

Modelo de Hopfield

Héctor Casas Ruiz

20 de mayo de 2022

Física Computacional del Grado en Física. Universidad de Granada.



Índice

1. Introducción	1
2. Resultados	3
2.1. Funcionamiento de la red con un patrón	3
2.1.1. $T = 10^{-4}$	3
2.1.2. Estudio para varias temperaturas	3
2.2. Funcionamiento de la red con varios patrones	7
2.3. Capacidad de la red	14
3. Conclusiones	14

1. Introducción

En este informe se va a utilizar el algoritmo de Metrópolis para estudiar el modelo de Hopfield de red neuronal. Este modelo consiste en la que cada neurona intercambia información con el resto de la red mediante el Hamiltoniano. Este modelo fue popularizado por Hopfield en 1982 pero previamente desarrollado por Little en 1974. Este modelo es una red neuronal recurrente, que se utiliza para simular sistemas biológicos, en la física es uno de los modelos para describir el vidrio de espín y también se utiliza para resolver diferentes problemas informáticos como el problema de las ocho reinas y el problema del viajante de comercio.[1][2]

Fundamento teórico

Consideramos una red cuadrada bidimensional de $N \times N$ partículas. Donde cada nodo representa una neurona activa ($s_{i,j} = 1$) o inactiva ($s_{i,j} = 0$). El Hamiltoniano para cada configuración del sistema es

$$H(s) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \omega_{ij,kl} s_{ij} s_{kl} + \sum_{i=1}^N \sum_{j=1}^N \theta_{i,j} s_{i,j} \quad (1)$$

Donde la interacción entre cada neurona viene dada por el peso sináptico $\omega_{ij,kl}$:

$$\omega_{ij,kl} = \begin{cases} \frac{1}{N^2} \sum_{\mu=1}^P (\xi_{i,j}^\mu - a^\mu)(\xi_{k,l}^\mu - a^\mu) & \text{si } (i,j) \neq (k,l) \\ 0 & \text{si } (i,j) = (k,l) \end{cases} \quad (2)$$

siendo $\xi^\mu = \{\xi_{i,j}^\mu\}$, es decir, los patrones previamente almacenados, con $\mu = 1, 2, \dots, P$, siendo P el número de patrones. Estos patrones corresponden al mínimo de energía para $T = 0$.

$$a^\mu = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \xi_{i,j}^\mu \quad (3)$$

Por otro lado $\theta_{i,j}$ es el umbral de disparo, que viene dado por

$$\theta_{i,j} = \frac{1}{2} \sum_{k=1}^N \sum_{l=1}^N \omega_{ij,kl} \quad (4)$$

A partir de estas ecuaciones se implementa la dinámica del sistema utilizando el algoritmo de Metrópolis, que consiste en calcular en cada iteración el mínimo entre $e^{-\Delta H/T}$ y 1 y comparar este valor con un número aleatorio $\xi \in [0, 1]$ y cambiar el sistema si ξ es menor que el mínimo.

Por otro lado, para cuantificar cómo la red se aproxima al patrón a medida que el tiempo aumenta se ha utilizado el solapamiento de la red con cada patrón ξ^μ , que viene dado por:

$$m^\mu(s) = \frac{1}{N^2 a^\mu (1 - a^\mu)} \sum_{\mu=1}^P (\xi_{i,j}^\mu - a^\mu)(s_{i,j} - a^\mu) \quad (5)$$

Además, para ahorrar tiempo de cálculo, se ha calculado ΔH analíticamente. En cada iteración la red neuronal cambia en una neurona (n, m) , al calcular el Hamiltoniano teniendo en cuenta este término y el Hamiltoniano en el siguiente posible estado de energía y teniendo en cuenta que $s'_{n,m} = s_{n,m} - 1$ siendo $s'_{n,m}$ la neurona de la nueva configuración de la red, se obtiene:

$$\Delta H = H' - H = (1 - 2s_{n,m}) \left(\theta_{n,m} - \sum_{i=1}^N \sum_{j=1}^N \omega_{ij,nm} s_{i,j} \right) \quad (6)$$

Sobre la implementación

La simulación del modelo se ha realizado en *Python* concretamente en un *Jupyter Notebook*, que se encuentra adjunto en la carpeta de *GitHub* (Enlace). Los cálculos para los primeros apartados se han realizado con *voluntario1.ipynb*, mientras que los cálculos del último apartado se han realizado con *voluntario1_4.ipynb* que es una versión modificada del anterior que permite variar el número de patrones dentro de la función principal.

2. Resultados

2.1. Funcionamiento de la red con un patrón

2.1.1. $T = 10^{-4}$

En primer lugar se ha estudiado el funcionamiento de la red para un único patrón, de $N = 30$. Partiendo de una condición inicial aleatoria se puede observar la evolución en la Figura 1, en función del paso Monte Carlo. También se ha estudiado su solapamiento, que se encuentra en la Figura 2. Como se puede observar en la gráfica, la red se iguala completamente al patrón en el paso 7, cuando su solapamiento es -1 . También se puede observar como la red tiende al patrón inverso, es decir al estado espúreo, aunque se observará más adelante que también tiende al propio patrón, dependiendo del estado inicial.

En segundo lugar se ha estudiado la red con una configuración inicial del patrón difuminado un 70 %. Encontrándose la evolución en la Figura 3 y el solapamiento en la Figura 4.

También se ha estudiado la evolución de la red para un estado inicial difuminado un 30 %. La evolución se puede observar en la Figura 5 y el solapamiento se encuentra en la Figura 6. Comparando las diferentes figuras sobre solapamiento, se puede observar que las configuraciones con distribuciones iniciales más parecidas al patrón, convergen más rápido a este, especialmente en los primeros pasos Monte Carlo.

