 4a). Le popolazioni eccitatorie forniscono input alla popolazione inibitoria, che a sua volta fornisce ad

esse un feedback inibitorio. Una configurazione di questo tipo realizza un meccanismo competitivo 'winner-take-all': se l'attività di una delle due popolazioni selettive supera quella dell'altra in misura superiore ad una soglia, la reazione inibitoria amplifica ulteriormente questa differenza, finché si raggiunge uno stato stabile in cui la popolazione eccitatoria 'vincente' e quella 'perdente' si attestano rispettivamente su uno stato di alta e bassa attività. Gli attrattori della dinamica complessiva di questo sistema sono dunque i due stati corrispondenti alla 'vittoria' dell'una o dell'altra popolazione.

Nella descrizione del processo di decisione, il sistema parte da una condizione simmetrica in cui entrambe le popolazioni eccitatorie sono in uno stato di bassa attività (attività spontanea). L'accensione dell'input visivo corrisponde all'attivazione di due input (rumorosi) a ciascuna delle popolazioni eccitatorie corrispondenti la cui

intensità media codifica il livello di coerenza per le due direzioni di movimento dei puntini. Nel caso massimamente ambiguo (50% di coerenza) la media dei due input sarà uguale, l'attività delle due popolazioni eccitatorie crescerà in media in modo uguale, guidata dalla risposta nonlineare di auto-eccitazione. La crescita sarà temperata, ma in modo uguale, dalla reazione inibitoria. Data la componente stocastica dell'input, l'attività crescente corrisponderà a fluttuazioni di dimensione crescente tra le attività delle due reti eccitatorie, finché una fluttuazione sarà sufficiente a determinare una reazione inibitoria in grado di rompere la simmetria e innescare in modo irreversibile la dinamica competitiva che porterà una delle due popolazioni alla vittoria. In questo caso, il ruolo delle fluttuazioni casuali nel determinare la decisione corrisponde ovviamente al fatto che per coerenza al 50% il soggetto può solo tirare a caso.

Diversi livelli di coerenza corrispondono a diversi sbilanciamenti negli input ricevuti dalle due popolazioni eccitatorie, e questo fornisce una componente deterministica che rompe dall'inizio la simmetria tra le due e produrrà decisioni più rapide e più affidabili al crescere dell'asimmetria.

Prima di illustrare brevemente il tipo più semplice di modello proposto per descrivere la dinamica neuronale della decisione percettiva, sottolineiamo che la crescita graduale dell'attività neuronale media osservata, a cui abbiamo accennato sopra, si sviluppa (a seconda della difficoltà del compito) su scale di tempo che arrivano facilmente al secondo ed oltre. Ottenere una dinamica così lenta in reti i cui neuroni hanno tempi caratteristici molto più brevi è difficile, e questo pone una sfida ai modelli teorici. Nel modello proposto in [8], e in diversi lavori seguenti, si enfatizza il possibile ruolo della componente più lenta della trasmissione sinaptica eccitatoria (che corrisponde all'attivazione dei cosiddetti recettori NMDA), la cui scala di tempo si stima in decine di ms. Accenneremo all'esistenza di opzioni diverse, riguardo sia al meccanismo dinamico di decisione che all'interpretazione degli andamenti osservati negli esperimenti.

Nello spirito di una descrizione di campo medio, identifichiamo lo stato del sistema ad un certo istante con i corrispondenti valori delle fre-

quenze medie di emissione di spike delle popolazioni coinvolte. Per le due popolazioni eccitatorie A e B che codificano la decisione sulla direzione di movimento possiamo scrivere una dinamica di campo medio (Figura 4b):

$$\begin{aligned}\tau_E \dot{\nu}_A &= -\nu_A + \Phi_E(J_{EE}\nu_A - J_{EI}\nu_I + I_A) \\ \tau_E \dot{\nu}_B &= -\nu_B + \Phi_E(J_{EE}\nu_B - J_{EI}\nu_I + I_B)\end{aligned}$$

A queste equazioni si dovrebbero aggiungere un'equazione analoga per l'attività media ν_I della popolazione inibitoria, caratterizzata da una scala di tempo τ_I . Per focalizzarci sulla dinamica delle due popolazioni eccitatorie, ν_A e ν_B (il che ci permette di utilizzare gli strumenti di analisi dei sistemi dinamici sul 'piano di fase') assumiamo che $\tau_I \ll \tau_E$; in altre parole ipotizziamo che la dinamica della popolazione inibitoria sia così veloce rispetto a quella eccitatoria che sulla scala di tempo τ_E essa raggiunga istantaneamente la condizione di equilibrio $\nu_I = \Phi_I(\nu_I, \nu_A, \nu_B)$ rispetto ai valori correnti di ν_A e ν_B . Questo ci permette di sostituire ν_I con $\Phi_I(\nu_I, \nu_A, \nu_B)$ nelle precedenti equazioni. Questa drastica semplificazione è parzialmente giustificata dal fatto che tipicamente i neuroni inibitori generano spike a frequenza più alta di quelli eccitatori, e che gli input sinaptici inibitori sono mediati da recettori veloci.

