



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
MAESTRÍA EN CIENCIAS (FÍSICA)

DESCOMPOSICIÓN DE PERTURBACIONES EN REDES BOOLEANAS
CON UMBRALES

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS (FÍSICA)

PRESENTA:
MAYRA PATRICIA GARCÍA ALCALÁ

TUTOR PRINCIPAL:
DR. MAXIMINO ALDANA GONZÁLEZ
INSTITUTO DE CIENCIAS FÍSICAS

COMITÉ TUTORAL:
DR. HERNÁN LARRALDE RIDAURA
INSTITUTO DE CIENCIAS FÍSICAS
DR. CARLOS VILLARREAL LUJÁN
INSTITUTO DE FÍSICA

CUERNAVACA, MORELOS, ENERO 2015

Índice general

1. Resumen	5
2. Terapia de Múltiples Fármacos	7
3. Maximización de Entropía	9
3.1. Formalismo matemático	9
3.2. Modelo de máxima entropía por parejas	10
3.3. Predicción de respuesta de múltiples antibióticos	12
4. Redes Booleanas con Umbrales	17
4.1. Redes Booleanas con Umbrales	17
4.2. Régimen dinámico	19
4.2.1. Análisis de transición de fase	21
4.2.2. Análisis usando teoría de la información	24
4.3. Conclusiones	27
5. Descomposición del Efecto de Perturbaciones en RBU	31
5.1. Perturbaciones en Pesos	31
5.2. Efecto de múltiples perturbaciones	34
5.3. Predicción del efecto de perturbaciones	36
5.3.1. Predicción con ecuación de Isserlis	36
5.3.2. Predicción utilizando promedios	38
5.4. Hipótesis de aditividad en las perturbaciones	43
5.4.1. Hipótesis de aditividad en ecuación de Isserlis	43
5.4.2. Hipótesis de aditividad en pares de perturbaciones	45
5.4.3. Conclusiones	46
6. Discusión	49
A. Análisis semi-analítico de transición de fase	51

1. Resumen

En el presente trabajo nos preguntamos si es posible predecir el efecto de una combinación de múltiples perturbaciones en la dinámica de un sistema, dado que se conoce el efecto al aplicar las perturbaciones por parejas e individualmente.

Wood et al. [1], abordaron esa cuestión, pero en particular para buscar si hay una forma de predecir el efecto total de cocteles de múltiples fármacos aplicados a células patógenas. Utilizando el principio de Maximización de Entropía obtuvieron una ecuación simple que es función únicamente del efecto de los fármacos aplicados por pares e individualmente, y que les permite hacer la predicción para el coctel en conjunto.

Aquí establecemos una analogía entre un conjunto de células y fármacos que las afectan, con un sistema de redes booleanas y perturbaciones que se apliquen sobre ellas. Las redes que utilizamos son redes booleanas con umbrales que tienen una regla dinámica que se supone las provee de una alta no-linealidad. Además, estas redes han sido utilizadas anteriormente para modelar el ciclo celular de la levadura, el cual es un fenómeno no lineal.

Analizamos si la descomposición por parejas que utilizan Wood et al. para el efecto de los fármacos es útil también para la predicción del efecto de múltiples perturbaciones en las redes.

Como parte de este análisis se estudiaron las propiedades dinámicas de las redes y encontramos que existe una transición de fase de segundo orden caracterizada por un punto crítico. También, para nuestra sorpresa, encontramos que dependiendo del régimen dinámico de las redes, la respuesta a las perturbaciones puede ser lineal o no lineal. Esta respuesta lineal es inesperada porque los nodos de la red interactúan a través de una regla altamente no lineal. Un resultado interesante es que la ecuación utilizada por Wood et al. para la predicción de la respuesta de la colonia de bacterias a cocteles de fármacos funciona también para la predicción en las redes booleanas. Sin embargo, encontramos en nuestro sistema que la validez de esa ecuación no está relacionada con la maximización de entropía, sino que se debe precisamente a la respuesta lineal de la red a las perturbaciones.

2. Terapia de Múltiples Fármacos

Las terapias de combinación de fármacos consisten en el consumo de más de un medicamento para el tratamiento de alguna enfermedad.

Estos tratamientos son de gran importancia para enfermedades que son difíciles de combatir, como la malaria, VIH, tuberculosis, y cáncer, entre otros. Estas enfermedades tienen una alta tendencia a desarrollar resistencia a fármacos ya que las bacterias o patógenos que las ocasionan tienen una alta tasa de mutación, lo cual les hace evolucionar rápidamente y adaptarse ante la presencia de fármacos. En la literatura se suelen encontrar dos razones generales para utilizar las terapias de múltiples fármacos para combatir estas enfermedades.

La primera razón es que la combinación de múltiples fármacos retrasa el surgimiento de resistencia. La hipótesis de por qué sucede esto es que la probabilidad de que el patógeno evolucione con las diferentes mutaciones que permiten la resistencia disminuye cuando los fármacos inhiben múltiples blancos o mecanismos de acción simultáneamente. Como consecuencia, la probabilidad de que el patógeno se resista a una combinación de antibióticos será menor porque dicha probabilidad surge de la multiplicación de las probabilidades de resistencia a los antibióticos individuales, haciendo la suposición de que estos sean eventos independientes [2]. Este efecto se ha observado para los casos de malaria [3], lepra [4] y tuberculosis [5].

La segunda razón es que la combinación de los fármacos puede tener un efecto sinérgico. En farmacología se dice que dos componentes son sinérgicos si al aplicarlos en combinación se incrementa la eficacia de cada uno. Esto permite utilizar concentraciones más bajas de los fármacos para obtener el efecto deseado, en comparación con la concentración que se utilizaría en el tratamiento con uno solo de los fármacos [6]. Sin embargo generalmente no se puede saber *a priori* si dos fármacos serán sinérgicos, aún conociendo la respuesta individual de ellos a distintas concentraciones.

En la figura 2.1 se muestra otra representación de las distintas interacciones entre pares de medicamentos. Estas gráficas se llaman Isobogramas y corresponden a un corte transversal de la gráfica de concentraciones y el efecto de la interacción, cuantificado, por ejemplo, con el número de células o luminiscencia a cierta concentración (a la derecha de la figura). La línea roja muestra el caso de aditividad. Según el modelo de Loewe [7] dos fármacos son aditivos, si el efecto de aplicarlos conjuntamente es simplemente la suma de los efectos que se tendrían al aplicarlos por separado. En este caso los fármacos no interactúan. Por otro lado, la interacción sinérgica corresponde al caso en que el efecto conjunto es mayor que la suma de los efectos por separado, y la interacción antagónica corresponde al caso contrario: el efecto conjunto es menor que la suma de los efectos separados. En estos dos últimos casos, los fármacos interactúan ya sea para potenciar su efecto (caso sinérgico) o para suprimirse mutuamente (caso antagónico).

Investigaciones recientes apuntan a que la combinación de medicamentos sinérgicos promueven el surgimiento de la resistencia de bacterias a los antibióticos, y por lo tanto no son la mejor elección al momento de diseñar cocteles de fármacos [9]. Quienes han observado este fenómeno sugieren utilizar combinaciones antagónicas [10], mientras que otros proponen reemplazar las combinaciones de medicamentos por la aplicación individual y secuencial de los fármacos [11]. Sin embargo aún no es claro cuál es el efecto de estas combinaciones en la evolución de resistencia.

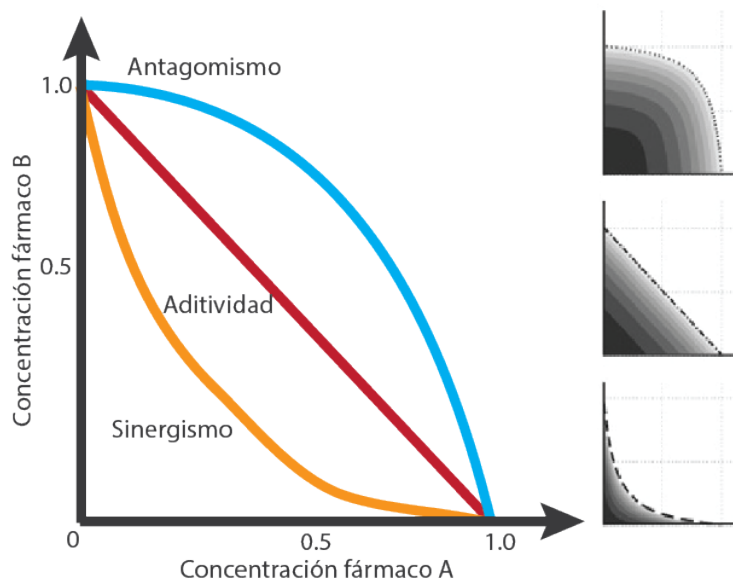


Figura 2.1: Representación de los distintos efectos que pueden tener dos antibióticos, sinergismo y antagonismo, aditividad. Las líneas representan curvas de nivel de las tasas de crecimiento $g(C_A, C_B)$ expresadas como función de las concentraciones de los fármacos. Del lado derecho se muestran mapas de contorno de las superficies $g(C_A, C_B)$ para los tres casos, aditivo, sinérgico y antagonista. Imagen adaptada de [8]

Estos tratamientos son difíciles de diseñar ya que el sistema en cuestión, las células y su interacción con fármacos, es un sistema complejo, no en el sentido de “complicado”, sino en el sentido de los Sistemas Complejos. Una característica de un sistema complejo es que está compuesto por diferentes partes que interaccionan entre sí y de cuya interacción emerge algún comportamiento colectivo, el cual no puede explicarse al conocer únicamente las características individuales de cada una de las partes del sistema. Así pues, es difícil saber el efecto que se obtendrá en las células, bacterias o tumores, cuando se administran estos cocteles, justamente por esta interacción, (además de que una célula en sí ya es un sistema complejo) ya que además no se conocen todos los mecanismos moleculares a través de los cuales interactúan los medicamentos. Estas interacciones son estudiadas principalmente de manera descriptiva a partir de experimentación en laboratorio pero al analizar estas combinaciones *in vivo* puede cambiar radicalmente el efecto, ser más costoso y llevar mayor tiempo hacer las mediciones [12].

Para atacar este problema Wood et al. proponen un método para predecir la respuesta de múltiples fármacos que es independiente del mecanismo molecular de éstos. Ellos hacen un análisis suponiendo que la respuesta de más de dos antibióticos obedece a reglas estadísticas en vez de químicas. Recurriendo al principio de máxima entropía encuentran que el efecto de múltiples medicamentos puede ser obtenido a partir del conocimiento de la respuesta de las células a las interacciones individuales y por pares de estos medicamentos, independientemente de que las interacciones sean sinérgicas, aditivas o antagónicas. En lo sucesivo, nos referiremos a este resultado como *descomposición por parejas*.

En el siguiente capítulo presentaré la forma en que el principio de máxima entropía permite hacer la descomposición por parejas tanto en los cocteles de antibióticos como en la dinámica de sistemas neuronales y en el movimiento colectivo observado en parvadas de pájaros [13, 14].

3. Maximización de Entropía

Para sistemas biológicos resulta difícil conocer las leyes de interacción entre los elementos que lo componen, por lo que se acude a modelos probabilísticos para describir su comportamiento. Por ejemplo, para poder descifrar la relación estímulo/respuesta en el cerebro se construyen distribuciones de probabilidad de los distintos patrones de actividad neuronal [15]. Otro ejemplo se da en biología molecular, donde se requiere de la probabilidad de plegamiento de una secuencia de aminoácidos para conocer la forma natural de plegamiento de una proteína [16, 17]. Encontrar las distribuciones de probabilidad de los eventos involucrados en estos sistemas puede ser muy difícil por las siguientes razones [18]:

1. Dado que los sistemas biológicos típicamente cuentan con un número grande de elementos, número de variables aleatorias necesarias para construir las funciones de probabilidad correspondientes es también muy grande.
2. Los elementos de estos sistemas interaccionan de forma no lineal entre ellos por lo que no se puede considerar la hipótesis de independencia estadística para construir la función de probabilidad conjunta que describe el estado del sistema completo.

