

3. Redes neuronales fuera del equilibrio

3.1. Introducción

- El modelo de Hopfield de memoria asociativa así como gran parte de los modelos basados en él, consiste de un conjunto de N neuronas binarias, cuya actividad puede tomar uno de los valores $s_x = \pm 1$ en cada instante de tiempo. En estos modelos, en general, la actividad evoluciona con el tiempo mediante algún tipo de proceso que puede ser determinista o estocástico (correspondiendo cada uno, respectivamente, a una situación en la que tenemos temperatura cero o temperatura finita distinta de cero). Las neuronas, que podemos imaginar (pero no es esencial) localizadas en los nudos de una red d -dimensional, Λ_d , están interactuando entre sí de acuerdo con una determinada regla de aprendizaje como, por ejemplo, la regla de Hebb [25]. Esto es, la neurona que ocupa el nudo que está en la posición \mathbf{x} está conectada con la neurona que ocupa el nudo cuya posición es \mathbf{y} mediante una sinapsis cuya intensidad es

$$J_{xy} = \frac{1}{N} \sum_{\mu=1}^P \xi_x^\mu \xi_y^\mu, \quad (74)$$

donde $\{\xi_x^\mu = \pm 1; \mathbf{x} \in \Lambda_d\} \equiv \xi^\mu$ representa $\mu = 1, \dots, P$ patrones almacenados. En estos modelos, los valores J_{xy} son independientes del tiempo. Esto suele ser interpretado de dos formas, bien considerando que las intensidades sinápticas han sido determinadas durante un proceso previo de aprendizaje (*plasticidad sináptica*), consistente en que, cada vez que el sistema memoriza un determinado patrón ξ^μ , las intensidades sinápticas se modifican en la forma

$$J_{xy}^{new} = J_{xy}^{old} + \frac{\xi_x^\mu \xi_y^\mu}{N}, \quad (75)$$

o bien considerando que existen dos escalas temporales, bien definidas y diferentes, para la evolución de las sinapsis y de las neuronas, respectivamente, la primera mucho más grande que la segunda.

En el modelo de Hopfield interesa evaluar la importancia que tienen los detalles microscópicos en las propiedades macroscópicas de la red, y así han surgido una serie de modelos, modificaciones del original, que tratan de explicar características esenciales de los sistemas biológicos; cf. [26, 27, 28, 29, 30, 31]. En este sentido no es realista el considerar sólo las variaciones temporales de las sinapsis que ocurren en el

proceso de aprendizaje. Una situación típica en los sistemas biológicos [36, 37, 38, 39, 40, 41, 43, 44, 32] consiste en que las intensidades sinápticas varían con el tiempo, también, en una escala temporal mucho menor que aquélla en la que tiene lugar la evolución de la actividad de las neuronas. Así, en un intervalo de tiempo Δt del orden del tiempo que transcurre entre la generación de dos potenciales de acción consecutivos, el estado de las sinapsis puede cambiar de forma sustancial, debido, por ejemplo, a las variaciones en el número de vacuolas en los botones sinápticos que inducen fluctuaciones locales en la concentración de neurotransmisores [46, 45]. La motivación para considerar variaciones locales de las intensidades sinápticas J_{xy} en los modelos de redes de neuronas, también tiene su origen en el hecho de que las neuronas biológicas están, en la práctica, conectadas por más de una sinapsis [47], cada una teniendo diferente naturaleza, es decir químicas o eléctricas, las cuales transmiten el potencial de acción a diferente velocidad. En este capítulo se presenta un modelo de red de neuronas [33] que, además de tener en cuenta la *lenta* plasticidad sináptica debida al aprendizaje, tiene en cuenta fluctuaciones locales, relativamente rápidas, de las intensidades sinápticas. Motivado por lo que ocurre en los sistemas biológicos, las intensidades J_{xy} varían aleatoriamente con el tiempo de tal forma que su promedio sobre Δt , \bar{J}_{xy} , tiene el valor correspondiente a una determinada regla de aprendizaje. Por ejemplo, asumiendo que dicha regla de aprendizaje es la regla de Hebb, J_{xy} constantemente *visita* al azar, durante toda la evolución, los patrones memorizados, tomando así valores instantáneos $J_{xy} = \frac{\xi_{\mathbf{x}}^{\mu} \xi_{\mathbf{y}}^{\mu}}{N}$ con una determinada probabilidad, de forma que \bar{J}_{xy} viene dado por (74).