Spesso si adotta un'altra approssimazione, non necessaria alla riduzione bidimensionale del modello, che semplifica le equazioni e consiste nell'assumere che la funzione di trasferimento Φ_I dei neuroni inibitori sia lineare ($\nu_I = \alpha(\nu_A + \nu_B)$), in modo tale che le equazioni dinamiche del modello si riducono a:

$$\begin{aligned}\tau_E \dot{\nu}_A &= -\nu_A + \Phi_E(\tilde{J}_{EE}\nu_A - J_{EI}\alpha\nu_B + I_A) \\ \tau_E \dot{\nu}_B &= -\nu_B + \Phi_E(\tilde{J}_{EE}\nu_B - J_{EI}\alpha\nu_A + I_B)\end{aligned}$$

con $\tilde{J}_{EE} \equiv J_{EE} - \alpha J_{EI}$. Questo sistema di equazioni può essere analizzato con strumenti standard dello studio dei sistemi dinamici bidimensionali, che permettono di studiare molte caratteristiche della dinamica senza dover ricavare esplicitamente le soluzioni delle equazioni (un ottimo riferimento è il testo di S. Strogatz [9]). In breve: nel piano generato da ν_A e ν_B si tracciano innanzitutto le due cosiddette *nullcline*, luoghi dei punti in cui $\dot{\nu}_A = 0$ e $\dot{\nu}_B = 0$, rispettivamente.

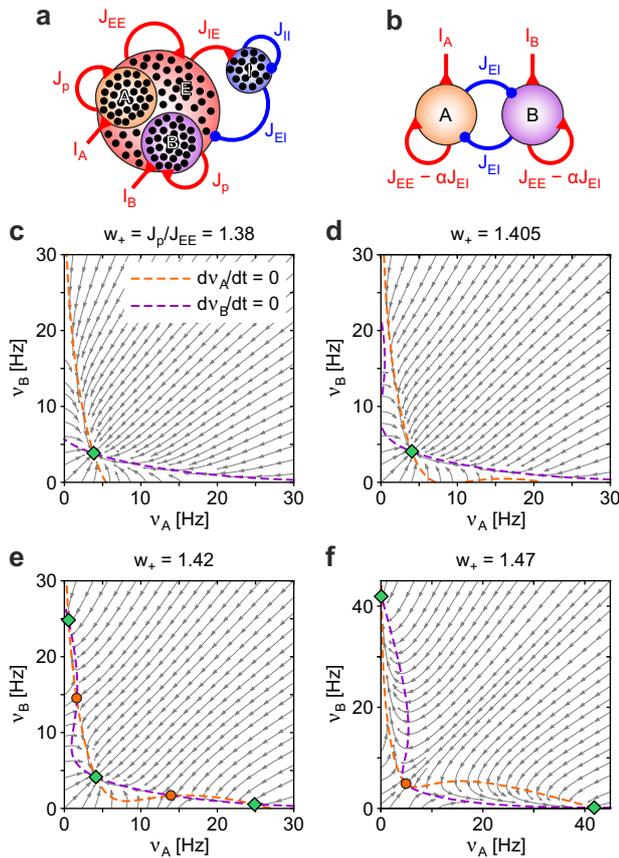


Figura 4: (a) Rete di neuroni IF (puntini neri) modello del modulo decisionale in LIP. I neuroni eccitatori sono divisi in tre gruppi: due codificano la scelta A o B, rispettivamente, e uno non è coinvolto attivamente nel compito (E). I neuroni in A e B ricevono un input I_A e I_B esterno associato alla frazione di punti che si muove nelle rispettive direzioni. Le efficacie sinaptiche $J_{\alpha\beta}$ indicano la forza di accoppiamento tra neuroni delle popolazioni α e β . (b) Riduzione a due popolazioni del modello in approssimazione di campo medio (vedi testo). (c-f) Piani delle fasi (v_A, v_B) delle frequenze di sparo della rete in (b) per diversi livelli di auto-accoppiamento sinaptico $w_+ = J_p/J_{EE}$. Linee tratteggiate, nullcline della dinamica di v_A e v_B . Freccie, linea di flusso del campo di velocità. Cerchi rossi, punti fissi instabili. Rombi verdi, fuochi stabili.