Una técnica que ha sido utilizada con éxito para estimar las probabilidades de los estados de algunos sistemas biológicos consiste en utilizar el *Principio de Máxima Entropía*. Este principio establece que sujeta a ciertas constricciones, la distribución de probabilidad que mejor representa el estado del sistema es la distribución que tiene máxima entropía. La justificación de utilizar este principio proviene de Teoría de la Información, donde la entropía S es una medida de la “incertidumbre” representada en una distribución de probabilidad. Una distribución que tiene un valor alto de entropía corresponde a un sistema cuyo estado es difícil de conocer con certeza. Por lo tanto, esta relación entre entropía e incertidumbre implica que al construir una distribución de probabilidad y pedir que la función de entropía sea máxima de manera consistente con las constricciones conocidas, la distribución obtenida será lo más imparcial posible [19].

La función de entropía, en términos de distribuciones de probabilidad, según el enfoque de la teoría de la información, fue introducida por Shannon en 1948 [20] y está definida como:

$$S(X) = - \sum_{i=1}^n p(x_i) \log p(x_i), \quad (3.1)$$

donde $S(X)$ es la entropía de una variable aleatoria X que puede tomar valores x_1, \dots, x_n , con probabilidad $p(x_1), \dots, p(x_n)$, respectivamente.

3.1. Formalismo matemático

Considérese una variable aleatoria que puede tener valores discretos $x = \{x_1, x_2, \dots, x_n\}$, y que la probabilidad de ocurrencia para cada uno de sus valores $\{p_1, p_2, \dots, p_n\}$ es desconocida, pero el valor esperado de la función $f(x)$ sí es conocido

$$\langle f(x) \rangle = \sum_{i=1}^n p_i f(x_i) = F. \quad (3.2)$$

Además, existe siempre la restricción de que las probabilidades p_i deben estar normalizadas:

$$\sum_{i=1}^n p_i = 1. \quad (3.3)$$

Nos interesa pues, encontrar los valores de las probabilidades p_i que maximizan la función entropía (3.1) y que están sujetos a las restricciones dadas por las relaciones (3.2) y (3.3). Para utilizar el principio de maximización de entropía recurrimos a multiplicadores de Lagrange. Por lo tanto maximizaremos la función:

$$\begin{aligned} \mathcal{L} &= S(p_1, p_2, \dots, p_n) - \lambda \left[\sum_{i=1}^n p_i - 1 \right] - \mu \left[\sum_{i=1}^n p_i f(x_i) - F \right] \\ &= \sum_{i=1}^n \left[p_i \ln \left(\frac{1}{p_i} \right) - \lambda p_i - \mu p_i f(x_i) \right] + \lambda + \mu F \end{aligned} \quad (3.4)$$

Haciendo $\frac{\partial \mathcal{L}}{\partial p_i} = 0$ se llega a la expresión de la probabilidad:

$$p_i = \exp(-\lambda - 1 - \mu f(x_i)), \quad (3.5)$$

donde las variables λ y μ son obtenidas al sustituir esta solución de p_i en las constricciones (3.2) y (3.3).

Similarmente, para el caso de cualquier número de funciones $f_1(x), f_2(x), \dots, f_n(x)$, dados los valores esperados $\langle f_1(x_i) \rangle, \langle f_2(x_i) \rangle, \dots, \langle f_n(x_i) \rangle$, la distribución de máxima entropía está dada por

$$p_i = e^{-H_i}, \quad (3.6)$$

con $H_i = [\lambda_0 + \lambda_1 f_1(x_i) + \dots + \lambda_n f_n(x_i)]$.

3.2. Modelo de máxima entropía por parejas

Considérese ahora un sistema de N variables aleatorias x_1, x_2, \dots, x_N continuas cuya distribución de probabilidad $P(x_1, x_2, \dots, x_N)$ es desconocida.

El teorema de momentos de Hausdorff dice que es posible conocer exactamente la distribución de probabilidad $P(x_1, \dots, x_N)$ del sistema si se conocen todos los momentos $\langle x_i \rangle, \langle x_i x_j \rangle, \langle x_i x_j x_k \rangle, \langle x_i x_j x_k x_l \rangle$, etcétera, (siendo las variables x_i acotadas en $[0,1]$, en el caso de que no estén acotadas, el problema es llamado de momentos de Hamburger). [21]. Ahora nos podemos preguntar cuál es la cantidad mínima de los momentos necesaria para tener una buena aproximación a la distribución de probabilidad. Trabajos recientes han utilizado el principio de máxima entropía, incluyendo solamente los primeros y segundos momentos, para construir una función aproximada $P_s(x_1, \dots, x_N)$ [14, 22, 13, 23, 24]. Sorprendentemente, al utilizar esta distribución aproximada para hacer predicciones del estado del sistema, o para calcular con ella momentos de orden mayor, o el promedio de alguna función $f(x_1, \dots, x_N)$, se obtienen resultados muy buenos respecto a lo que se obtendría con la distribución exacta $P(x_1, \dots, x_N)$.

Más específicamente, para crear la distribución aproximada $P_s(x_1, \dots, x_N)$ se maximiza la entropía utilizan el principio de máxima entropía

$$S(\dots) = \int \dots \int P(x_1, \dots, x_N) \ln \left(\frac{1}{P(x_1, \dots, x_N)} \right) dx_1 \dots dx_N. \quad (3.7)$$

poniendo como constricciones únicamente los primeros y segundos momentos (y la normalización):

$$\langle x_i \rangle = \int_{\{x_1, x_2, \dots, x_N\}} x_i P(x_1, x_2, \dots, x_N) dx_1 \dots dx_N = f_i, \quad (3.8)$$

$$\langle x_i x_j \rangle = \int_{\{x_1, x_2, \dots, x_N\}} x_i x_j P(x_1, x_2, \dots, x_N) dx_1 \dots dx_N = f_{ij}, \quad (3.9)$$

$$\int_{\{x_1, x_2, \dots, x_N\}} P(x_1, \dots, x_N) dx_1 \dots dx_N = 1. \quad (3.10)$$

El resultado de este proceso de maximización es la distribución gaussiana multivariada [13]:

$$P_s(x_1, \dots, x_N) = \frac{1}{Z} \exp \left[\sum_{i=1}^n \lambda_i x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n J_{ij} x_i x_j \right], \quad (3.11)$$

donde Z es la constante de normalización y λ_i y J_{ij} son constantes que están relacionadas con los valores de las constricciones (3.8) y (3.9) de la forma:

$$Z = \int_{-\infty}^{\infty} \exp \left[\sum_{i=1}^n \lambda_i x_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n J_{ij} x_i x_j \right] dx_1 \dots dx_N$$

$$-\frac{\partial \ln Z}{\partial \lambda_i} = f_i$$

$$-\frac{\partial \ln Z}{\partial J_{ij}} = f_{ij}$$

Trabajos recientes [14, 13, 22, 23, 24, 1] utilizan este método para aproximar las distribuciones de probabilidad de sus sistemas, encontrando que esta aproximación se acerca en gran medida a la distribución real, mostrando con ello que es suficiente incluir a lo más las interacciones entre pares de sus componentes para describir el comportamiento del sistema, y por lo tanto las interacciones de mayor orden pueden despreciarse, no incluyen mucha más información de la que se agrega con los primeros y segundos momentos como restricción.

Por ejemplo en el trabajo de Bialek et al., [14], se utiliza este formalismo para analizar el movimiento colectivo emergente de parvadas de pájaros construyendo la distribución de probabilidad de velocidades conjunta $P(\{\vec{s}_i\})$ donde \vec{v}_i es el vector de velocidad que se le atribuye al pájaro i a cada paso de tiempo y $\vec{s}_i = \vec{v}_i / \|\vec{v}_i\|$. Utilizando el principio de máxima entropía con la restricción de que la correlación $C_{ij} = \langle \vec{s}_i \cdot \vec{s}_j \rangle$ sea igual a la obtenida experimentalmente, obtienen la distribución de máxima entropía $P_S(\{\vec{s}_i\})$, la cual les permite predecir correctamente la propagación del parámetro de orden direccional a través de la parvada.

En otro ejemplo, esta vez en neurociencia, Schneidman et al. [13] utilizan el modelo de máxima entropía para crear la distribución de probabilidad $P(\sigma_i)$ de los patrones de expresión de un conjunto de neuronas retinales de vertebrados. σ_i es una variable aleatoria asociada a una neurona i , que toma valores 1 y 0, dependiendo de si la neurona generó un potencial de acción o no, respectivamente, en una pequeña ventana de tiempo. Ellos incluyen la restricción de que el valor esperado y las correlaciones por pares entre estas variables aleatorias sean los mismos que se observan en el experimento. Otra vez, con la distribución aproximada de probabilidad ellos reproducen la tasa de expresión de cada patrón y encuentran que es una muy buena predicción al compararlo con las tasas reales.

3.3. Predicción de respuesta de múltiples antibióticos

Wood et al. se preguntan en [1] si es posible predecir la respuesta del crecimiento bacteriano ante la aplicación de más de dos medicamentos sin considerar el mecanismo con el que estos actúan, acudiendo únicamente a un marco estadístico.

Para lograr esto, miden la tasa de crecimiento g_i de una colonia de bacterias después de que se le ha aplicado el antibiótico D_i a cierta concentración C , siendo g_i la pendiente de la curva de crecimiento bacteriano en el tiempo. Ahora suponen que g_i es una tasa de crecimiento promedio de todas las bacterias que conforman la colonia. Se supone entonces que la tasa de crecimiento x_i de cada célula individual, cuando el antibiótico D_i se aplica, es una variable aleatoria cuyo promedio es $\langle x_i \rangle = g_i$. La variable aleatoria x_i está normalizada por la tasa de crecimiento G de las bacterias sin tratar. Se asocia entonces el efecto de cada medicamento D_i a una variable aleatoria x_i , de modo que los valores que puede tener esta variable provienen de una distribución de probabilidad $P(x_1, \dots, x_n)$. En el caso de aplicar dos medicamentos D_i con una concentración C y D_j con concentración C' conjuntamente, la tasa de crecimiento de la colonia estaría dada por $g_{ij} = \langle x_i x_j \rangle$, y cuando se apliquen tres medicamentos D_i, D_j, D_k , simultáneamente, con concentraciones fijas, la tasa de crecimiento promedio correspondiente sería $g_{ijk} = \langle x_i x_j x_k \rangle$, etcétera. Todas las variables asociadas a las tasas de crecimiento están normalizadas por la tasa de crecimiento G . Dado este planteamiento, si se tiene la distribución de probabilidad conjunta $P(x_1, \dots, x_n)$ se puede encontrar el efecto de cualquier combinación de los medicamentos D_1, D_2, \dots, D_n , para distintas concentraciones, a través de los momentos de la distribución.

Para encontrar esta distribución utilizan el modelo de Máxima Entropía, poniendo como restricciones las tasas de crecimiento $g_i(C_a) = \langle x_i(C_a) \rangle$ y $g_{ij}(C_a, C_b) = \langle x_i(C_a) x_j(C_b) \rangle$ (con $i, j = 1, 2, 3, \dots, n$, índice que denota los antibióticos distintos y siendo C_a y C_b la concentración con que son aplicados los fármacos) observadas experimentalmente cuando se aplican los antibióticos individualmente y por parejas, respectivamente, además de la normalización de la probabilidad 3.10.

Esto lleva a la distribución $P_s(x_1, \dots, x_n)$ de probabilidad como la mostrada en la ecuación (3.11), donde λ_i y J_{ij} en este caso son parámetros relacionados con los efectos g_i y g_{ij} . En la figura 3.1 se muestra un esquema de este proceso para encontrar la respuesta a los múltiples antibióticos.