Es interesante ver la influencia de tales fluctuaciones en propiedades emergentes del sistema. Vamos a considerar distribuciones específicas para las fluctuaciones que producen algunos resultados exactos, como hamiltonianos efectivos. Partiendo de estos hamiltonianos, se puede usar la técnica de *réplicas* para obtener resultados explícitos.

3.2. Definición del modelo

Consideremos un sistema constituido por N neuronas con actividad representamos por $s_{\mathbf{x}}$, conectadas entre sí mediante sinapsis cuya intensidad denotamos mediante la variable J_{xy} . Denotemos por $\mathbf{s} \equiv \{s_{\mathbf{x}} = \pm 1; \mathbf{x} \in \Lambda_d\}$ y $\mathbf{J} \equiv \{J_{xy} \in \mathfrak{R}; \mathbf{x}, \mathbf{y} \in \Lambda_d\}$ una configuración de neuronas y de sinapsis, respectivamente. La probabilidad del estado (\mathbf{s}, \mathbf{J}) en el instante de tiempo t varía de forma estocástica de acuerdo con la

ecuación maestra:

$$\partial_t P_t(\mathbf{s}, \mathbf{J}) = \sum_{\mathbf{s}', \mathbf{J}'} [P_t(\mathbf{s}', \mathbf{J}') c(\mathbf{s}', \mathbf{J}' \rightarrow \mathbf{s}, \mathbf{J}) - P_t(\mathbf{s}, \mathbf{J}) c(\mathbf{s}, \mathbf{J} \rightarrow \mathbf{s}', \mathbf{J}')], \quad (76)$$

donde la doble sumatoria está extendida a todas las configuraciones de neuronas y de sinapsis, respectivamente. Esto es, consideramos que determinados agentes, entre los que se incluye un baño térmico a temperatura T , inducen cambios $(\mathbf{s}, \mathbf{J}) \rightarrow (\mathbf{s}', \mathbf{J}')$ cuya probabilidad por unidad de tiempo (*rate*) es

$$c(\mathbf{s}, \mathbf{J} \rightarrow \mathbf{s}', \mathbf{J}') = p c_1(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{J}) \delta_{\mathbf{J}, \mathbf{J}'} + (1 - p) c_2(\mathbf{J} \rightarrow \mathbf{J}') \delta_{\mathbf{s}, \mathbf{s}'}, \quad (77)$$

donde $\delta_{X, X'}$ representa la función delta de Kronecker,

$$c_1(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{J}) = \sum_{\mathbf{y}} \varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{y}) \delta_{s_{\mathbf{y}}, -s'_{\mathbf{y}}} \prod_{\mathbf{x} \neq \mathbf{y}} \delta_{s_{\mathbf{x}}, s'_{\mathbf{x}}} \quad (78)$$

y

$$c_2(\mathbf{J} \rightarrow \mathbf{J}') = \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} \varpi(J_{xy} \rightarrow J'_{xy}) (1 - \delta_{J_{xy}, J'_{xy}}) \prod_{\substack{\mathbf{u}, \mathbf{v} \\ (\mathbf{u}, \mathbf{v}) \neq (\mathbf{x}, \mathbf{y})}} \delta_{J_{uv}, J'_{uv}}, \quad (79)$$

donde las sumas y productos se extienden a todo el sistema con las restricciones mostradas explícitamente en cada caso. Esto es, la evolución temporal tiene lugar mediante la superposición de dos clases diferentes de procesos elementales (sólo uno en la unidad de tiempo), que son cambios $s_{\mathbf{x}} \rightarrow -s_{\mathbf{x}}$ que ocurren con *velocidad* $p \varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x})$ para una configuración de sinapsis dada, y cambios $J_{xy} \rightarrow J'_{xy}$ que ocurren con *velocidad* $(1 - p) \varpi(J_{xy} \rightarrow J'_{xy})$, independiente de la configuración de neuronas que en ese instante hay en el sistema. Aquí el parámetro p mide la probabilidad del proceso de evolución de las neuronas con respecto al de evolución de las sinapsis. En estas condiciones, haciendo uso de las expresiones (77)-(79) e introduciéndolas en (76), podemos escribir la ecuación maestra (76) en la forma