Le nullcline forniscono una sorta di sistema di riferimento sul piano di fase, che lo partiziona in regioni all'interno delle quali le componenti del vettore (\dot{v}_A, \dot{v}_B) hanno segno definito; in questo modo si ricava un quadro complessivo delle direzioni del flusso vettoriale rispetto alle nullcline. I punti di intersezione delle nullcline, se esistono, sono punti fissi della dinamica, in quanto soddisfano la condizione $\dot{v}_A = \dot{v}_B = 0$. Un punto fisso può essere stabile o instabile, tale cioè che, par-

tendo da una condizione iniziale ad esso vicina (analisi lineare), la dinamica riporta il sistema al punto fisso o lo allontana, rispettivamente. È possibile che un punto fisso risulti stabile rispetto ad una direzione di allontanamento e stabile rispetto ad un'altra (queste direzioni sono identificate dagli autovettori della matrice che linearizza la dinamica): si tratta in questo caso di un 'punto di sella'. L'analisi di stabilità descrive la dinamica linearizzata del sistema in prossimità del punto fisso; se questo risulta instabile, la dinamica a lungo termine del sistema va analizzata con altri strumenti. È comunque possibile ricavare dall'analisi lineare alcune informazioni 'anticipatorie' del destino a lungo termine del sistema, ad esempio nella distinzione tra una fuga rettilinea ed una spirale divergente.

Uno dei vantaggi della riduzione bidimensionale di un sistema dinamico è che in due dimensioni il repertorio di regimi dinamici a lungo termine si riduce a due casi (due tipologie di attrattore): punti fissi o 'cicli limite' (orbite chiuse attrattive). Al di là dell'analisi lineare, ogni attrattore (punto fisso o ciclo limite) è associato ad un 'bacino di attrazione', cioè alla regione del piano di fase tale che per una condizione iniziale in esso contenuta la dinamica porta asintoticamente il sistema nell'attrattore.

Le equazioni che abbiamo scritto per il modello di decisione dipendono da diversi parametri, che a loro volta influenzano la forma delle nullcline e quindi la posizione e stabilità dei punti fissi, e la natura degli attrattori della dinamica.

Due grandezze giocano un ruolo chiave in questo scenario: l'entità dell'auto-eccitazione delle due popolazioni selettive, ed il livello di coerenza dello stimolo. Lo studio sistematico del modo in cui la dinamica del sistema cambia qualitativamente al variare di uno o più parametri è la *teoria delle biforcazioni*; senza pretesa di completezza, illustriamo le conseguenze principali di questa analisi. L'auto-eccitazione (quantificata dal potenziamento sinaptico relativo w_+ in Figura 4) determina la nonlinearità della risposta allo stimolo e, secondo quanto discusso alla fine della sezione precedente, il repertorio e la natura dei punti fissi della dinamica di campo medio, a parità di stimolo: in assenza di stimolo, per bassa auto-eccitazione l'unico punto fisso è un attrattore a bassa attività (Figura 4c). Al crescere

di w_+ ad un certo punto (di 'biforcazione pitchfork') compaiono, insieme al precedente punto fisso che rimane stabile, due altri punti fissi asimmetrici, caratterizzati da alta attività di una popolazione selettiva e bassa attività dell'altra (Figura 4e). Un ulteriore aumento di w_+ arriva a destabilizzare il punto fisso di bassa attività, lasciando come unici punti fissi stabili i due attrattori asimmetrici (Figura 4f). A seconda del valore di w_+ , si possono avere diversi comportamenti dinamici al variare dello stimolo, esemplificati in Figura 5 nella quale, per due valori di w_+ , si descrive la rappresentazione del sistema nel piano di fase prima, durante e dopo la stimolazione. Per il valore maggiore di w_+ (Figura 5a), il sistema possiede, come abbiamo visto sopra, un punto fisso simmetrico stabile di bassa attività, che costituisce la condizione pre-stimolazione, e due punti fissi stabili asimmetrici. Durante la stimolazione, di coerenza tale da favorire uno degli attrattori asimmetrici, il paesaggio di fase viene temporaneamente deformato insieme ad i suoi punti fissi, ed il sistema viene spinto verso una versione 'potenziata' dell'attrattore asimmetrico favorito dallo stimolo. Dopo la stimolazione il paesaggio di fase torna alla condizione pre-stimolo, ma il sistema ora rimane 'intrappolato' nell'attrattore asimmetrico. Per un basso valore di w_+ (Figura 5b), il paesaggio pre-stimolo espone un solo punto fisso stabile a bassa attività e, sebbene la stimolazione spinga temporaneamente il sistema verso uno stato asimmetrico, nella fase post-stimolo il sistema rilassa verso la stessa condizione pre-stimolo.