Utilizan ahora la distribución $P_s(x_1, \dots, x_n)$ para predecir las tasas de crecimiento g_{ijk} y g_{ijkl} que se obtendrían al aplicar tres y cuatro antibióticos simultáneamente. Denotamos como \hat{g}_{ijk} y \hat{g}_{ijkl} a las predicciones hechas utilizando la función $P_s(x_1, \dots, x_n)$. Al comparar las tasas de crecimiento predichas \hat{g}_{ijk} y \hat{g}_{ijkl} con las tasas reales g_{ijk} y g_{ijkl} observadas en el experimento se encuentra un muy buen acuerdo, como se muestra en la figura 3.2, en donde cada punto de la gráfica izquierda corresponde a una combinación de tres antibióticos a una concentración dada. Los mapas de color de la parte derecha muestran la tasa de crecimiento g_{ijk} reales y las predichas \hat{g}_{ijk} para las com-

binaciones de tres medicamentos aplicados con distintas concentraciones. Estos mapas permiten observar que la predicción es buena no importando si el resultado de la combinación es sinérgico o antagónico.

Como se observa en la ecuación 3.11, la distribución que se obtiene por máxima entropía restringiendo a los primeros y segundos momentos es una gaussiana multivariada. Por lo tanto, podemos utilizar el teorema de Isserlis (también conocido como teorema de Wick) que dice que es posible obtener los momentos de orden mayor a partir solamente de sus segundos y primeros momentos. Por ejemplo para el tercer y cuarto momentos se cumple las relaciones:

$$\begin{aligned} \langle x_i x_j x_k \rangle &= \langle x_i \rangle \langle x_j x_k \rangle + \langle x_j \rangle \langle x_i x_k \rangle + \langle x_k \rangle \langle x_i x_j \rangle \\ &\quad - 2 \langle x_i \rangle \langle x_j \rangle \langle x_k \rangle \end{aligned} \quad (3.12)$$

$$\begin{aligned} \langle x_i x_j x_k x_l \rangle &= \langle x_i x_j \rangle \langle x_k x_l \rangle + \langle x_i x_k \rangle \langle x_j x_l \rangle + \langle x_i x_l \rangle \langle x_j x_k \rangle \\ &\quad - 2 \langle x_i \rangle \langle x_j \rangle \langle x_k \rangle \langle x_l \rangle \end{aligned} \quad (3.13)$$

Utilizando estas ecuaciones directamente para hacer las predicciones se obtienen resultados idénticos al obtenido con todo el marco estadístico antes mencionado, simplemente haciendo la sustitución para las tasas de crecimiento de tres y cuatro antibióticos:

$$\hat{g}_{ijk} = g_i g_{jk} + g_j g_{ik} + g_k g_{ij} - 2g_i g_j g_k, \quad (3.14)$$

$$\hat{g}_{ijkl} = g_{ij} g_{kl} + g_{ik} g_{jl} + g_{il} g_{jk} - 2g_i g_j g_k g_l. \quad (3.15)$$

De hecho, utilizando estas ecuaciones se obtienen los mismos resultados que los reportados en la Fig. 3.2.

Para generalizar la metodología anterior, consideremos un sistema que es modificado por diferentes perturbaciones. Si estas perturbaciones no interactúan, entonces el efecto total en el sistema es igual a la suma del efecto de las perturbaciones aplicadas de manera independiente. Cuando las perturbaciones interactúan entre ellas, es difícil de conocer el efecto neto que generan éstas en la red. En nuestro caso la información de estas interacciones está expresada en el Hamiltoniano que aparece en la ecuación (3.6).

Supongamos entonces que el sistema está conformado por N cuerpos (los fármacos o perturbaciones, por ejemplo) y todos ellos interactúan entre sí. Un Hamiltoniano que describe al sistema sería de la forma:

$$H = - \sum_i \lambda_i x_i - \sum_{ij} J_{ij} x_i x_j - \sum_{ijk} J_{ijk} x_i x_j x_k - \dots - \sum_{i\dots N} J_{i\dots N} x_i \dots x_N \quad (3.16)$$

Si conocemos todos los coeficientes de acoplamiento, λ_i , J_{ij} , etc., decimos que tenemos el sistema bien caracterizado. En el caso de estar en un sistema probabilístico, conocer estos coeficientes es equivalente a conocer todos los momentos de la distribución de probabilidad asociada al sistema, como ya se mencionó.

Sin embargo la jerarquía completa de momentos no siempre se tiene. Para simplificar el problema, entonces, se busca cortar esa serie dejando únicamente la mínima cantidad de términos necesaria para caracterizar la distribución. Para el caso de los múltiples antibióticos se sabe que no es suficiente el conocimiento de los primeros momentos debido a que los antibióticos interactúan, por lo que la siguiente posibilidad es considerar las interacciones a segundo orden. Esto es un equivalente de lo que hacen Wood et al. para la predicción del efecto de la respuesta ante múltiples fármacos. El éxito de sus resultados y la sencillez de las ecuaciones que utilizan abre la pregunta

de bajo qué condiciones estas funciones, construidas solamente a partir de los primeros y segundos momentos, pueden ser utilizadas para hacer predicciones más generales y si estas predicciones son escalables a otros sistemas. Para analizar esto, en el presente trabajo tomaremos un sistema que es conocido por su respuesta no lineal ante perturbaciones, haciendo con eso una analogía con las bacterias y los antibióticos. Nuestro objetivo será contestar de forma muy general la pregunta:

¿es posible predecir el efecto de múltiples perturbaciones conociendo sólo el efecto de perturbaciones aplicadas por parejas e individualmente?

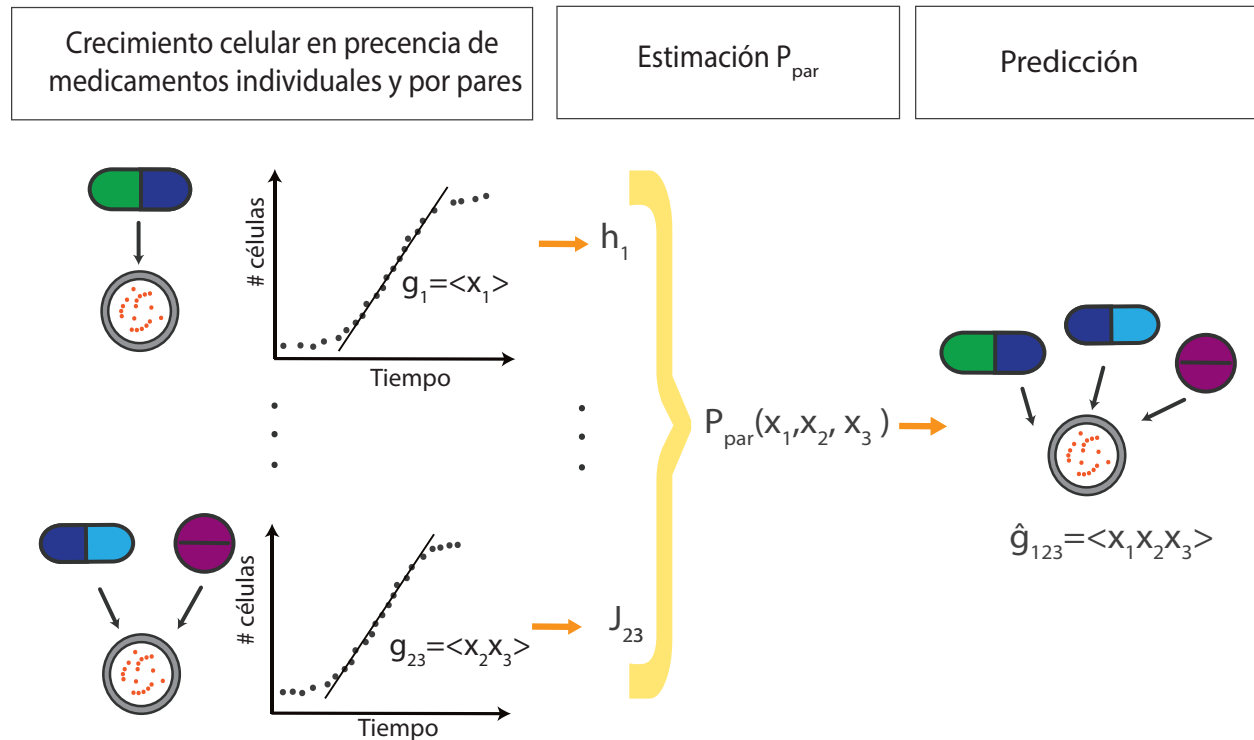


Figura 3.1: Esquema del procedimiento realizado en [1] para predecir el efecto de múltiples antibióticos en el crecimiento bacteriano. Se muestra el ejemplo de aplicación de tres medicamentos. Primeramente se obtiene la curva de crecimiento habiendo aplicado los medicamentos individualmente y por pares. Se asigna a cada antibiótico D_i una variable aleatoria x_i , tal que el promedio de esta variable es igual al crecimiento g_i . Se construye una distribución de máxima entropía P_s con esos datos. Los coeficientes h_i y J_{ij} de P_s son escogidos de tal manera que los promedios $\langle x_i \rangle$ y $\langle x_i x_j \rangle$ de la distribución de pares coincida con las tasas reales de crecimiento g_i y g_{ij} a cada concentración. Luego, la distribución P_s se utiliza para hacer la predicción de haber aplicado D_1 , D_2 y D_3 juntos calculando el momento $\langle x_1 x_2 x_3 \rangle$ de la distribución

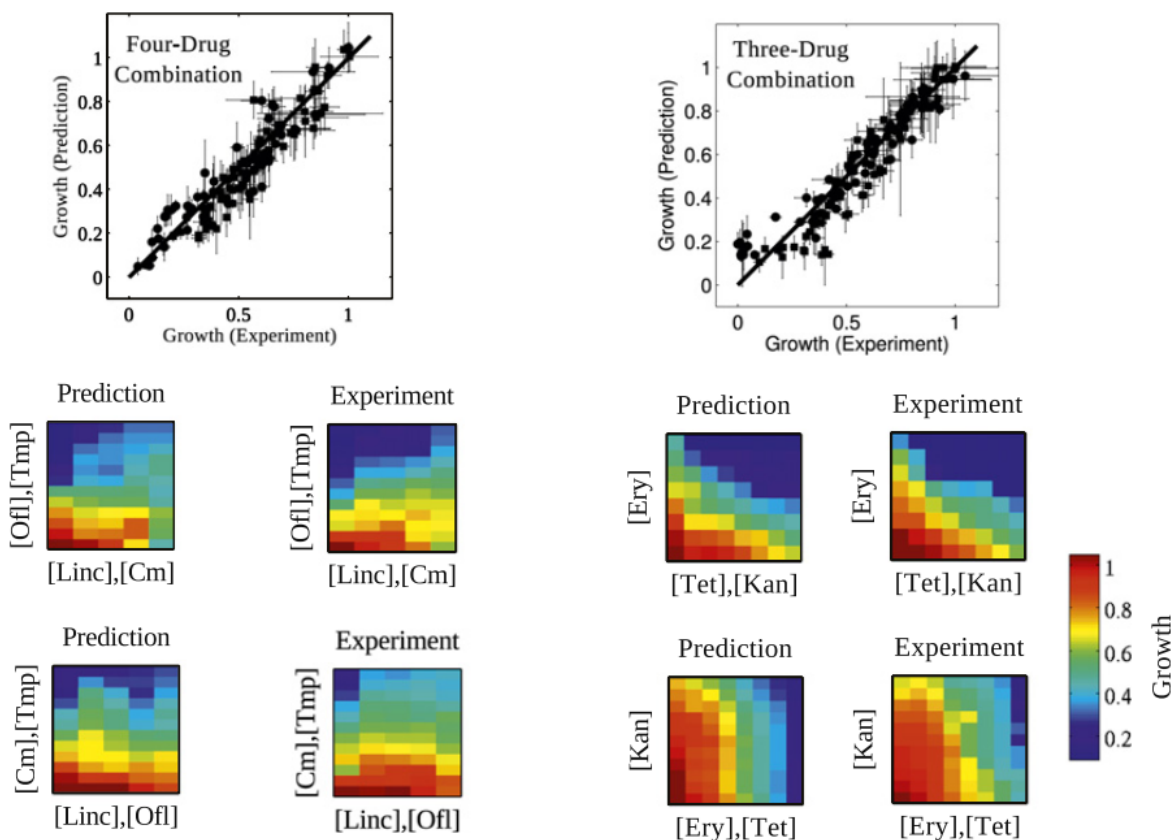


Figura 3.2: Comparación de el efecto real g_{ijk} de tres antibióticos en el crecimiento bacteriano con la predicción \hat{g}_{ijk} de dicho efecto obtenida con la distribución de máxima entropía por pares P_s . *Izquierda*: Cada punto corresponde a la combinación de tres antibióticos a una dosis específica. La línea recta representa la identidad. *Derecha*: mapas de calor para los antibióticos. En ellos se muestra la tasa de crecimiento real y predicha después de aplicar los tres antibióticos a distintas concentraciones. Los ejes que hacen referencia a dos fármacos corresponde a la variación de la concentración de igual forma para ellos. En este ejemplo se observan el comportamiento sinérgico y antagónico de las combinaciones. Las etiquetas corresponden a abreviaciones de los nombres de los antibióticos: [Ofl], ofloxacin; [Linc], lincomycin; [Tmp], trimethoprim; [Cm], chloramphenicol; [Tet], tetracycline; [Ery], erythromycin; [Kan], kanamycin. Imagen tomada de [1].