2.1.2. Estudio para varias temperaturas

En este caso se va a estudiar el solapamiento para diferentes temperaturas. En primer lugar se ha realizado la Figura 7, que representa el solapamiento en función del número de pasos Monte Carlo para el patrón anterior, de 900 neuronas y con un estado inicial aleatorio fijo. En esta se observa que para $T > 0,001$ el sistema no converge al patrón, si no a un estado intermedio, hasta una temperatura aproximada de $T = 0,04$ en la que la red no se solapa con el patrón y cuyo solapamiento se debe a fluctuaciones aleatorias. Cabe destacar que el solapamiento se ha tomado en valor absoluto, para visualizarlo, ya que, a veces el sistema converge al estado espúreo, dando lugar un solapamiento negativo.

A continuación se encuentra la Figura 8 en la que se puede observar el solapamiento esta vez en función del número de pasos Monte Carlo, tomándose 50 temperaturas equiespaciadas logarítmicamente en $10^{-4} < T < 1$, cal-

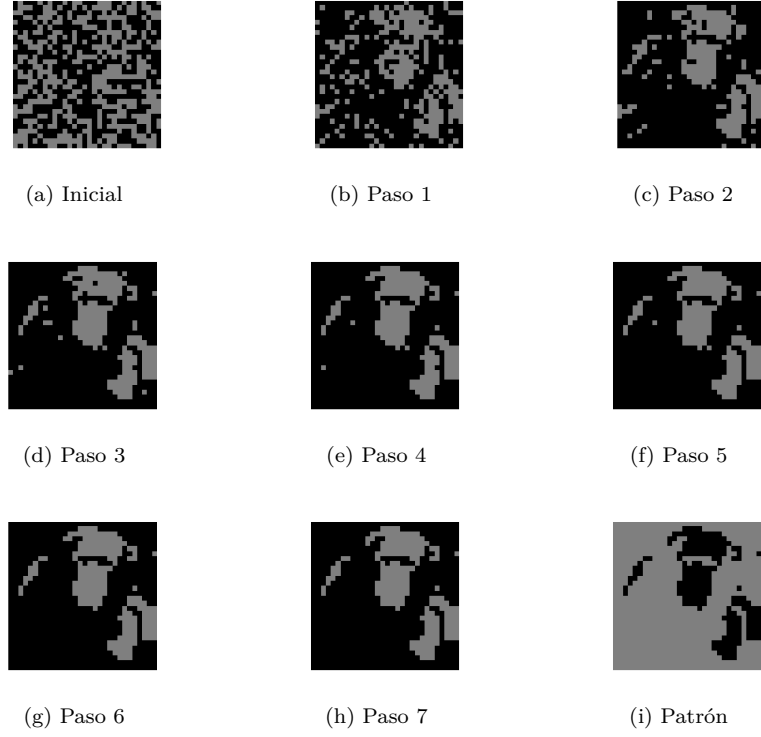


Figura 1: Evolución para un patrón de $N = 30$ con estado inicial aleatorio

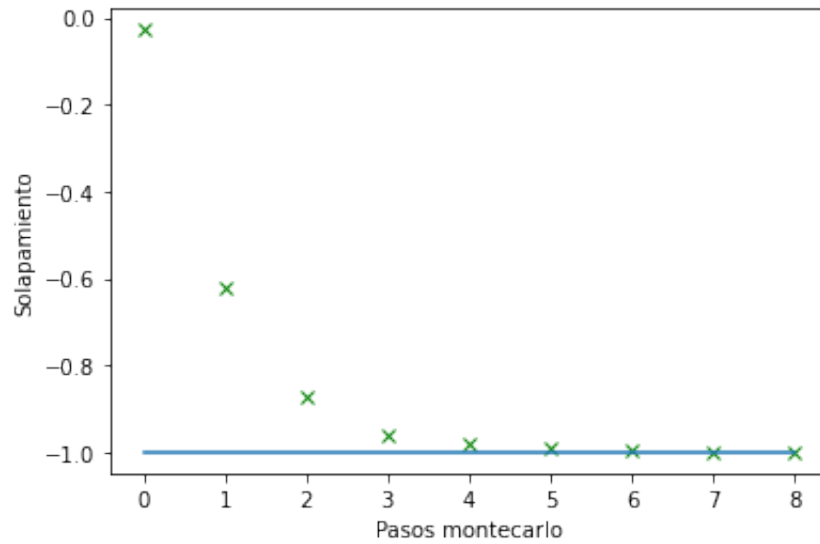


Figura 2: Solapamiento en función del número de pasos Monte Carlo

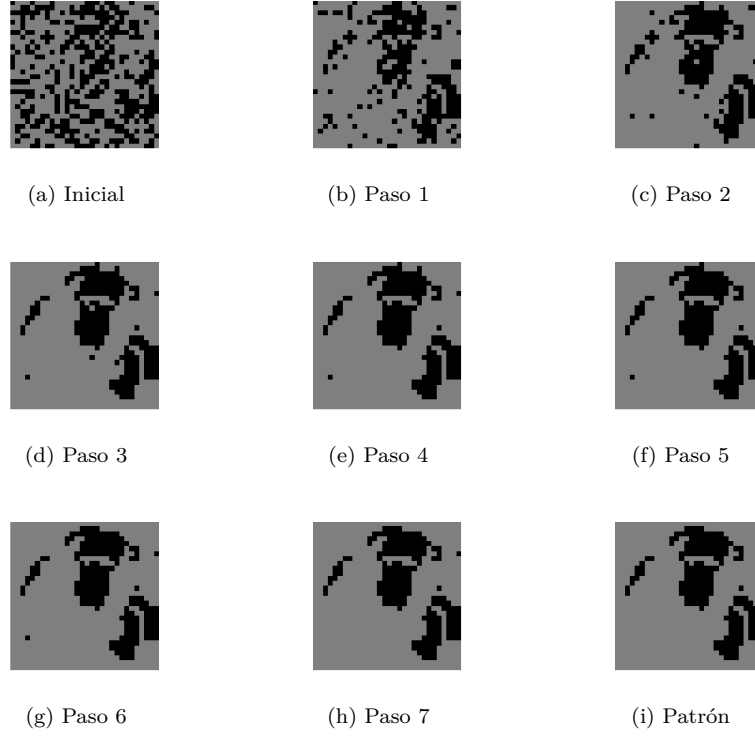


Figura 3: Evolución para un patrón de $N = 30$ con estado inicial del patrón difuminado 70 %