I modelli di decisione percettiva del tipo descritto in [8] si basano sul fatto che l'arrivo dello stimolo (il display dei puntini in movimento) modifica il paesaggio del sistema nel piano di fase in modo da destabilizzare lo stato simmetrico di bassa attività, trasformandolo in una sella. In assenza di rumore e per coerenza nulla il sistema 'scivolerebbe' sul crinale della soglia, continuando a rispettare la simmetria, ma anche una modesta fluttuazione romperebbe la simmetria e dal momento in cui il sistema lascia il crinale della soglia la sua dinamica è determinata dalla pendenza del paesaggio che lo accompagna verso uno dei due attrattori asimmetrici. Per coerenza non nulla, il sistema viene subito allontanato dal crinale della sella, e la traiettoria verso l'at-

trattore asimmetrico è dominata dalla dinamica competitiva descritta e guidata in modo essenzialmente deterministico dalla forma del paesaggio; l'eventuale rumore presente ha un effetto marginale.

Notiamo di passaggio che è possibile stabilire una corrispondenza formale tra i modelli bidimensionali appena descritti e la categoria di modelli che sopra abbiamo definito 'astratti', unidimensionali, essenzialmente basati su un meccanismo di integrazione (con rumore) del segnale di input in cui la decisione corrisponde all'attraversamento di una soglia da parte dell'integratore. In questa corrispondenza, l'innescò del meccanismo winner-take-all del modello bidimensionale corrisponde al raggiungimento della soglia.

Tornando ai modelli bidimensionali, il tempo di decisione, cioè il tempo tra la reazione allo stimolo al raggiungimento dell'attrattore, è influenzato dai tempi caratteristici della trasmissione sinaptica, e come abbiamo accennato è stato originariamente proposto che la scala di tempo, comparativamente lunga, del recettore NMDA, fosse l'elemento chiave per dar conto dei tempi di decisione osservati. È però possibile immaginare, per gli stessi modelli descritti, un diverso meccanismo dinamico, che ha il pregio di assicurare in modo naturale una notevole flessibilità nei tempi di decisione su range molto ampi, e fornisce anche un supporto teorico ad alcune recenti analisi dell'attività neurale durante la decisione. L'idea è semplice: uno stimolo non abbastanza forte da destabilizzare il punto fisso di bassa attività ne diminuirà comunque la stabilità (cioè modificherà comunque il paesaggio, diminuendo la profondità della 'valle' di bassa attività). In assenza di rumore quindi il sistema rimarrebbe ancorato a questo punto fisso, ma il rumore può favorirne la fuga nella valle corrispondente ad uno dei due attrattori asimmetrici. In fisica questo tipo di situazione corrisponde ad un problema noto (*noise-driven transitions*), ed un risultato fondamentale è che la probabilità per unità di tempo di fuga da una valle dipende in modo esponenziale dal rapporto tra l'altezza della barriera che dev'essere superata e la varianza del rumore (equazione di Arrhenius). Nel nostro caso, l'altezza della barriera è la quantità che viene modulata dall'intensità dello stimolo, e si capisce quindi come questa dipendenza esponenziale