4. Redes Booleanas con Umbrales

4.1. Redes Booleanas con Umbrales

Para modelar redes biológicas se suelen utilizar Redes Booleanas, propuestas para el estudio de redes de regulación genéticas por S. Kauffman en 1969 [25]. Este enfoque utiliza únicamente las reglas de interacción entre los elementos del sistema, las cuales dictan el estado de los elementos con sólo dos posibles valores. A estas redes booleanas se les provee de una dinámica determinista y por lo tanto, al explorar la configuración de la red en el tiempo se llega a estados, llamados atractores, que se repiten. Estos atractores corresponden, en el marco del modelo, a los fenotipos celulares (la configuración de genes prendidos y apagados específico de una célula).

Algunos ejemplos de redes genéticas que se han modelado con redes booleanas son la red del desarrollo en *Arabidopsis Thaliana* [26, 27], la red de interacción responsable de los patrones de segmentación en *Drosophila melanogaster* [28], la red de regulación del ciclo celular de la levadura *Saccharomyces cerevisiae* [29] y de *Saccharomyces pombe* [30]. El éxito de estos trabajos radica en que los patrones de expresión de los fenotipos reales de los organismos se logran reproducir con los atractores, mostrando que estas redes, a pesar de ser simples, logran capturar las propiedades esenciales de las redes biológicas.

En los últimos dos artículos referidos antes se utiliza una variante de Red Booleana llamada Red Booleana con Umbrales (RBU), la cual está definida por un conjunto de N nodos $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$, cada uno con dos posibles estados, $\sigma_i = 1$ (nodo activado o prendido) y $\sigma_i = 0$ (nodo inhibido o apagado). Cada nodo σ_i tiene asignado un conjunto de k_i nodos que pertenecen a la misma red y que regulan su estado. Si σ_{i_j} pertenece al conjunto de reguladores de σ_i (siendo éste el j -ésimo vecino de σ_i) se dice que σ_i tiene una conexión de entrada desde σ_{i_j} . Esta conexión tiene además una intensidad asociada ω_{ii_j} y es un real que está tomado de una distribución $P(\omega)$.

El estado de σ_i está determinado por sus k_i reguladores $\{\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_{k_i}}\}$ de la forma:

$$\sigma_i(t+1) = \begin{cases} 1, & \sum_{j=1}^{k_i} \omega_{ii_j} \sigma_{i_j}(t) > \theta \\ 0, & \sum_{j=1}^{k_i} \omega_{ii_j} \sigma_{i_j}(t) < \theta \\ \sigma_i(t), & \sum_{j=1}^{k_i} \omega_{ii_j} \sigma_{i_j}(t) = \theta \end{cases} \quad (4.1)$$

θ es llamado *umbral de activación*, es el valor que necesariamente debe alcanzarse en la suma $\sum_{j=1}^{k_i} \omega_{ii_j} \sigma_{i_j}(t)$ para que la activación de σ_i se lleve a cabo en el tiempo $t+1$. Este umbral θ es el mismo para todos los nodos. En el presente trabajo más que nada utilizamos la misma conectividad de entrada igual para todos los nodos $k_1 = \dots = k_n = K$, aunque más adelante se utilizará otra distribución, de Poisson. En la figura 4.1 se representa una de estas redes, con $N = 6$. Cabe destacar que esta regla dinámica, debido a ser una función escalón, en general causa que las redes tengan una respuesta no lineal al cambiar alguno de los parámetros que definen la red. Li *et al.*, en [29], presentan la construcción de la red de proteínas relacionadas con el ciclo celular a partir de

la información tomada de la literatura utilizando la regla dinámica 4.1. Ellos encuentran que los atractores de la red son estados que tienen significado biológico y corresponden a los estados de las distintas fases del ciclo celular. Es decir, dada la regla de interacción entre elementos de la célula, emerge como una ruta natural el ciclo celular.

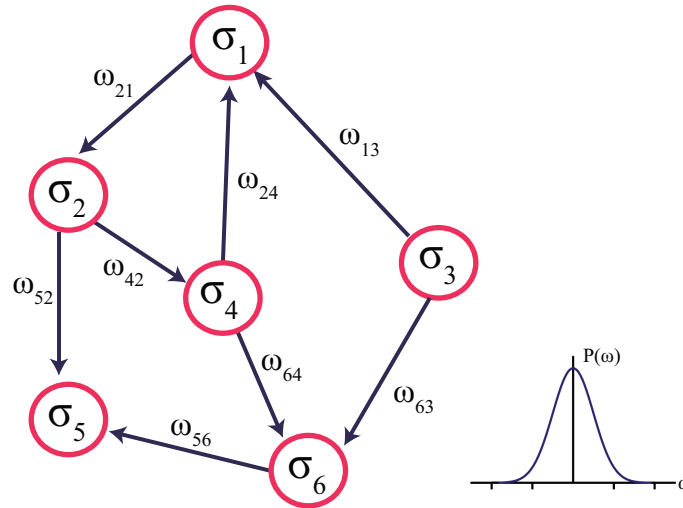


Figura 4.1: Imagen de una red con umbrales con $N = 6$ nodos. Cada vértice $\{ij\}$ representa una conexión dirigida del nodo i al j y ω_{ij} corresponde al peso de la conexión entre estos dos nodos. El valor de ω_{ij} está tomado de una distribución de probabilidad $P(\omega_{ij})$ gaussiana.

En el presente trabajo se utilizan redes booleanas con umbrales como sistema al que se aplican perturbaciones combinadas para contestar la pregunta planteada al final del capítulo anterior. Aunque no se hace un análisis de los atractores, este sistema fue elegido porque los trabajos antes mencionados apuntan a que estas redes son un buen mecanismo para modelar redes biológicas reales, y que además permiten no solo describir al sistema, sino también hacer análisis y predicciones sobre él. Otra razón para usar RBUs es porque estas redes booleanas son prototipos de sistemas dinámicos no lineales, como lo son muchos sistemas biológicos. Las conexiones entre los nodos son escogidas aleatoriamente, todos con conectividad $k_i = K$, de modo que su promedio es también K). Además nos restringimos al valor del umbral $\theta = 0$. La distribución de pesos que principalmente utilizamos es una distribución $P(\omega)$ gaussiana con media $\mu = 0$ y varianza $\sigma = 1$, pero también exploramos resultados al utilizar una distribución uniforme entre $[-1, 1]$ que no se incluyen en el presente trabajo, puesto que los resultados obtenidos con esta distribución son similares a los obtenidos con la distribución gaussiana. La distribución de pesos fue escogida simétrica alrededor de cero para considerar el hecho de que la regulación de un gen puede ser negativa o positiva. La transcripción es el primer proceso para la expresión del gen, el cual comienza solo si ciertas proteínas, llamadas factores de transcripción, están ancladas a la región promotora de un gen, que es el sitio en el ADN en el que se indica dónde comenzar la transcripción. De ser así, la regulación es positiva. La regulación negativa se da si se evita que los factores de transcripción se anclen y por lo tanto no se lleve a cabo la regulación.

Entonces el peso ω_{ij} hace alusión a la afinidad de ciertos factores de transcripción a la región promotora de un gen σ_i , considerando que el un gen representado por σ_j es un gen que está involucrada con la producción de dichos factores de transcripción.

4.2. Régimen dinámico

Se considera que un sistema dinámico posee una dinámica caótica si una pequeña perturbación en la condición inicial cambia mucho la configuración del sistema en el tiempo y que posee una dinámica ordenada si la perturbación desaparece rápidamente en el tiempo o cambia muy poco el estado de la red. Se ha demostrado que las redes booleanas presentan estas dos fases, además de una transición de segundo orden entre ellas [25], siendo el parámetro de control la conectividad de entrada promedio K y el parámetro de orden la distancia de Hamming correspondiente a una red (con conectividad K), entre su configuración en el tiempo de antes y después de ser perturbada una condición inicial.

Sin embargo, la existencia de esta transición de fase se obtuvo para redes booleanas más generales que las RBU que estamos utilizando. De hecho, la transición de fase entre dinámicas ordenadas y caóticas no estaba bien caracterizada para las RBU, por lo que uno de los resultados de este trabajo es precisamente llevar a cabo esta caracterización. Nos interesa hacer esta caracterización porque es necesaria para conocer la respuesta de la red ante perturbaciones.

Nuestras simulaciones numéricas demuestran que efectivamente existen dinámicas ordenadas y caóticas en las RBU, y una transición de fase entre ellas que parece ser de segundo orden. En la figura 4.2 se muestra la configuración de tres RBU operando en las tres fases dinámicas: caótica, crítica y ordenada. Cada configuración corresponde al estado de la red en un tiempo $T = 500$. Los cuadros pequeños corresponden a un nodo de la red y están coloreados de color azul si el nodo está en estado 0 y en amarillo si está en estado 1. Los cuadros color naranja corresponden a los nodos que cambiaron de estado al haber cambiado muy poco la condición inicial, en comparación con su estado, en ese mismo tiempo, que tiene red si no se perturba.

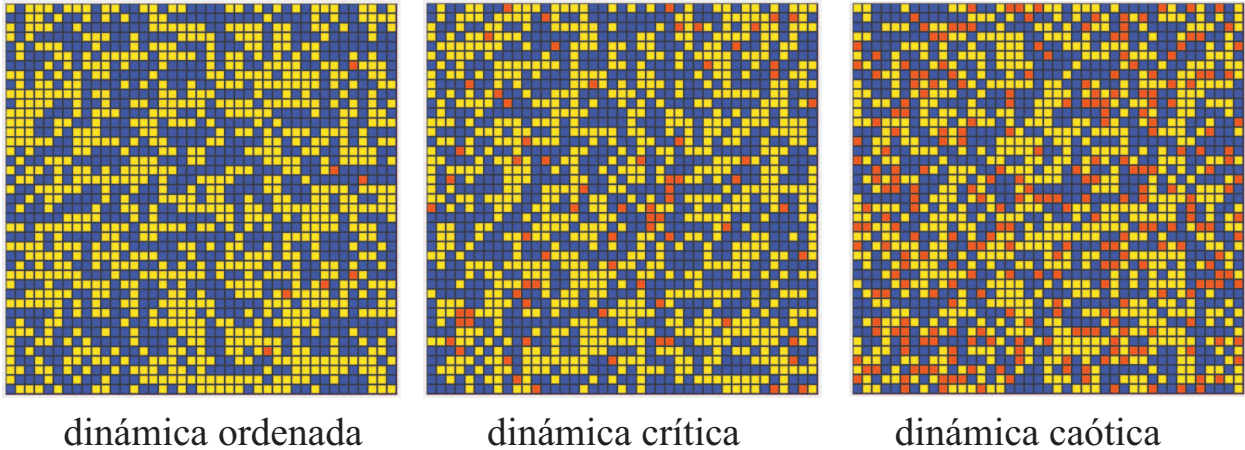


Figura 4.2: Cambios en el estado de redes al inicializarlas con dos condiciones iniciales distintas que difieren una fracción $d = 0,01$ en su configuración. En naranja se muestran los nodos que difieren en su estado una vez cambiada la condición inicial. El azul y amarillo corresponden a nodos con estado prendido y apagado, respectivamente. Las tres redes tienen $N = 1600$ nodos, corresponden a la configuración en un tiempo $T = 500$ y tienen una conectividad (cada nodo): $K = 2$ para la que tiene dinámica ordenada, $K = 5$ para dinámica crítica y $K = 10$ para el caso caótico. Las conexiones de cada nodo son escogidas aleatoriamente de toda red con distribución uniforme.