$$\begin{aligned} \partial_t P_t(\mathbf{s}, \mathbf{J}) = & p \sum_{\mathbf{x}} [\varpi_{\mathbf{J}}(\mathbf{s}^{\mathbf{x}}; \mathbf{x}) P_t(\mathbf{s}^{\mathbf{x}}, \mathbf{J}) - \varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x}) P_t(\mathbf{s}, \mathbf{J})] \\ & + (1 - p) \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} \sum_{J'_{xy}} [\varpi(J'_{xy} \rightarrow J_{xy}) P_t(\mathbf{s}, \mathbf{J}^{xy}) - \varpi(J_{xy} \rightarrow J'_{xy}) P_t(\mathbf{s}, \mathbf{J})], \end{aligned} \quad (80)$$

donde \mathbf{s}^x representa la configuración de neuronas obtenida de \mathbf{s} después de cambiar $s_x \rightarrow -s_x$, y \mathbf{J}^{xy} es la configuración de sinapsis que se obtiene a partir de \mathbf{J} después de que haya tenido lugar el cambio $J_{xy} \rightarrow J'_{xy}$. Por simplicidad, a partir de este momento nos restringimos a la situación particular en la que se tiene

$$\varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x}) = \Psi [2\beta s_x h_x(\mathbf{s}, \mathbf{J})], \quad (81)$$

donde $\beta = (k_B T)^{-1}$, k_B es la constante de Boltzmann, y

$$h_x(\mathbf{s}, \mathbf{J}) = \sum_{y \neq x} J_{xy} s_y - \theta_x \quad (82)$$

es el campo local asociado a la neurona que ocupa la posición \mathbf{x} con θ_x la energía umbral necesaria para activar esa neurona. En (81) Ψ es una función arbitraria con la única condición de que verifique $\Psi(0) = 1$, $\lim_{X \rightarrow +\infty} \Psi(X) = 0$, y $\Psi(X) = \Psi(-X) \exp(-X)$; esta última condición se conoce con el nombre de condición de balance detallado. Ejemplos de este tipo de funciones son

$$\Psi(X) = \begin{cases} e^{-\frac{X}{2}}, & \text{rate } V \\ 2(1 + e^X)^{-1}, & \text{rate } K \\ \min\{1, e^{-X}\}, & \text{rate } M. \end{cases} \quad (83)$$

En el caso de ciertos límites bien definidos, el considerar estas elecciones simplificatorias, es decir (81), (82), y la independencia de ϖ ($J_{xy} \rightarrow J'_{xy}$) en \mathbf{s} , reduce el modelo a situaciones más familiares, haciendo que el presente modelo sea una generalización de las mismas, válido para estudiar otras situaciones hasta ahora no tenidas en cuenta. Así, para $p \equiv 1$, $c(\mathbf{s}, \mathbf{J} \rightarrow \mathbf{s}', \mathbf{J}') = c_1(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{J}) \delta_{\mathbf{J}, \mathbf{J}'}$, y (80), la ecuación maestra dada en (76) queda transformada en

$$\partial_t P_t(\mathbf{s}, \mathbf{J}) = \sum_{\mathbf{y}} [\varpi_{\mathbf{J}}(\mathbf{s}^{\mathbf{y}}; \mathbf{y}) P_t(\mathbf{s}^{\mathbf{y}}, \mathbf{J}) - \varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{y}) P_t(\mathbf{s}, \mathbf{J})]. \quad (84)$$

Esta ecuación describe una situación en el sistema en la que las sinapsis permanecen *congeladas* (*quenched*) durante el intervalo de tiempo que caracteriza la variación temporal de \mathbf{s} . Así, el valor congelado de \mathbf{J} estaría determinado, por ejemplo, mediante un proceso previo de memorización de P patrones ξ^μ , tras el cual no ocurre ninguna modificación sináptica adicional (en la escala de tiempo de variación de

s). Varios casos particulares, asociados con diferentes simetrías y reglas de aprendizaje, han sido considerados antes por diferentes autores, por ejemplo:

(i) $J_{xy} = J_{yx}$. En general esta simetría no es realista si se trata de modelar sistemas biológicos [47], pero tiene la ventaja de que reduce el problema original a una forma más sencilla. Esto es así porque la condición de balance detallado implica entonces que el estado estacionario se corresponda con un estado de equilibrio termodinámico a temperatura T con energía

$$\mathcal{H}_{\mathbf{J}}(\mathbf{s}) = -\frac{1}{2} \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} J_{xy} s_{\mathbf{x}} s_{\mathbf{y}} + \sum_{\mathbf{x}} \theta_{\mathbf{x}} s_{\mathbf{x}}. \quad (85)$$