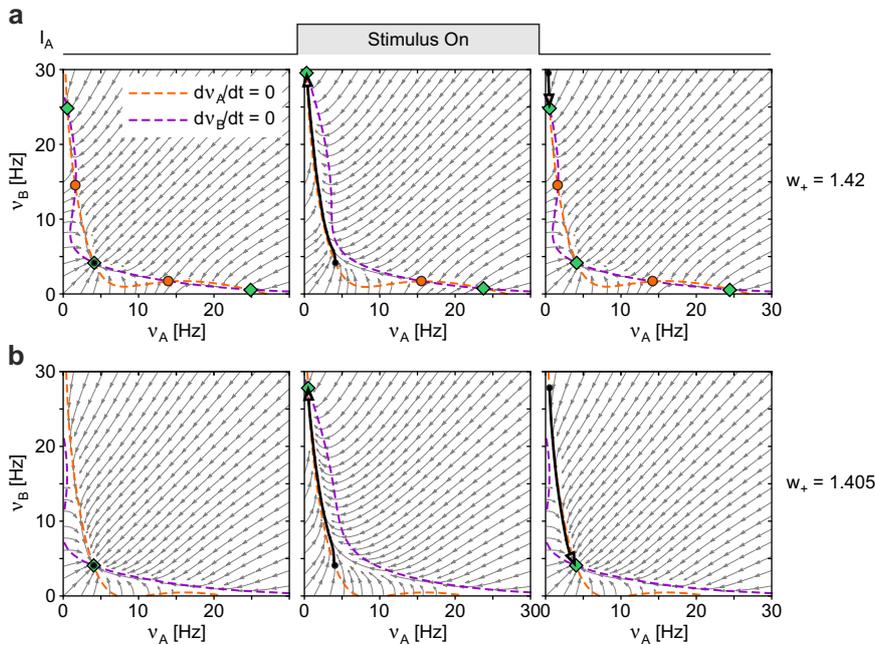


Figura 5: Piani delle fasi (v_A, v_B) della rete in Figura 4b con l'aggiunta della traiettoria da essa seguita (curve con freccia nera). Le tre fasi del compito illustrate sono: i) il periodo che precede l'inizio del compito; ii) il periodo in cui appaiono i puntini in movimento nella direzione A portando ad un incremento della sola I_A ; iii) la fase finale del compito in cui la stimolazione scompare. Le reti illustrate sono: (a) una con forte auto-accoppiamento w_+ (tre attrattori stabili), e (b) l'altra con le sottopopolazioni A e B debolmente accoppiate (il solo attrattore a bassa frequenza ν è presente). Simboli non descritti, come in Figura 4.

consenta di modulare il tempo medio di fuga (il tempo di decisione nel nostro caso) su un range molto ampio anche con modeste variazioni dell'intensità dello stimolo. Non è facile estendere la trattazione analitica classica del problema unidimensionale al caso in due dimensioni (sebbene sia possibile una riduzione approssimata ad un caso unidimensionale in prossimità della biforcazione), e l'opzione descritta è stata studiata in [10] essenzialmente attraverso simulazioni numeriche, caratterizzando la dipendenza della distribuzione dei tempi di decisione dall'intensità dello stimolo e da quella del rumore.

In questo scenario, dunque, il rumore gioca un ruolo costruttivo essenziale. Se da un lato l'inclusione del rumore è del tutto naturale, per le motivazioni discusse sopra, resta da vedere se i dati sperimentali offrono un chiaro supporto all'uno o l'altro scenario.

La questione è articolata e complessa, e non possiamo discuterla qui in dettaglio; ci focalizziamo su un solo aspetto, che riveste un interesse generale al di là del processo di decisione oggetto dei modelli.

Abbiamo ricordato che, dalle registrazioni elettrofisiologiche effettuate durante il processo di decisione, l'attività dei neuroni coinvolti cresce lungo una 'rampa' fino a raggiungere un valore che sembra esprimere una soglia, e che la pendenza della rampa, non la soglia, sembra modulata dalla difficoltà del compito (dalla coerenza). Bisogna però sottolineare che queste 'rampe' ven-

gono chiaramente osservate quando si effettua la media dell'attività neuronale su molte ripetizioni del protocollo sperimentale ('trial'); riconoscerle nell'attività durante una singola ripetizione è molto più difficile. D'altra parte, il meccanismo guidato dal rumore appena descritto non predice certo 'rampe' nell'attività neuronale: questa rimarrebbe bassa e stazionaria finché il sistema non viene spinto da una fluttuazione oltre la barriera, e dopo la fuga aumenterebbe rapidamente fino ad attestarsi intorno al valore elevato corrispondente ad uno dei due attrattori asimmetrici; invece di una rampa, quasi un 'gradino'. Ma essendo ora il processo guidato dal rumore, a parità di input l'istante di tempo in cui il sistema supera la barriera varierà molto (con una distribuzione approssimativamente esponenziale). La 'rampa' osservata negli esperimenti potrebbe essere quindi una conseguenza (in un certo senso, un artefatto) di una media effettuata su attività neuronali che, nel singolo trial, esprimono un andamento 'a gradino'; in [10] viene esemplificata la plausibilità di questo meccanismo (Figura 6). Un esempio di questi due tipi di attività osservabile a livello di singolo trial è stata illustrata non solo in corteccia parietale [11] ma anche nelle cortecce motorie [12]. Decidere tra le due opzioni sulla base dei dati può apparire facile ma non lo è, e il dibattito è ancora acceso (si veda ad esempio [13]).