Para determinar el régimen dinámico de una red con N nodos y una conectividad K , asigno a cada nodo una configuración aleatoria $\Sigma_0 = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ y evoluciono la red por un tiempo T . Luego inicializo la misma red con otra condición inicial $\tilde{\Sigma}_0 = \{\tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_N\}$ que difiera

en una fracción d de elementos de Σ . Dejo evolucionar la dinámica de las redes por el mismo número de pasos de tiempo. Luego comparo las configuraciones de estos dos estados a cada paso de tiempo utilizando la distancia de Hamming, que está definida como:

$$h(t) = \frac{1}{N} \sum_{i=0}^N |\sigma_i(t) - \tilde{\sigma}_i(t)| \quad (4.2)$$

Intuitivamente, si se obtiene una distancia de Hamming $h(t)$ grande, calculada con una d pequeña significa que la red está en un régimen caótico. En cambio cuando $h(t) \rightarrow 0$ se trata de redes con dinámica ordenada. Finalmente, si $h(t) \sim d$ la dinámica es crítica. En la figura 4.3 se muestra la distancia de Hamming en el tiempo, en escala logarítmica, para tres redes con $N = 1000$, $K = 2$, 5 y 10 respectivamente, habiendo cambiado la configuración inicial una fracción de nodos $d = 0,01$.

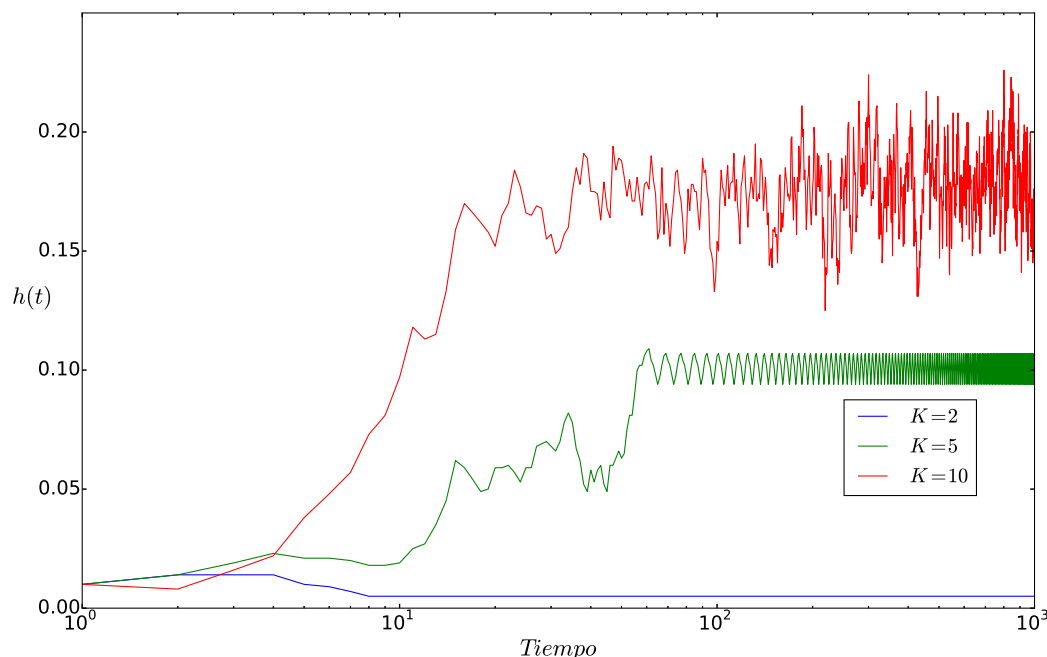


Figura 4.3: Distancia de Hamming en el tiempo, correspondiente a tres redes con $N = 1000$ y conectividades de $K = 2$ (color azul), 5 (verde) y 10 (rojo). La distancia se midió entre el estado de las redes que difieren únicamente en una fracción de $d = 0,01$ de elementos en la condición inicial. Para la red con $K = 2$, a pesar de haberse perturbado la red evolucionó al mismo estado. En el caso $K = 5$ la red rápidamente llega a un estado periódico. La red con $K = 10$, además de tener en promedio una distancia de Hamming mayor, ésta fluctúa aleatoriamente alrededor de un valor promedio estacionario.

Para conectividad $K = 2$ la distancia Hamming adquiere un valor constante muy cercano a 0, lo que indica que la dinámica es ordenada.

Cuando se aumenta la conectividad, la distancia de Hamming también aumenta y además muestra variabilidad en el tiempo. Para $K = 5$ la distancia de Hamming tiene periodicidad que es debida a que para cada una de las condiciones iniciales la red llega a un atractor (que puede ser

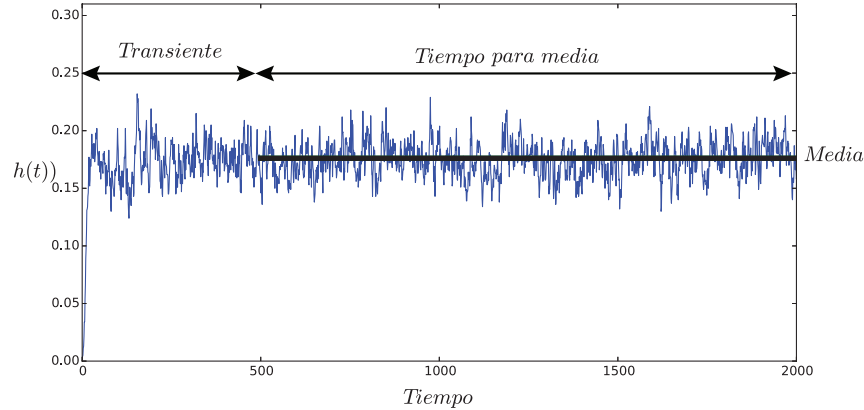


Figura 4.4: Serie temporal de la distancia de Hamming entre la configuración de una red y la configuración de la misma habiendo cambiado su estado inicial ligeramente. Se muestra en la imagen que tras dejar pasar un transiente T_{Trans} , la serie llega a un estado estacionario. El promedio temporal de esta serie desde T_{Trans} se considera como la medida de la perturbación a de la red.

distinto en cada condición inicial). Para $K = 10$ no parece haber un patrón periódico que revele que la red ha llegado a un atractor. Por el contrario, la distancia Hamming se mantiene fluctuando de manera más o menos aleatoria alrededor de un valor promedio que parece haber llegado a un estado estacionario.

La distancia de Hamming promedio en el tiempo $\langle h \rangle_T$ se define como

$$\langle h \rangle_T = \frac{1}{T} \int_0^T h(t) dt$$

Cuando $\langle h \rangle_T$ ya no cambia con T , el sistema ha llegado a un estado estacionario. En la práctica, para calcular el valor estacionario de la distancia Hamming dejamos correr la dinámica de la red un tiempo transitorio T_{trans} y después promediamos la distancia Hamming desde T_{tran} hasta T , como se indica en la figura 4.4. Al valor estacionario de la distancia Hamming lo denotaremos simplemente como h .

En la figura 4.5 se muestra la distancia de Hamming promedio h como función de la conectividad K de la red. Al variar K se observa que hay una transición de fase de segundo orden en donde el punto crítico parece estar en $K_c = 5$. Para hacer esta gráfica se promedió $h(t)$ desde $T_{trans} = 500$ hasta $T = 1000$, y luego promediando sobre 20 redes y 10 condiciones iniciales para cada una. La fracción de elementos que se cambian en las condiciones iniciales es $d = 0,01$.

Llamo ahora a *nodo congelado* a un nodo que no cambia su estado a lo largo del tiempo al utilizar la regla dada por la eq. 4.1. En la figura 4.6 se muestra $1 - f$, donde f es la fracción de nodos congelados en una red, como función de su conectividad K . Para calcular calcular el número de nodos congelados inicializamos una red con conectividad promedio K con 10 condiciones iniciales distintas y cada vez se deja evolucionar la red y se cuenta el número de nodos que después de cierto tiempo $T_{trans} = 500$ permanecen en el mismo estado (ya sea 0 o 1) en el tiempo. La gráfica que se obtiene de este cálculo muestra también en un comportamiento similar a la gráfica de la distancia de Hamming: una transición de fase que parece ocurrir en $K_c = 5$.

4.2.1. Análisis de transición de fase

Los resultados antes mostrados sugieren que las RBU tienen una transición de fase con un punto crítico en $K_c = 5$. Sin embargo esta transición no ha sido caracterizada antes, a pesar del amplio

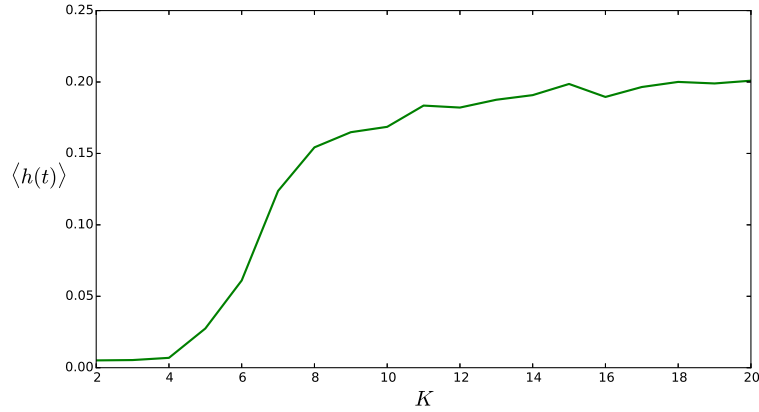


Figura 4.5: Promedio de la distancia de Hamming $\langle h(t) \rangle$ en estado estacionario, tomando un rango de tiempo desde $T_{trans} = 500$ hasta $T = 1000$ para las distintas conectividades. Para cada conectividad la distancia de Hamming corresponde a un promedio sobre 40 redes y 20 condiciones iniciales distintas cada una. La fracción de nodos perturbados en la condición inicial es $d = 0,01$. Esta gráfica muestra la transición de fase de las redes booleanas con umbrales, como función de su conectividad $K = 5$.

uso que se le ha dado a estas redes en modelos de regulación genética. Por esta razón nos interesa caracterizar bien esta transición de fase. En particular, nos interesa saber si efectivamente existe una transición de fase de segundo orden, o si las curvas mostradas en las figuras 4.5 y 4.6 crecen continuamente desde $K = 0$ y la aparente transición que se observa es simplemente un efecto de tamaño finito.

Cálculo semi-analítico de la transición

En el apéndice A muestro un cálculo semi-analítico de la distancia de Hamming h_t en una red con conectividad uniforme K . La gráfica resultante se muestra en 4.7. Esta gráfica valida la existencia de dos fases y un punto crítico en $K = 5$ como se sugería de la distancia de Hamming calculada de la simulación de las redes booleanas.