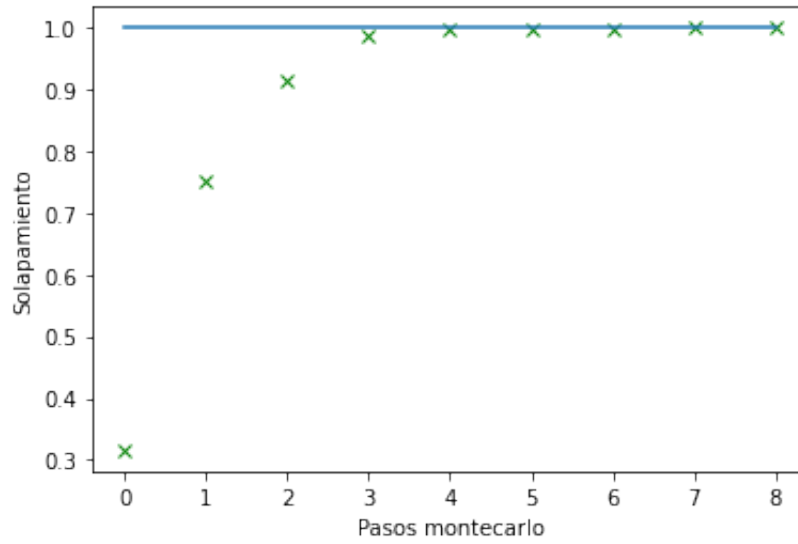


Figura 4: Solapamiento en función del número de pasos Monte Carlo con patrón difuminado 70 %

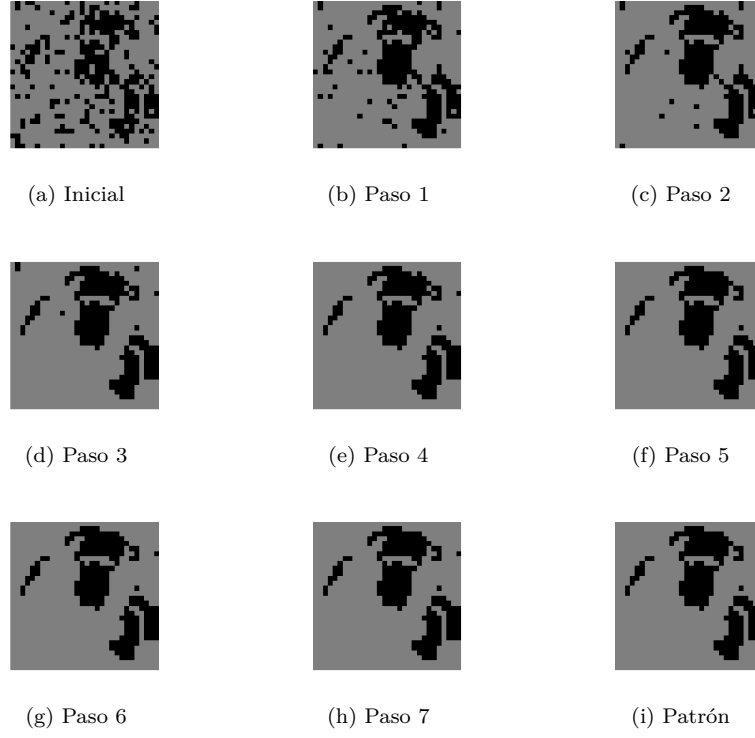


Figura 5: Evolución para un patrón de $N = 30$ con estado inicial del patrón difuminado 30 %

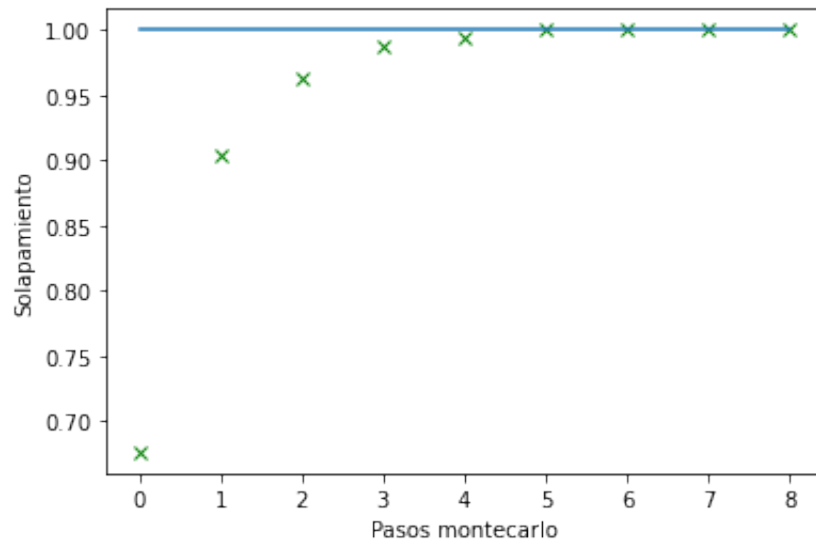


Figura 6: Solapamiento en función del número de pasos Monte Carlo con patrón difuminado 30 %