Este es el hamiltoniano de Ising si $J_{xy} = J$ para cualquier par de vecinos próximos y cero en cualquier otro caso. Si consideramos, por ejemplo, una distribución espacialmente aleatoria para las sinapsis J_{xy} , podemos obtener el modelo de Edwards-Anderson de vidrios de espines [5], o bien cualquiera de las familiares versiones de él. Alternativamente, tomando como regla de aprendizaje la regla de Hebb (74), da lugar al modelo de Hopfield [56]. Podemos considerar también una regla de aprendizaje no lineal [48, 49, 50] como, por ejemplo, $J_{xy} = N^{-1} P^{1/2} \phi(P^{1/2} \xi_{\mathbf{x}} \cdot \xi_{\mathbf{y}})$, donde $\xi_{\mathbf{x}} \equiv \{\xi_{\mathbf{x}}^{\mu} = \pm 1; \mu = 1, \dots, P\}$; este caso se reduce al modelo de Hopfield para $\phi(X) = X$.

(ii) $J_{xy} \neq J_{yx}$. Para esta situación, no es posible, en general, hacer una descripción hamiltoniana, pero podemos obtener alguna información interesante a partir de la ecuación maestra, por ejemplo, multiplicando los dos miembros de (84) por $s_{\mathbf{x}}$ y sumando sobre todas las posibles configuraciones \mathbf{s} de actividades neuronales, se obtiene que

$$\partial_t \langle s_{\mathbf{x}} \rangle = -2 \langle s_{\mathbf{x}} \Psi [2\beta s_{\mathbf{x}} h_{\mathbf{x}}(\mathbf{s}, \mathbf{J})] \rangle. \quad (86)$$

Después de introducir el campo local medio $m_{\mathbf{x}}(\mathbf{J}) \equiv \sum_{\mathbf{y}} J_{xy} \langle s_{\mathbf{y}} \rangle$ (por simplicidad, tomamos $\theta_{\mathbf{x}} = 0 \forall \mathbf{x}$ de aquí en adelante), y considerando la aproximación de campo medio $s_{\mathbf{x}} \approx \langle s_{\mathbf{x}} \rangle$, que implica $h_{\mathbf{x}}(\mathbf{s}, \mathbf{J}) \approx m_{\mathbf{x}}(\mathbf{J})$, se obtiene que la solución a (86) que caracteriza el estado estacionario del sistema verifica $\langle s_{\mathbf{x}} \rangle \approx \tanh \{\beta m_{\mathbf{x}}(\mathbf{J})\}$, donde el campo local medio se obtiene mediante la relación de autoconsistencia

$$m_{\mathbf{x}}(\mathbf{J}) = \sum_{\mathbf{y} \neq \mathbf{x}} J_{xy} \tanh \{\beta m_{\mathbf{x}}(\mathbf{J})\}, \quad J_{xy} \neq J_{yx}. \quad (87)$$

Ésta es la solución estacionaria obtenida antes por Crisanti y Sompolinsky para su modelo dinámico de red de Hopfield con acoplamientos asimétricos [60].

La discusión precedente indica que la ecuación (80), que caracteriza una clase muy interesante de modelos tipo Hopfield, puede permitirnos ilustrar, en cierta forma, la influencia, en el comportamiento emergente de los sistemas de redes de neuronas, de algunas características o propiedades de los mismos no consideradas hasta el momento y, como consecuencia, evaluar su papel en sistemas más realistas como son los sistemas neuronales biológicos. Como un primer paso, estamos particularmente interesados aquí en el límite $p \rightarrow 0$. Esto se corresponde con una situación en la que, una vez que el proceso de aprendizaje ha sido completado (o durante el mismo pero en una escala de tiempo muy inferior a la de la variación de las sinapsis consecuencia del aprendizaje), las sinapsis cambian muy rápidamente (fluctuaciones rápidas) comparado con los cambios en la actividad de las neuronas, de tal forma que se pueden distinguir dos escalas de tiempo bien definidas [61]. Por una parte hay una escala temporal microscópica, τ , para las fluctuaciones rápidas de las sinapsis, en la que la actividad de las neuronas no evoluciona de forma apreciable, y por otra parte hay una escala temporal, $t = p\tau$ para $p \rightarrow 0$ y $\tau \rightarrow \infty$, diferente de la anterior, en la que las neuronas evolucionan bajo una distribución estacionaria para las sinapsis. Bajo tal condición, para la que hay ciertas motivaciones biológicas, la ecuación (80) se transforma en