Autocorrelación

Como una prueba más para corroborar la naturaleza no trivial del punto crítico en la transición de fase mostrada en las figuras 4.5, 4.6 y 4.7, calculé la autocorrelación $A(k)$ de las series temporales de la distancia de Hamming, como las que se muestran en la figura 4.3. La función de autocorrelación de la función $h(t)$ está definida como:

$$A(k) = \frac{1}{(n-k)\sigma^2} \sum_{t=1}^{n-k} (h(t) - h)(h(t+k) - h). \quad (4.3)$$

donde h es el promedio de $h(t)$ definido anteriormente y σ^2 la varianza alrededor de este promedio. k es llamado el desplazamiento temporal. En el régimen caótico y ordenado de un sistema dinámico la función de autocorrelación decae exponencialmente de la manera:

$$A(k) \sim e^{-k/\tau} \quad (4.4)$$

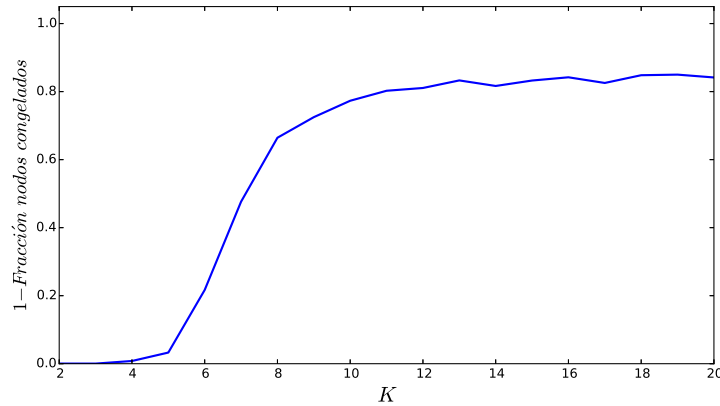


Figura 4.6: $1 - f$, donde f es la fracción de nodos congelados en la dinámica de redes con número de nodos $N = 2000$ y con conectividad uniforme K . Cada punto de la gráfica es un promedio sobre 40 redes y 20 condiciones iniciales para cada red. Se observa una transición de fase en donde todos los nodos están congelados para $K < K_c = 5$ y comienzan gradualmente a descongelarse a partir de allí.

El parámetro τ se llama el tiempo de relajación. En la transición de fase la autocorrelación decae como una ley de potencias, por lo que el ajuste a la exponencial mostraría que τ diverge o, al menos, que tiene el valor más grande. A este fenómeno se le llama *Critical Slowing Down* [31, 32].

La autocorrelación $A(k)$ se muestran en la figura 4.8, para conectividades $K = 2, 4, 5, 6, 7$ y 10. Para cada conectividad se hizo un promedio de la autocorrelación de la serie de la distancia de Hamming de 20 condiciones iniciales distintas.

Para determinar el tiempo de relajación τ para las distintas conectividades, a cada función de autocorrelación le ajusté una exponencial de la misma forma que la ecuación 4.4, utilizando mínimos cuadrados. El tiempo de relajación obtenido de este ajuste se muestra en la figura 4.9, en color azul.

Otra forma de calcular el tiempo de relajación τ es utilizando la ecuación:

$$\tau = - \left\langle \frac{k}{\log \left| \frac{A(k)}{A(0)} \right|} \right\rangle. \quad (4.5)$$

en donde el promedio se hace sobre el desplazamiento temporal k [33]. El resultado de τ obtenido con esta ecuación se muestra también en la figura 4.9 en color verde. En ambos casos, utilizando el ajuste o la ecuación (4.5) para calcular τ , se obtiene que la conectividad para la cual el tiempo de relajación alcanza un máximo corresponde a $K = 6$. Esto sugiere que el punto crítico está en $K = 6$ y no en $K = 5$, como parecía de todas las figuras anteriores (figuras 4.6, 4.5 y 4.6 y en 4.7).

Tenemos pues que de acuerdo al cálculo del tiempo de relajación, la transición de fase ocurre en $K_c = 6$, mientras que las gráficas de la distancia Hamming y la fracción de nodos congelados indican que la transición de fase ocurre en $K_c = 5$. Una posible explicación de esta aparente contradicción podría ser que el punto crítico no sea un entero, sino entre $K = 5$ y $K = 6$. Para comprobar esta hipótesis calculamos en parámetro de orden para redes con topología de Poisson, las cuales se caracterizan por poder tener una conectividad promedio arbitraria, es decir, no necesariamente entera. Esto nos permitió calcular el parámetro de orden para conectividades promedio entre 5 y 6. En la figura 4.10 se muestra el resultado. Aunque los efectos de tamaño finito hacen difícil

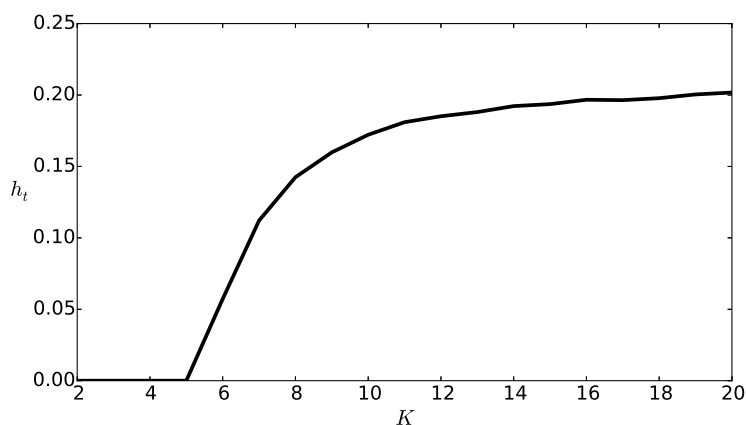


Figura 4.7: Gráfica de la ecuación de la distancia de Hamming (ecuación (A.3)) obtenida del cálculo semi-analítico, como función de la conectividad K . Esta gráfica valida la transición de fase y punto crítico en $K = 5$, como se preveía de la simulación (figura 4.5).

ver exactamente donde está el punto crítico, la curva para redes de Poisson se desplazó hacia la derecha respecto a la curva con conectividades enteras, lo cual es consistente con nuestra hipótesis de que el punto crítico ocurre entre $K = 5$ y $K = 6$. En lo sucesivo solamente considero redes con conectividad entera.

La transición de fase debería poder caracterizarse calculando las correlaciones temporales de cualquier variable dinámica global del sistema. Por lo tanto, en vez de considerar la serie temporal de la distancia de Hamming, consideremos ahora la serie temporal de la fracción de nodos prendidos (en estado 1), y calculemos la función de autocorrelación de esta serie temporal. El tiempo de relajación de esta nueva serie temporal alcanza su máximo en $K_c = 5$, como se expone en la fig. 4.11.

4.2.2. Análisis usando teoría de la información

Existe evidencia de que muchos sistemas físicos y biológicos operan en régimen crítico, por ejemplo avalanchas en la dinámica neuronal del cerebro [Chialvo2007, 34, 35, 36], terremotos [37], precipitación atmosférica [38], regulación genética [39, 40] entre otros. En el caso de los sistemas biológicos, varios grupos de investigación aseveran que la razón de que su dinámica sea crítica es porque esto les permite tener balance entre dos cualidades dinámicas: robustez ante perturbaciones (típico de sistemas ordenados), que permite a un sistema mantener propiedades indispensables para él; la otra es flexibilidad o adaptabilidad (típica de régimen caótico), osea que el sistema puede modificarse una vez que fue perturbado (por ejemplo por cambios en el ambiente) y de este modo adaptarse a las nuevas condiciones en las que se encuentra [41, 39, 40, 42]. Los sistemas críticos no están congelados pero tampoco responden excesivamente a las perturbaciones. Además, la longitud de correlación en algunos sistemas críticos diverge, lo cual les permite responder de forma colectiva a las perturbaciones externas.

Una de las propiedades especiales que se les ha adjudicado a los sistemas operando en punto crítico es tener la capacidad de almacenamiento y transferencia de información más eficaz que en los regímenes caótico y ordenado [43]. Sin embargo hay pocos trabajos que demuestren esto [44]. Dado que nosotros queremos caracterizar la respuesta dinámica de red ante múltiples perturbaciones, nos interesa saber cómo se propaga la perturbación a lo largo de la red. Es decir, si perturbamos un

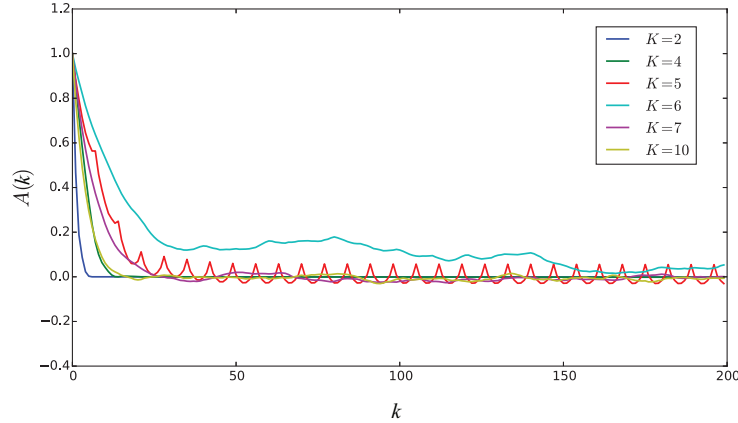


Figura 4.8: Función de autocorrelación $A(k)$ de la distancia de Hamming en el tiempo para redes con distintas conectividades. Cada una de las series corresponden a un promedio sobre 20 condiciones iniciales. Para redes con conectividad $K = 5$ y $K = 6$ la autocorrelación tiene un decaimiento menor que para redes de conectividad más alta o baja.

conjunto de nodos, esa información cómo se propagará al resto de la red. Por esta razón, analizo la capacidad de almacenamiento y transferencia de información de las RBU operando en los tres regímenes dinámicos.

Almacenamiento de información

La complejidad de Kolmogorov es una medida de los recursos de computabilidad mínimos necesarios para especificar un objeto o, dicho de otra forma, una medida de la información necesaria para describir dicho objeto.

Específicamente la Complejidad de Kolmogorov se define como el largo del programa p más corto que describa al objeto x en un máquina de Turing T tal que $T(p) = x$:

$$K_T := \min_p \{l(p) : T(p) = x\} \quad (4.6)$$

donde $l(p)$ es el largo de p medido en bits. Una manera aproximada de medir esta complejidad es utilizando compresores sin pérdida, por ejemplo *7-Zip*. Estos programas comprimen a una cadena binaria buscando “palabras” repetidas y quedándose solamente con aquellas que son esenciales para reconstruir la cadena. Una cadena de texto generada aleatoriamente, en principio, no se puede comprimir, mientras que para una cadena de texto de un mismo símbolo repetido, el tamaño de la cadena comprimida será muy pequeña. Esto habla pues de la complejidad de estas cadenas y de la información que pueden almacenar: entre más grande es la complejidad de Kolmogorov de la cadena, menos compresible es y más información podría almacenar, y viceversa.