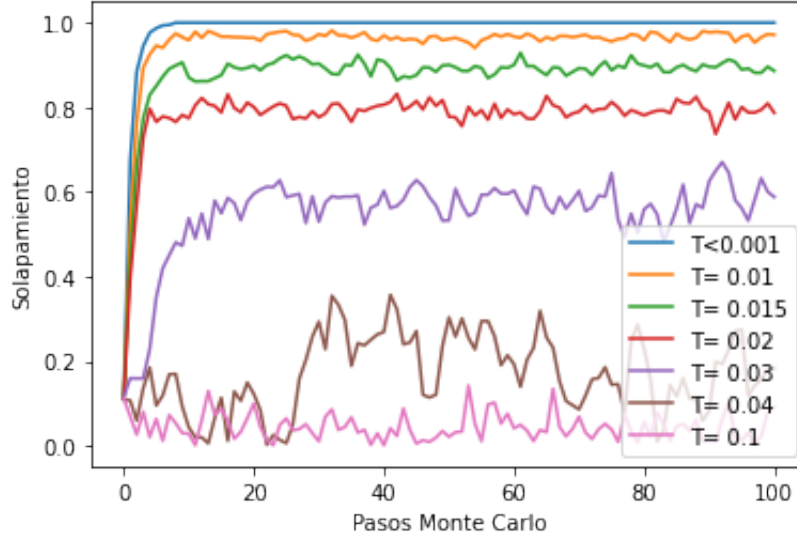


Figura 7: Solapamiento en función del número de pasos Monte Carlo para diferentes temperaturas

culándose el solapamiento 10 veces para cada una y promediando. Entre los 50 puntos que se han tomado se observa que la zona en la que converge se ubica en $10^{-4} \leq T \leq 3,55 \cdot 10^{-3}$. La zona de transición de fase se encuentra en $3,55 \cdot 10^{-3} < T \leq 4,09 \cdot 10^{-2}$ y la zona en la que no converge es para $T > 4,09 \cdot 10^{-2}$.

La zona de transición de fase se debe al número limitado de neuronas, la cual se estrecha conforme $N \rightarrow \infty$, aproximándose a una función escalón. A partir de los datos en detalle de la anterior gráfica, se puede estimar la temperatura crítica alrededor de $T \approx 0,03$. En las Figuras 9 y 10 se observa la misma gráfica para $N = 60$ y $N = 90$, comprobándose este comportamiento.

También se ha estudiado el comportamiento de la red en función de la temperatura para la red con patrones iniciales difuminados, con el 70 % y del 30 % de las neuronas cambiadas respecto al patrón, las gráficas del solapamiento frente al número de pasos Monte Carlo para diferentes temperaturas se encuentran en las Figuras 11 y 12. Como se puede observar, el solapamiento inicial es de alrededor de 0.3 en el primer caso y de 0.7 en el segundo, pero los sistemas tienden a estados parecidos a los del estado inicial aleatorio.

2.2. Funcionamiento de la red con varios patrones

En primer lugar se va a estudiar el funcionamiento de una red con dos patrones. Cabe destacar que como comenta Hopfield en [3], para que la red converja a uno de los patrones, los estados deben ser ortogonales, es decir que $N^2/2$ de las partículas sean diferentes o la red minimizará su energía con un estado mixto de los patrones. En la Figura 13 se puede observar la evolución de

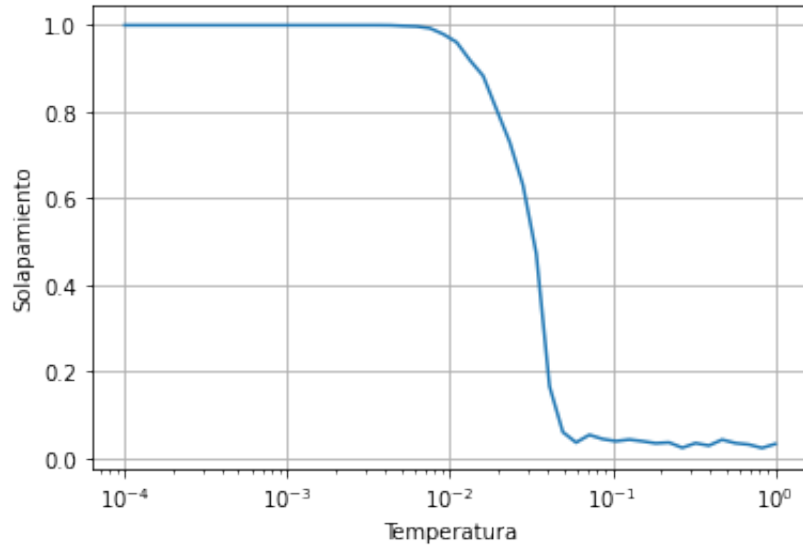


Figura 8: Solapamiento en función de la temperatura para N=30 y 100 pasos Monte Carlo

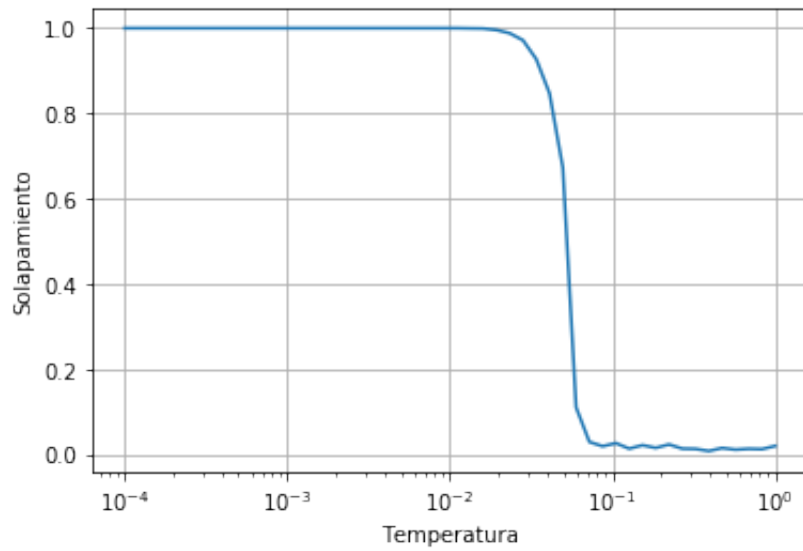


Figura 9: Solapamiento en función de la temperatura para N=60 y 20 pasos montecarlo