$$\partial_\tau f_\tau(\mathbf{J}) = \sum_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} \sum_{J'_{xy}} [\varpi(J'_{xy} \rightarrow J_{xy}) f_\tau(\mathbf{J}^{xy}) - \varpi(J_{xy} \rightarrow J'_{xy}) f_\tau(\mathbf{J})] \quad (88)$$

para las sinapsis, y

$$\partial_t P_t(\mathbf{s}) = \sum_{\mathbf{x}} [\bar{\varpi}(\mathbf{s}^{\mathbf{x}}; \mathbf{x}) P_t(\mathbf{s}^{\mathbf{x}}) - \bar{\varpi}(\mathbf{s}; \mathbf{x}) P_t(\mathbf{s})] \quad (89)$$

para las neuronas, donde

$$\bar{\varpi}(\mathbf{s}; \mathbf{x}) = \int d\mathbf{J} f(\mathbf{J}) \varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x}) \quad (90)$$

con $f(\mathbf{J})$ la solución estacionaria a (88), es decir $f(\mathbf{J}) = \lim_{\tau \rightarrow \infty} f_\tau(\mathbf{J})$. En otras palabras, el sistema se reduce en el límite $p \rightarrow 0$ al proceso marcoviano (89) con velocidades efectivas dadas por (90).

Este sistema que puede estudiar con detalle para diferentes elecciones de $f(\mathbf{J})$ y de $\varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x})$. Solamente vamos a considerar aquí velocidades elementales $\varpi_{\mathbf{J}}(\mathbf{s}; \mathbf{x})$ elegidas de forma que se verifiquen (81) y (82) (aunque pueden ser consideradas otro tipo de situaciones sencillas). La elección para $f(\mathbf{J})$ debe hacerse de forma que tenga en cuenta la naturaleza de simetría o asimetría de J_{xy} . A la hora de considerar variaciones del modelo, correspondientes a situaciones más familiares, suponemos que las fluctuaciones descritas por la distribución $f(\mathbf{J})$ son alrededor de valores medios correspondientes a una determinada regla de aprendizaje, esto es, por ejemplo,

$$\bar{J}_{xy} \equiv \int d\mathbf{J} f(\mathbf{J}) J_{xy} = \frac{1}{N} \sum_{\mu=1}^P \xi_{\mathbf{x}}^{\mu} \xi_{\mathbf{y}}^{\mu}, \quad (91)$$

para el caso de la regla de Hebb.

- Una elección particular que es consistente con esta consideración consiste en tomar

$$f(\mathbf{J}) = \prod_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} \delta \left(J_{xy} - \frac{1}{N} \sum_{\mu=1}^P \xi_{\mathbf{x}}^{\mu} \xi_{\mathbf{y}}^{\mu} \right); \quad (92)$$

sin embargo esta elección impide en la práctica la posibilidad de cualquier tipo de fluctuación, y el modelo se reduce al modelo de Hopfield con $J_{xy} = J_{yx}$ dados por la regla de Hebb.

- Más interesante es la elección

$$f(\mathbf{J}) = \prod_{\substack{\mathbf{x}, \mathbf{y} \\ \mathbf{x} \neq \mathbf{y}}} \sum_{\mu=1}^P a_{\mu} \delta (J_{xy} - \eta_{xy}^{\mu}), \quad (93)$$

que implica (91) para

$$\eta_{xy}^{\mu} = \lambda_{\mu} \xi_{\mathbf{x}}^{\mu} \xi_{\mathbf{y}}^{\mu}, \quad \lambda_{\mu} = (a_{\mu} N)^{-1}. \quad (94)$$

La elección (93) se corresponde con la situación en la que los patrones memorizados contribuyen con una cantidad η_{xy}^{μ} con probabilidad a_{μ} a la intensidad sináptica J_{xy} , independientemente de las otras, de forma que no existe correlación espacial entre las diferentes sinapsis. Una variedad de resultados para esta situación serán mostrados en las próximas secciones de este capítulo.

- Otra situación de interés que implica (91) consiste en tomar

$$f(\mathbf{J}) = \sum_{\mu=1}^P a_{\mu} \prod_{\substack{x,y \\ x \neq y}} \delta(J_{xy} - \eta_{xy}^{\mu}), \quad (95)$$

es decir, cada patrón memorizado ξ^{μ} contribuye con probabilidad a_{μ} a una configuración completa de sinapsis \mathbf{J} , permitiendo, a diferencia de la anterior elección, es decir (93), que existan correlaciones espaciales entre los diferentes patrones almacenados.