La metodología que utilizo para calcular la complejidad del estado las redes es la siguiente: genero la trayectoria temporal de una red con N nodos, dejando pasar primero un tiempo transiente T_{trans} y hasta un tiempo T . Esto genera una matriz de ceros y unos de tamaño $N \times (T_{trans} - T)$ que guardo en un archivo. Posteriormente comprimo este archivo con el compresor *7-Zip* (que es considerado como el “state of the art” entre compresores que existen actualmente). La complejidad de Kolmogorov C_{kol} la calculamos entonces de forma aproximada como:

$$C_{kol} = \frac{\text{Tamaño del archivo comprimido (bytes)}}{\text{Tamaño del archivo original (bytes)}} \quad (4.7)$$

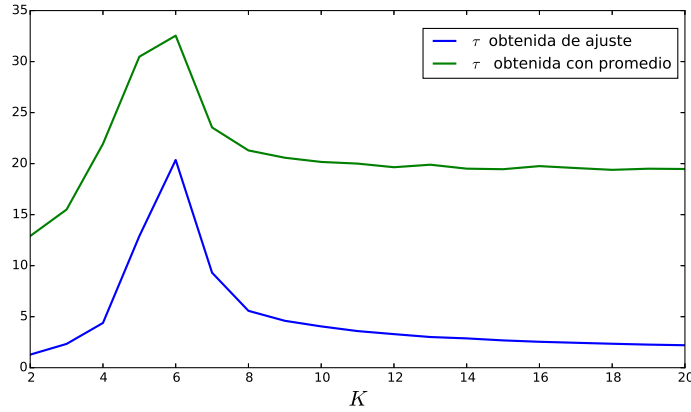


Figura 4.9: Tiempo de Relajación τ calculado de serie temporal de la distancia de Hamming para una redes con conectividad K . En color azul se muestra τ obtenida haciendo un ajuste de una función exponencial a la función de autocorrelación de cada red. En verde se gráfica τ obtenida a partir de la ecuación (4.5), en el promedio se consideró hasta el desfase $k = 200$. Usando ambos métodos se obtiene que para conectividad $K = 5$ y $K = 6$ el tiempo de relajación es máximo, indicando con ello que alrededor de ese punto se encuentra el punto crítico.

En la figura 4.12 muestro C_{kol} como función de la conectividad K de la red. Esta gráfica muestra de nuevo la transición de fase que ya habíamos observado utilizando la distancia de Hamming promedio. Para conectividad $K \leq 5$, la complejidad es cercana a $C_{kol} = 0$ y a partir de ahí ésta comienza a aumentar. Lo que es importante de esta gráfica es que muestra que la capacidad de almacenamiento de información de la red en el punto crítico no es máxima, como ocurre en otras transiciones de fase [43].

Transferencia de información

Para medir ahora la capacidad que tiene la red de transmitir información entre sus nodos, utilizo la *información mutua*, la cual se define entre dos variables aleatorias discretas x y y como:

$$\text{MI}(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (4.8)$$

donde $p(x)$ y $p(y)$ son las probabilidades marginales de las variables x y y respectivamente, y $p(x, y)$ es su distribución conjunta.

Para calcular la información mutua MI_{nm} entre dos nodos específicos de la red, σ_n y σ_m , definimos las variables aleatorias x y y como $x = \sigma_n(t)$ y $y = \sigma_m(t + 1)$. Nótese que en este caso la probabilidad conjunta $p(x, y)$ que aparece en la ecuación (4.8) se calcula entre estados defasados un paso de tiempo. Este desfaseamiento es lo que permite denominar en este caso a la información mutua MI_{nm} como la información *transferida* del nodo σ_n al σ_m . La información mutua de toda la red la definimos como el promedio de MI_{nm} sobre todas las parejas de nodos:

$$\text{MI}_R = \frac{1}{N(N-1)} \sum_{i \neq j}^N \text{MI}_{ij} . \quad (4.9)$$

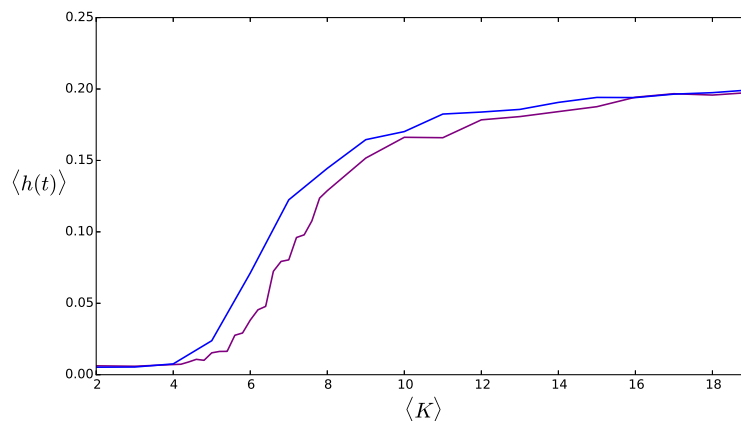


Figura 4.10: Promedio de la distancia de Hamming $\langle h(t) \rangle$ en estado estacionario para dos tipos de redes: redes con conectividad entera (azul), $\langle K \rangle = K$ y para redes con conectividad de sus nodos variable, con distribución de Poisson (morado). Nótese que la curva correspondiente a las redes de Poisson está recorrida hacia la derecha, lo cual parece indicar que el punto crítico ocurre entre $K = 5$ y $K = 6$.

La figura 4.13 muestra la información mutua promedio sobre 70 redes con $N = 3000$ nodos, como función de la conectividad K . Esta gráfica muestra de nuevo la misma forma que la gráfica de la transición de fase obtenida con la distancia de Hamming (fig. 4.10), mas no se observa que haya un máximo en la transferencia de información en el punto crítico, $K_c = 5$, como se presume en la literatura sobre la transferencia de información en los puntos críticos.

4.3. Conclusiones

El análisis presentado en este capítulo muestra que las RBU exhiben una transición de fase de segundo orden determinada por la conectividad de la red. Esta transición se manifiesta utilizando diversos parámetros de orden: la distancia Hamming, la fracción de nodos congelados, la complejidad de Kolmogorov y la información mutua. En todos estos casos, las gráficas del parámetro de orden como función de la conectividad parecen indicar que la transición ocurre en $K_c = 5$. Sin embargo, cuando calculamos los tiempos de relajación de la serie temporal de la distancia Hamming (a partir de la función de autocorrelación), no queda claro si la transición de fase ocurre en $K_c = 5$ o en $K_c = 6$, o en algún punto intermedio. Independientemente del valor exacto de K_c , lo que sí es cierto es que existe una transición de fase de segundo orden entre dinámicas ordenadas y dinámicas caóticas que se refleja de muchas formas. Esto es importante porque la respuesta de la red ante perturbaciones claramente dependerá de la fase dinámica en la que esté operando.

No se observa en el punto crítico un valor máximo de almacenamiento ni transferencia de información en las RBUs, calculadas con la complejidad de Kolmogorov y la información mutua, respectivamente, como se hubiera esperado por lo que se dice en la literatura.

La fracción de nodos congelados ayuda a entender las gráficas de la distancia de Hamming y de la complejidad de Kolmogorov. Para conectividades bajas, como una gran cantidad de nodos están congelados, es claro que en ese caso la complejidad sea casi cero, ya que es necesaria poca información para describir estos estados. También es comprensible que en esta fase donde hay muchos nodos congelados las perturbaciones no se propaguen, o se propaguen muy poco. Por lo

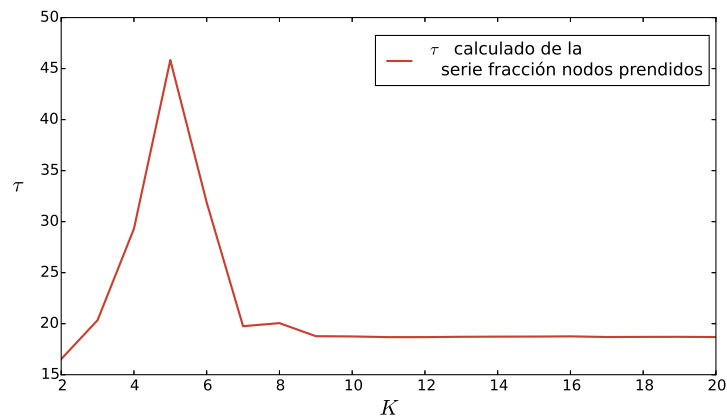


Figura 4.11: Tiempo de Relajación τ de la serie temporal de la fracción de nodos prendidos en la configuración de la red. Cálculo obtenido con la ecuación 4.5. Nótese que el punto máximo ocurre en $K_c = 5$.

tanto, cualquier perturbación inicial, o desaparece o se queda confinada a un conjunto pequeño de nodos, lo que significa que la distancia de Hamming o se hace cero o permanecerá muy cercana a la fracción de nodos que fueron perturbados inicialmente.

A medida que los nodos comienzan a descongelarse, la configuración de las redes va teniendo más variedad o aleatoriedad, por lo que la complejidad de Kolmogorov aumenta y las perturbaciones se propagan más fácilmente.

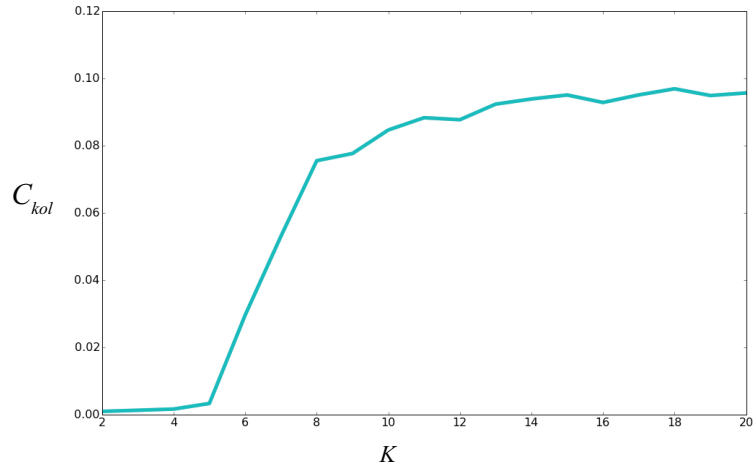


Figura 4.12: Complejidad de Kolmogorov C_{kol} del estado de una red en el tiempo como función de su conectividad K , medida como refiere la relación (4.7). Esta gráfica refleja que no hay un almacenamiento máximo de información en el punto crítico, sino que la complejidad aumenta a partir de ese punto y se mantiene constantes como función de K .

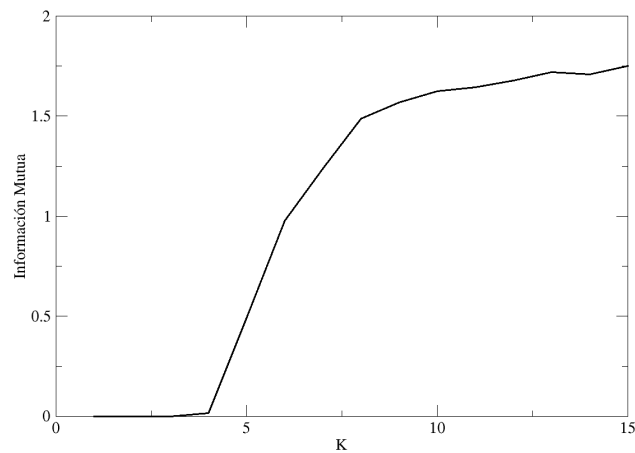


Figura 4.13: Información mutua de redes booleanas con umbrales como función de su conectividad promedio K . No se observa que la información transferida sea máxima en el punto crítico $K_c = 5$, sino que aumenta con la conectividad, de igual manera que la distancia de Hamming.

5. Descomposición del Efecto de Perturbaciones en RBU

En el capítulo 4 describimos que si el estado de un gen, por ejemplo σ_i , es regulado por el estado de otro gen, σ_j , esta regulación está dada en función de la afinidad del factor de transcripción a la región promotora del gen, que está representada en este modelo de redes booleanas con el peso ω_{ij} . El mecanismo de ciertos antibióticos consiste en modificar esa afinidad para evitar o promover que genes específicos se expresen. Por esta razón en este modelo las perturbaciones consisten en modificar los pesos ω_{ij} .

5.1. Perturbaciones en Pesos

Una perturbación en pesos, que denotaremos como $A_P(\tau, C)$, cambia los pesos de los enlaces de salida de un conjunto de τ nodos blanco, $\{\sigma_{n_1}, \sigma_{n_2}, \dots, \sigma_{n_\tau}\}$, y la magnitud de este cambio está determinado por la *intensidad* o *concentración* C , de la manera que sigue:

Si σ_{n_i} es un blanco de la perturbación $A_P(\tau, C)$ y a la vez es regulador del nodo σ_j con una peso ω_{jn_i} , al aplicar la perturbación, este peso cambia como $\omega_{jn_i} \rightarrow \omega_{jn_i} + C \xi_{jn_i}$, donde ξ_{jn_i} es un número real tomado de la misma distribución que ω_{jn_i} y es un valor fijo asociado a la perturbación, como se muestra en la figura 5.1. C es un parámetro real que describe la intensidad de la perturbación y su importancia radica en que permite aplicar con distinta intensidad una misma perturbación, por esa razón y en analogía con los antibióticos, en adelante se le hace referencia como concentración.