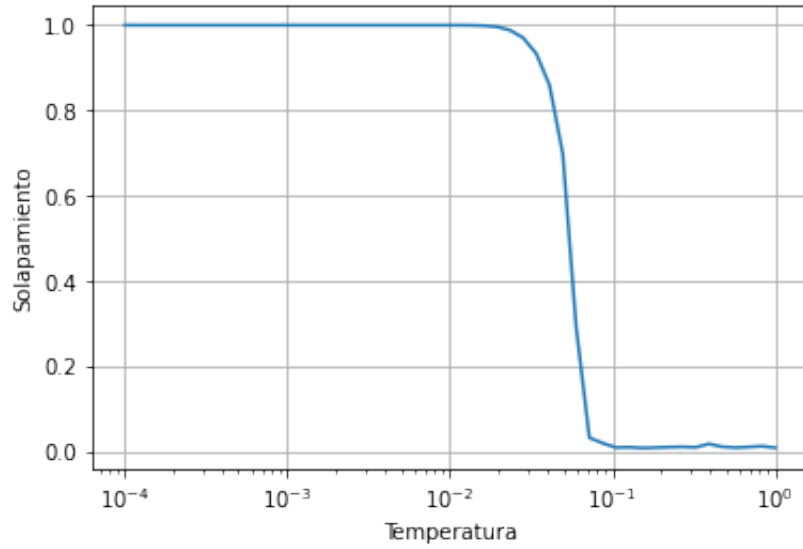


Figura 10: Solapamiento en función de la temperatura para $N=90$ y 20 pasos montecarlo

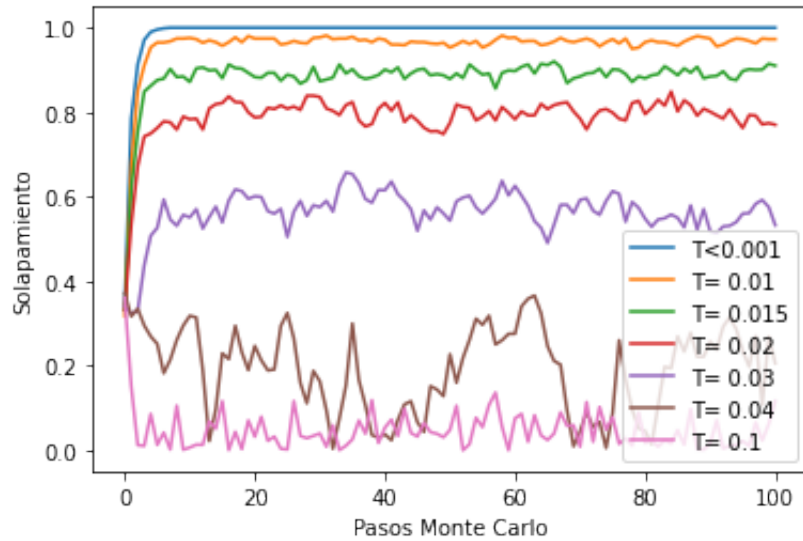


Figura 11: Solapamiento en función del número de pasos Monte Carlo para diferentes temperaturas con difuminado del 70 %

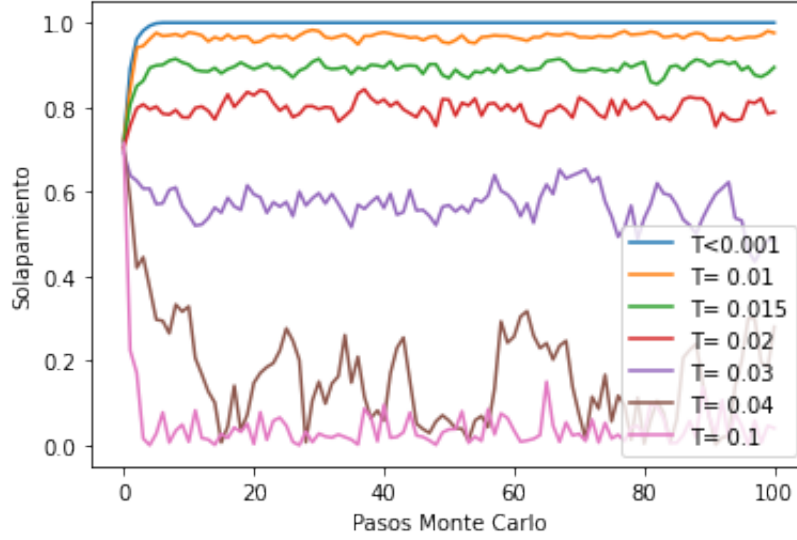


Figura 12: Solapamiento en función del número de pasos Monte Carlo para diferentes temperaturas con difuminado del 30 %

estos patrones y en la Figura 14 se puede observar el solapamiento.

Se han modificado los patrones para que fuesen ortogonales, la evolución y solapamiento de la red con estos patrones se encuentran en las Figuras 15 y 16, respectivamente. En este caso si se produce el solapamiento esperado.

Como se puede observar la red converge a uno de los patrones, aunque los patrones son algo parecidos por lo que la red también se solapa levemente de forma espúrea con el otro. A continuación se muestra el solapamiento de los patrones para varias temperaturas (Figura 17), es evidente que en general la red converge a uno de los dos patrones para temperaturas menores a 10^{-2} , salvo en algunos casos, que probablemente se deban a la aleatoriedad del método. Para $T > 10^{-2}$ se observa que la red converge a patrones mixtos, hasta $T = 6 \cdot 10^{-2}$, a partir de la cual la red no se solapa con ninguno de los patrones.

También se ha estudiado el comportamiento de la red para cuatro patrones, tomándose aleatorios para asegurar su ortogonalidad. En primer lugar se tiene la gráfica de solapamiento en función de temperatura, que se muestra en la Figura 18. En este caso se tiene un resultado similar al anterior.

También se ha estudiado el caso en el que el estado inicial es el primer patrón con el 90 % de las neuronas cambiadas (Figura 19), dando lugar a que la red converja al primer patrón para $T < 2 \cdot 10^{-2}$.