Notiamo che, indipendentemente dalla formulazione dinamica dei modelli, l'analisi comparati-

va delle due ipotesi si può formulare puramente in termini di analisi statistica del segnale, in cui si quantifica la verosimiglianza dei due modelli statistici (la probabilità condizionata di ottenere il dato osservato, dato il modello statistico descrittivo adottato).

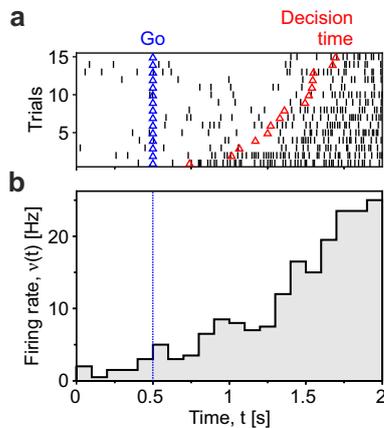


Figura 6: *Dinamica di singolo neurone nella rete in Figura 4a in regime di decisione guidata dal rumore. (a) Spike emessi dallo stesso neurone della rete simulata nei diversi trial. Ogni trattino verticale è uno spike. Simboli verdi e rossi rappresentano l'inizio della stimolazione (blue) e il tempo in cui viene presa la decisione (rosso). (b) Istogramma normalizzato (cioè la frequenza di emissione $\nu(t)$) degli spike emessi ed illustrati in (a). Adattata da [10].*

Si diceva della generalità di alcune implicazioni dei meccanismi dinamici discussi; un esempio stimolante è fornito da una categoria di fenomeni percettivi curiosi e per certi versi paradossali, esemplificati dalla 'rivalità binoculare' e dalla percezione degli stimoli ambigui. Di questi ultimi conosciamo quasi tutti degli esempi (la 'giovane-vecchia, il cubo di Necker ecc.). La rivalità binoculare è il fenomeno percettivo per cui se due immagini vengono mostrate, simultaneamente e separatamente, ai due occhi, la nostra percezione visiva cosciente alterna (e in modo apparentemente casuale) tra la piena 'visione' dell'una o dell'altra immagine (invece di percepire, come pure verrebbe naturale supporre, un mescolamento delle due immagini). Si osserva che gli intervalli di tempo di 'dominanza' di ciascuna delle due percezioni coscienti hanno una distribuzione statistica che mantiene invariate alcune caratteristiche anche per condizioni sperimentali molto diverse, a suggerire che un meccanismo comune potrebbe esserne alla base [14]. Si noti

che la variabilità statistica di questi processi si manifesta a fronte di uno stimolo costantemente presente e invariato, il che quasi irresistibilmente spinge a pensare ad un meccanismo stocastico interno.

Questi fenomeni percettivi sono studiati da decenni in psicofisica, e diversi modelli sono stati formulati per tentarne una spiegazione; senza tentarne una rassegna, per la nostra discussione è interessante osservare che è possibile spiegare diverse caratteristiche statistiche dell'alternanza percettiva attraverso modelli teorici che di fatto estendono il modello di decisione basato sulle fluttuazioni che abbiamo appena discusso [15],[14]; in effetti, l'alternanza percettiva può essere vista come una 'decisione involontaria'.

Sottolineiamo infine alcune assunzioni/limitazioni importanti, sia degli esperimenti che dei modelli di decisione percettiva: *i)* la scelta è obbligatoria, il repertorio di scelte è fissato e le scelte disponibili sono due (*two-alternative forced choice task*; questa limitazione può apparire banale ma non lo è); *ii)* l'input alle due popolazioni che codificano la decisione è completamente segregato in partenza rispetto alla direzione di movimento; *iii)* si assume che la rappresentazione neuronale delle decisioni disponibili sia pre-formata e il modello non si occupa di come essa si possa generare dall'esperienza (fedeli all'impegno di non trattare i processi di apprendimento, non discuteremo questo punto).

Conclusioni

Lo sviluppo di modelli fisico-matematici dell'attività nervosa ha una lunga storia (il fondamentale modello di Hodgkin-Huxley della generazione dello spike è del 1952), ed il ruolo dei fisici e dei matematici in questo ambito ha costruito da subito un germe di interdisciplinarietà nello studio del cervello. D'altra parte, nell'ambito delle (allora giovani) discipline psicologiche, da subito si avvertì l'utilità metodologica di astrarre uno schema concettuale dei rapporti funzionali tra diverse componenti del sistema nervoso, discolorato dal dettaglio biologico (è il caso, per esempio, del 'sistema nervoso concettuale', a cui si riferirono Hebb, Skinner ed altri, e forse lo

schema delineato da Freud nel 'Progetto di una psicologia' ne è un antesignano).