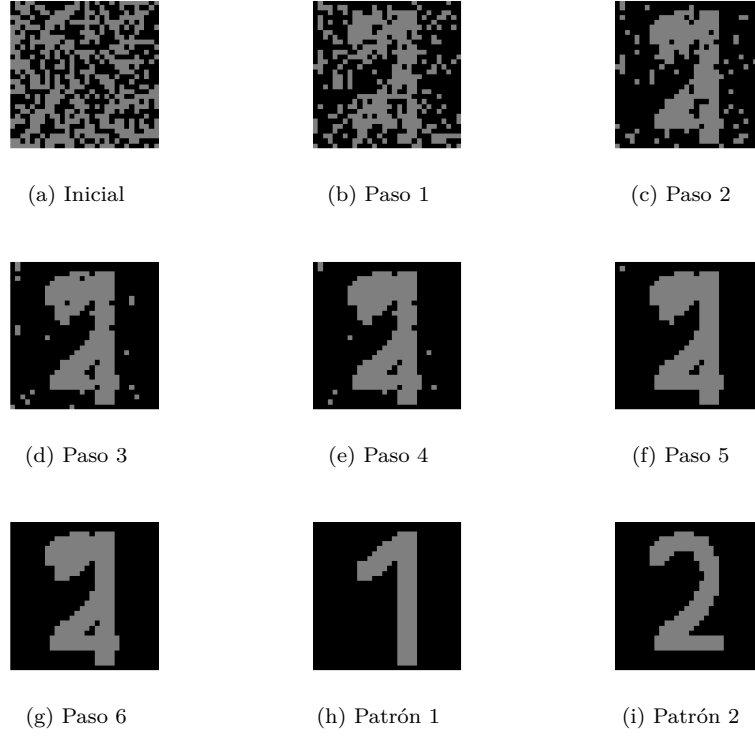


Figura 13: Evolución para dos patrones de $N = 30$ no ortogonales

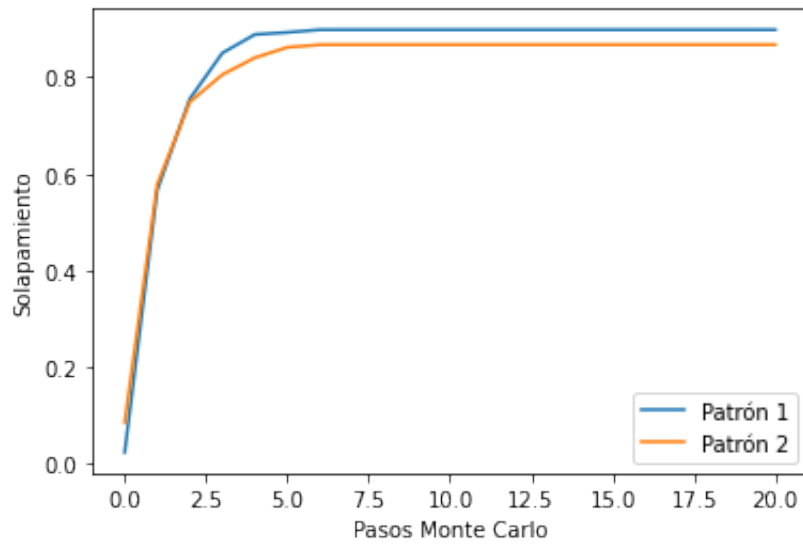


Figura 14: Solapamiento para dos patrones no ortogonales

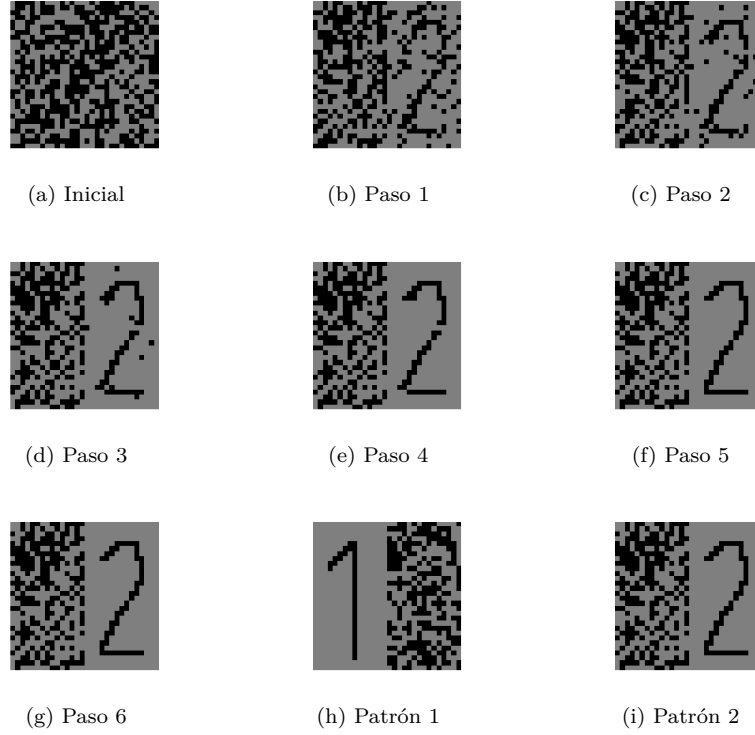


Figura 15: Evolución para dos patrones de $N = 30$ ortogonales

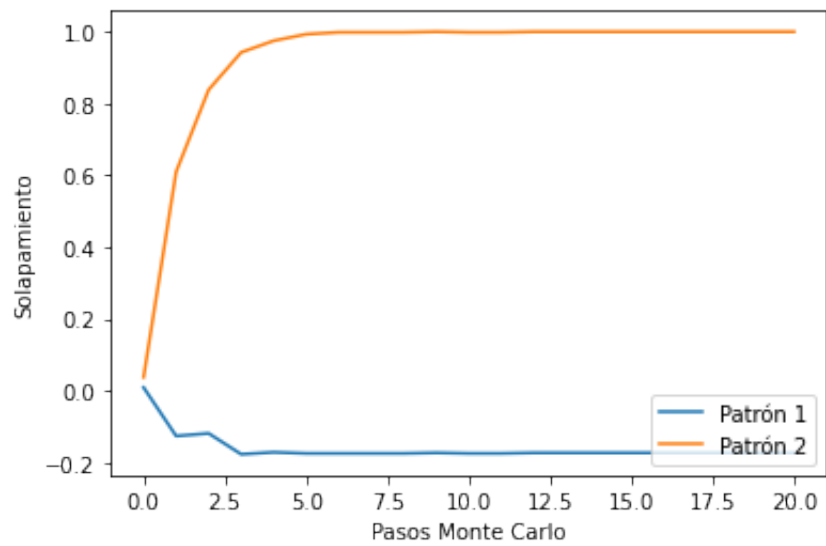


Figura 16: Solapamiento para dos patrones ortogonales

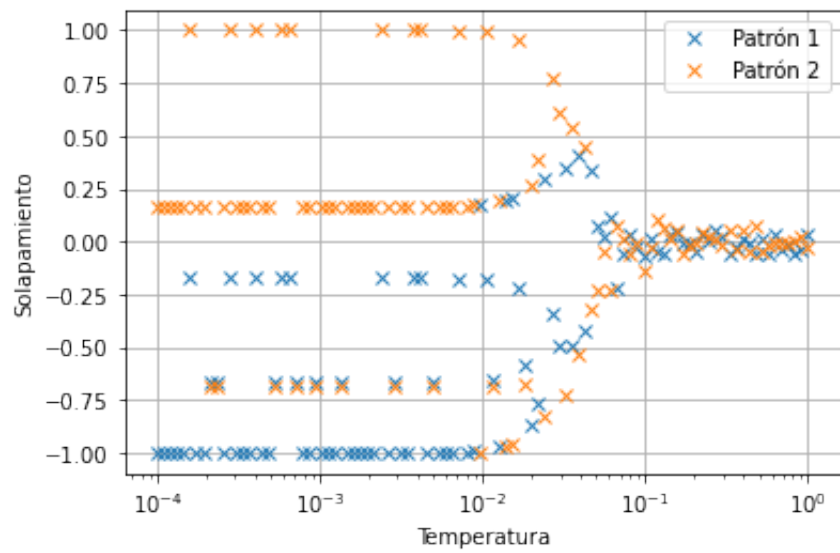


Figura 17: Solapamiento en función de la temperatura.

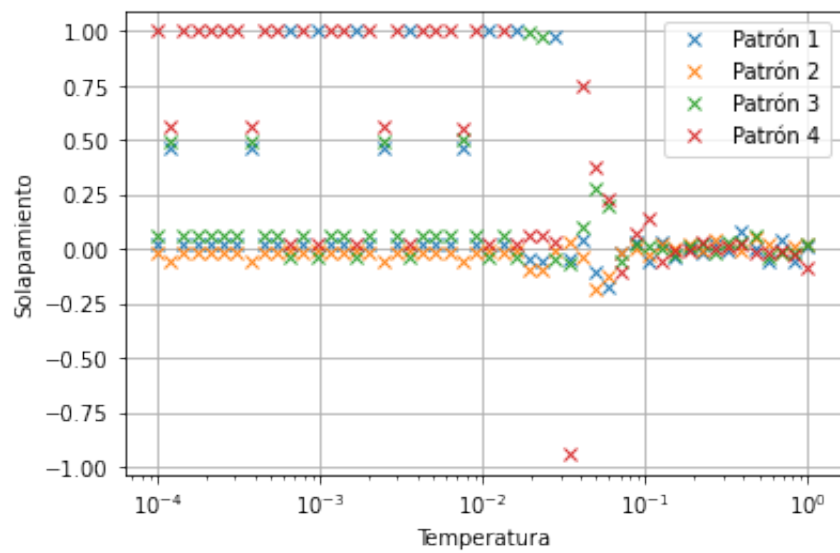


Figura 18: Solapamiento de 4 patrones para diferentes temperaturas

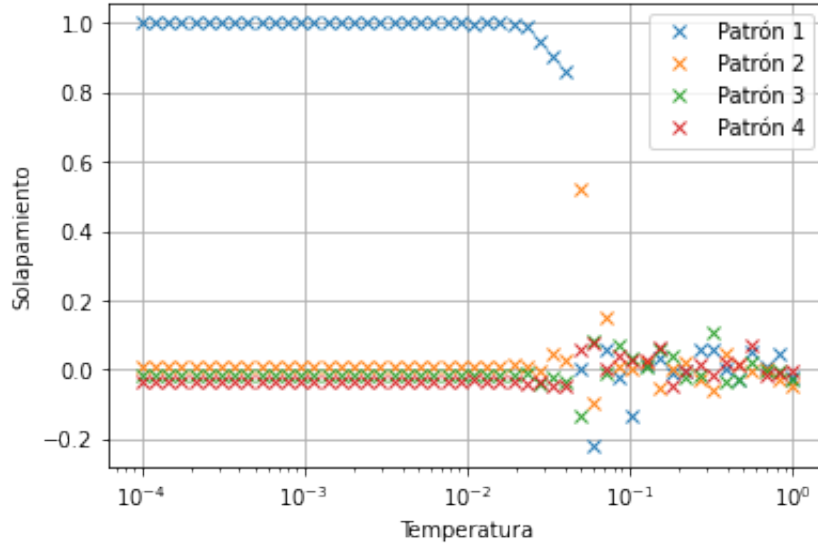


Figura 19: Solapamiento de 4 patrones para diferentes temperaturas (con estado inicial parecido al primer patrón)

2.3. Capacidad de la red

En este apartado se ha estudiado la capacidad de la red de 'recordar' patrones, considerando que la red recuerda un patrón cuando el solapamiento $|m^\mu(s)| > 0,75$. Para ello se ha utilizado una red de $N^2 = 400$ neuronas con estados iniciales aleatorios y patrones aleatorios, realizándose 20 pasos Monte Carlo a $T = 10^{-4}$ para un número de hasta 50 patrones. En la figura 20 se encuentra el número de patrones que ha recordado la red. Aquí se observa que la red recuerda un patrón hasta alrededor de los 15 patrones, cuando deja de recordar.