I progressi degli ultimi decenni, però, stanno mutando qualitativamente il quadro. Da una situazione in cui il flusso culturale tra biologia e 'scienze dure', pur fecondo, manteneva i due ambiti nettamente distinti, con le seconde 'a servizio' della prima per sistematizzare attraverso modelli quantitativi la quantità di dati empirici raccolti, si stanno ora consolidando nuovi profili scientifici, offerte formative e contesti di ricerca, grazie ad una osmosi che sta creando una 'scienza del cervello' multi-disciplinare.

Una conseguenza di questa trasformazione è il fatto che lo studio del cervello pone nuovi problemi alle scienze fisiche, stimolando lo sviluppo di nuovi strumenti teorici e computazionali. Un aspetto di questo, che speriamo emerga almeno in parte da quanto scritto, è la natura multi-scala dello studio del cervello.

La fisica si è sempre avvantaggiata della possibilità di separare le scale nella descrizione di un fenomeno: la dinamica dei fluidi necessaria a progettare la chiglia di una barca non tiene conto della struttura molecolare dell'acqua, se non nella determinazione di alcuni parametri che entrano nella descrizione macroscopica del problema. Analogamente, è spesso importante poter separare le scale di tempo in modo tale che, in un sistema i cui gradi di libertà evolvono su scale di tempo molto diverse, si possa descrivere la dinamica dei gradi di libertà veloci nel background 'statico' di quelli lenti. I problemi fisici in cui la separazione di scale non è possibile sono notoriamente difficili, e la meccanica statistica dei fenomeni critici ne è un esempio.

Lo studio teorico del cervello aggiunge a questo quadro diversi livelli di complessità: diverse scale spaziali e temporali sono accoppiate in modo difficilmente separabile, ma non costituiscono un continuo (si pensi alla modularità spaziale di cui abbiamo discusso), e la molteplicità di scale non è realisticamente trattabile con gli strumenti della meccanica statistica dei fenomeni critici, adatti a studiare fluttuazioni multi-scala in sistemi omogenei. Il sistema in generale non è assimilabile ad un sistema fisico, per quanto complesso, all'equilibrio, e la fisica dei sistemi a molti corpi fuori dall'equilibrio è molto più difficile da trattare (un fisico trova facilmente libri e artico-

li di 'meccanica statistica di non-equilibrio', ma spesso il titolo apparentemente generale corrisponde ad una scelta idiosincratice di metodi adatti a situazioni specifiche - qualcuno ha scritto che definire la 'meccanica statistica di non-equilibrio' è un po' come definire la 'zoologia dei non-elefanti'). In relazione all'osservazione precedente, il cervello è chiaramente un 'sistema aperto', le cui interazioni col mondo esterno sono molto più difficili da schematizzare teoricamente di un 'bagno termico' o di un campo magnetico variabile.

Queste osservazioni suggeriscono che, dal punto di vista dello studio teorico, il cervello si colloca in una 'terra di mezzo'. Con ammirevole preveggenza, Warren Weaver nel 1948 [16] la identificò riflettendo sulle sfide inedite che i sistemi biologici complessi avrebbero posto alla fisica, e distingueva problemi 'semplici' (non vuol dire facili da risolvere, ma trattabili con astrazioni che semplificano la logica dei modelli, come nel calcolo deterministico delle orbite dei pianeti), e problemi di 'complessità disorganizzata' (per esempio quelli in cui le fluttuazioni casuali sono importanti ma si possono assumere omogenee nello spazio e nel tempo, e si riesce a descrivere il comportamento del sistema attraverso l'evoluzione delle distribuzioni di probabilità di poche osservabili). 'In mezzo', appunto, stanno i problemi di 'complessità organizzata'. Nella classificazione di Weaver, l'approccio teorico al cervello ricade senz'altro nell'ultima categoria.