También se tiene la fracción máxima $\alpha_c = P_c/N^2$ en la que se divide el número de patrones recordados por el número de neuronas, esta gráfica se encuentra en la Figura 21.

3. Conclusiones

En primer lugar, sobre el comportamiento de la red con un único patrón se destaca el estudio del solapamiento frente a la temperatura. En este se puede observar la brusca caída del solapamiento al alcanzar temperaturas del orden de $T \approx 10^{-2}$, dando paso de un solapamiento completo a la nula convergencia del sistema al patrón previamente almacenado. Igualmente se ha estudiado como era el solapamiento frente a la temperatura para redes neuronales de mayor tamaño, observándose que conforme este aumentaba la caída era mas brusca. Desde un punto de vista físico, esto equivale al cambio de fase termodinámico de un sistema, que, en el límite termodinámico sería discontinuo.

En segundo lugar, se comentará el funcionamiento de la red con varios patro-

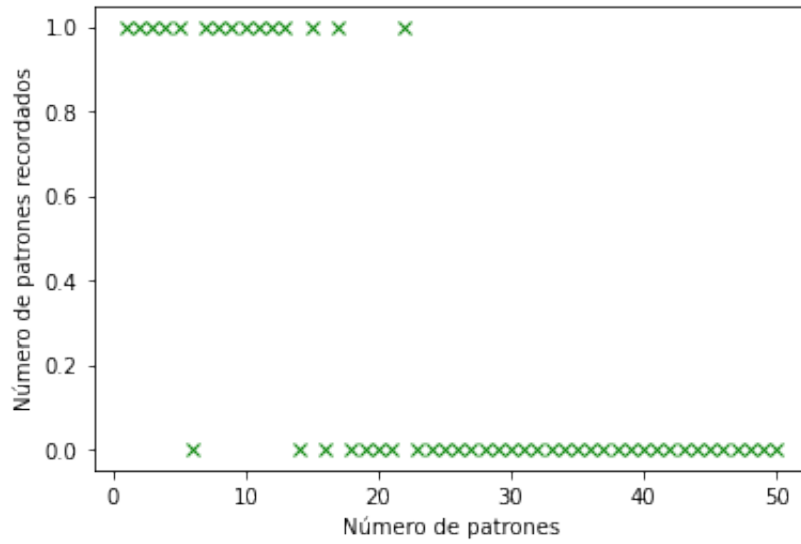


Figura 20: Patrones recordados en función del número de patrones

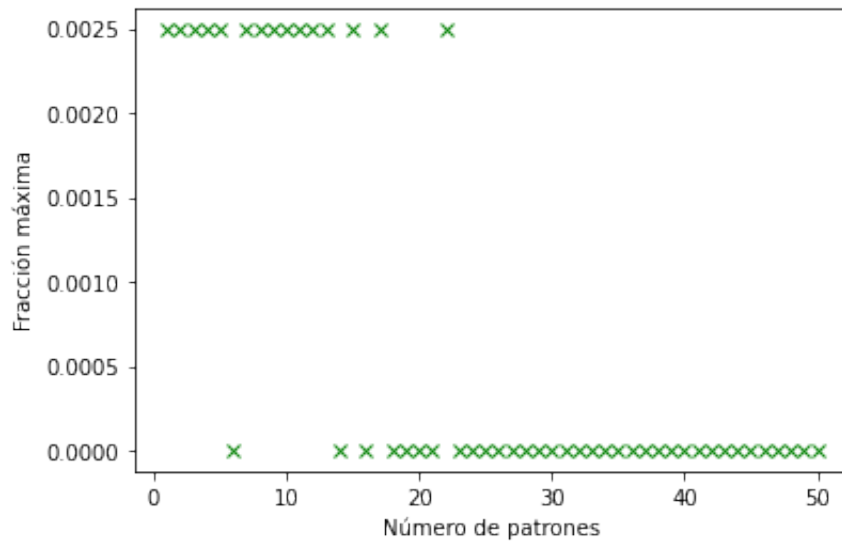


Figura 21: Fracción máxima en función el número de patrones

nes almacenados. Se ha analizado la evolución de los patrones dependiendo de su ortogonalidad, observándose la importancia de que difieran en $N^2/2$, ya que, si no la red encuentra un estado de energía mínima mixto de los dos patrones. Por otro lado, al variar la temperatura se han obtenido resultados similares a los anteriores, extrapolándolos a que había más patrones, es decir, se ha observado que para ciertas temperaturas el solapamiento era pleno, seguido de una zona de transición debida al tamaño del sistema y una zona en la que la red no se solapa con ninguno de los patrones. Sin embargo, tanto con 2 y 4 patrones para ciertas temperaturas dentro de la zona de solapamiento, la red ha minimizado su energía en un estado mixto, esto se puede deber al carácter aleatorio del método, pero en términos generales el comportamiento ha sido el esperado.

Por último, se ha comprobado la capacidad de la red para 'recordar' patrones. Se ha observado que el número de patrones recordados ha disminuido con el número de patrones almacenados. Además, el número máximo de patrones recordados ha sido 1, lo que probablemente se deba a que si el sistema era capaz de converger a uno de los patrones, el solapamiento se hacía 1 mientras que el solapamiento con el resto de patrones sería prácticamente nulo. Mientras que, en el caso de que el sistema tendiese a un estado mixto, el solapamiento con ninguno de los patrones alcanzaba la condición de que $|m| > 0,75$. En este apartado cabe destacar la limitación que suponía el tiempo de cálculo al aumentar el número de patrones en memoria, haciendo que no se pudiesen estudiar sistemas con más neuronas que 400.

Referencias

- [1] Hopfield Network Wikipedia
- [2] R. Rojas: Neural Networks, Springer-Verlag, Berlin, 1996
- [3] Hopfield J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences of the United States of America, 79(8), 2554-2558. DOI
- [4] Amit, Daniel Gutfreund, Hanoach & Sompolinsky, Haim. (1987). Information storage in neural networks with low levels of activity. 35. 10.1103/PhysRevA.35.2293.