Per i motivi che abbiamo sopra riassunto, la costruzione di modelli realistici di attività del cervello è difficile, e la loro soluzione analitica supera spesso le nostre capacità. Questo fatto motiva (come, di nuovo, lo stesso Weaver aveva predetto) un ruolo centrale delle simulazioni al computer dei modelli. Nello studio teorico del cervello gli approcci computazionali hanno assunto un ruolo crescente (anche grazie, ovviamente, alle crescenti potenze di calcolo disponibili), ed il loro ruolo euristico è oggi spesso assimilato a veri e propri 'esperimenti *in silico*'; ne è testimonianza il fatto che gli sviluppi computazionali sono centrali in diverse iniziative progettuali multi-nazionali; in particolare nello Human Brain Project, iniziativa europea decennale a cui l'Italia dà un notevole contributo, che dopo la sua conclusione nel 2023 si evolverà in una in-

infrastruttura europea (EBRAINS) che metterà a servizio della comunità scientifica il ricco repertorio di strumenti computazionali sviluppati nel progetto [17]. È auspicabile, e probabile, che questi sviluppi accelereranno in modo importante l'evoluzione della comunità multi-disciplinare impegnata nella affascinante frontiera costituita dallo studio del cervello.



[1] V. B. Mountcastle: *The columnar organization of the neocortex*, Brain, 120 (1997) 101.

[2] P. Del Giudice and M. Mattia: *Stochastic population dynamics of spiking neurons* In E. Korutcheva and R. Cuenno, editors, *Advances in Condensed Matter and Statistical Physics*, chapter 10, pages 125–153. Nova Science Publishers, New York, (2004).

[3] D. J. Amit and N. Brunel. *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex*. Cereb. Cortex, 7 (1997) 237.

[4] C. van Vreeswijk and H. Sompolinsky. *Chaos in neuronal networks with balanced excitatory and inhibitory activity*. Science, 274 (1996) 1724.

[5] R. Ratcliff, P. L. Smith, S. D. Brown, and G. McKoon: *Diffusion decision model: Current issues and history* Trends in cognitive sciences, 20 (2016) 260.

[6] M. Mattia and P. Del Giudice. *Population dynamics of interacting spiking neurons*. Phys. Rev. E, 66 (2002) 051917.

[7] J. I. Gold and M. N. Shadlen: *The neural basis of decision making*. Annu. Rev. Neurosci. 30 (2007) 74.

[8] X.-J. Wang. *Probabilistic decision making by slow reverberation in cortical circuits*. Neuron, 36 (2002) 955.

[9] S. H. Strogatz: *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press, (2000).

[10] D. Martí, G. Deco, M. Mattia, G. Gigante, and P. Del Giudice. *A fluctuation-driven mechanism for slow decision processes in reverberant networks*. PloS one, 3 (2008) e2534.

[11] T. A. Engel, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, and K. Boahen. *Selective modulation of cortical state during spatial attention*. Science, 354 (2016) 1140.

[12] M. Mattia, P. Pani, G. Mirabella, S. Costa, P. Del Giudice, and S. Ferraina. *Heterogeneous attractor cell assemblies for motor planning in premotor cortex*. J. Neurosci., 33 (2013) 11155.

[13] D. M. Zoltowski, K. W. Latimer, J. L. Yates, A. C. Huk, and J. W. Pillow. *Discrete stepping and nonlinear ramping dynamics underlie spiking responses of lip neurons during decision-making* Neuron, 102 (2019) 1249.

[14] R. Cao, A. Pastukhov, M. Mattia, and J. Braun. *Collective activity of many bistable assemblies reproduces characteristic dynamics of multistable perception*. Journal of Neuroscience 36 (2016) 6957.

[15] G. Gigante, M. Mattia, J. Braun, and P. Del Giudice. *Bistable perception modeled as competing stochastic integrations at two levels*. PLoS Comput Biol, 5 (2009) e1000430.

[16] W. Weaver. *Science and complexity*. Am. Sci., 36 (1948) 536.

[17] K. Amunts, C. Ebell, J. Muller, M. Telefont, A. Knoll, and T. Lippert. *The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain*. Neuron, 92 (2016) 574.



Paolo Del Giudice: è Primo Ricercatore dell'Istituto Superiore di Sanità a Roma e Professore incaricato di Reti Neurali presso il Dipartimento di Fisica dell'Università di Roma³. Si occupa dello studio teorico di modelli neuronali e della loro inferenza dai dati elettrofisiologici.

Maurizio Mattia: è Ricercatore dell'Istituto Superiore di Sanità a Roma e Professore incaricato di Neural Networks presso il Dipartimento di Fisica dell'Università di Roma 'Sapienza'. Si occupa dello studio teorico delle reti neuronali e dell'analisi dell'attività nervosa in collaborazione con laboratori sperimentali nell'ambito dello 'Human Brain Project' (finanziamento "EU H2020 Research and Innovation", grant n. 945539 - HBP SGA3).