En la imagen 5.2 se muestra cómo una vez que se aplican perturbaciones individualmente a una red ésta cambia su estado en el tiempo en comparación a su estado sin perturbar, esto medido con la distancia de Hamming. La imagen corresponde a la misma red, pero en cada caso se aplicó una perturbación individual que difiere con el número de blancos de ésta, $A_1(2, 1)$, $A_2(5, 1)$ y $A_3(10, 1)$. Se aprecia que las perturbaciones con mayor número de nodos blanco son las que generan una mayor distancia de Hamming. La red a la que se aplicaron las perturbaciones tiene una conectividad $K = 10$, por lo que ese resultado es comprensible, ya que es una red en régimen caótico y las perturbaciones se propagan mejor en ese régimen. La elección de esta cantidad de nodos blanco, $\tau_1 = 2$, $\tau_2 = 5$ y $\tau_3 = 10$ para las tres perturbaciones, respectivamente, la hice para mostrar cómo al tener mayor número de nodos blanco el efecto de la perturbación es mayor, además de para hacer a estas perturbaciones lo más distintas entre ellas posible.

El efecto de la perturbación $A(\tau, C)$ en una red, lo mido con el promedio de la distancia de Hamming $h(t)$ entre el estado de la red sin perturbar y el estado de ésta tras aplicar la perturbación, desde un tiempo T_{trans} hasta un tiempo T , además de promediar también sobre varias condiciones iniciales. La razón para utilizar el tiempo transiente es, como se hizo en el caso de perturbaciones en las condiciones iniciales en la sección 4.2, para calcular el promedio de la distancia de Hamming una vez que se llegó a un estado estacionario. En adelante a este promedio de la distancia de Hamming tras la perturbación A_i lo denoto como h_i ; si se aplican simultáneamente dos perturbaciones A_i y A_j , lo cual denotamos como $A_{ij} = A_i A_j$, la distancia de Hamming asociada será h_{ij} , y con

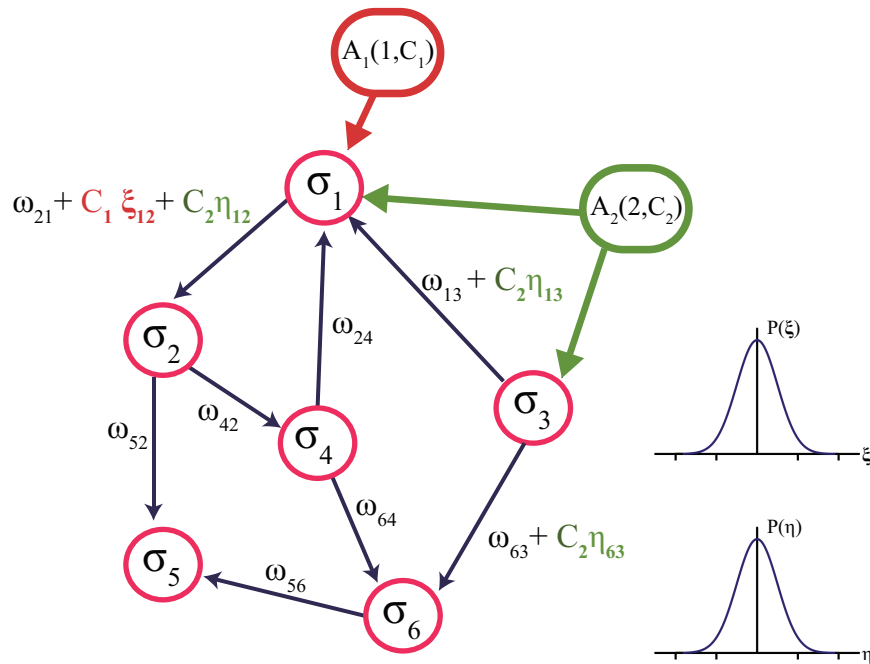


Figura 5.1: Red booleana con umbrales a la que se aplica perturbaciones en los pesos $A_i(\tau_i, C_i)$. La perturbación $A_1(1, C_1)$ modifica los pesos ω_{ij} de las conexiones de salida de sus $\tau = 1$ nodos blanco, sumándoles una nueva variable ξ_{ij} , que está tomada de una distribución gaussiana, y multiplicada por un parámetro C_1 que actúa como la intensidad de la perturbación. La otra perturbación, $A_2(2, C_2)$, actúa de manera igual sobre sus $\tau = 2$ nodos blanco. Las dos perturbaciones tienen un nodo blanco en común, por lo que los términos extra en pesos de sus conexiones de salida se suman.

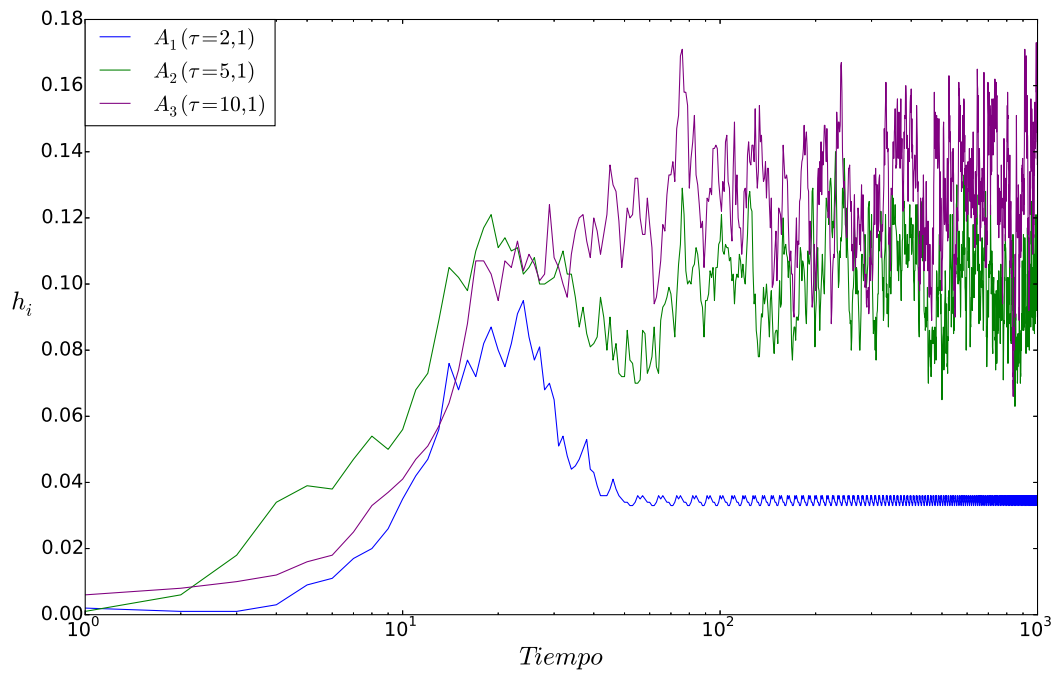


Figura 5.2: Variación temporal de la distancia de Hamming h_i de la configuración de una red booleana tras aplicarse independientemente perturbaciones con distinto número de blancos. Las perturbaciones son $A_1(2, 1)$, $A_2(5, 1)$ y $A_3(10, 1)$ y son aplicadas a una red que tiene $N = 1000$ nodos y con conectividad $K = 10$. Se observa que para las perturbaciones con mayor número de nodos la distancia de Hamming alcanza valores más altos, además de que para la perturbación A_1 la red llega rápidamente a un estado estacionario de periodicidad pequeña, lo cual ya no pasa para las perturbaciones con más blancos. Este resultado es típico para redes en la región caótica. El eje del tiempo se muestra en escala logarítmica,

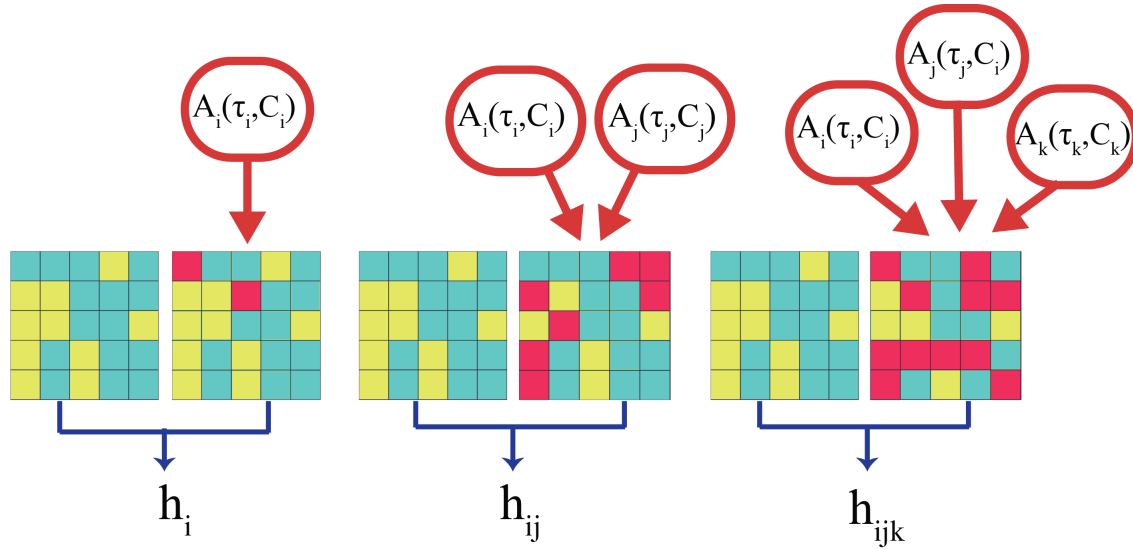


Figura 5.3: Medida de las perturbaciones en pesos en las redes booleanas con umbrales. Cuando se aplica una sola perturbación $A_i(\tau_i, C_i)$, se obtiene h_i , la distancia de Hamming promedio de la configuración en el tiempo de la red perturbada con la red sin perturbar. Al aplicar dos perturbaciones $A_i(\tau_i, C_i)$ y $A_j(\tau_j, C_j)$ simultáneamente y compararla con la red sin perturbar se obtiene h_{ij} . De manera similar al aplicar tres perturbaciones simultáneamente.

tres perturbaciones $A_{ijk} = A_i A_j A_k$, la distancia Hamming asociada será h_{ijk} , etc. Esta asignación está representada en la imagen 5.3.

En la figura 5.4 se muestra cómo se modifica la distancia de Hamming h_i para tres perturbaciones distintas aplicadas a una red con $N = 1000$ y $K = 10$, como función de la concentración C . Conforme se aumenta la C el efecto h_i aumenta, sin embargo se observa también que después de cierto crecimiento, aunque la concentración aumente, la distancia de Hamming ya no aumenta más: la red se satura. La concentración para la que esta distancia ya no parece crecer, depende tanto de la red como de la perturbación, particularmente del número de nodos blanco que ésta tenga asignados. Por esa razón en adelante solamente tomo en cuenta concentraciones hasta el valor máximo $C_{max} = 2$.

En la figura 5.5 se observa el efecto de aplicar distintas perturbaciones en redes con diferentes conectividades. En cada red las tres perturbaciones aplicadas son $A_1(2, 1)$, $A_2(5, 1)$ y $A_3(10, 1)$. Para cada red se muestra el promedio sobre 10 condiciones iniciales. La figura permite observar que la tendencia de la distancia de Hamming tras aplicar las perturbaciones es de aumentar conforme la conectividad de la red aumenta, independientemente del número de nodos blanco que afecte la perturbación. Se observa también que no hay un crecimiento constante de h_i con la conectividad, sino que alrededor de $K = 10$, la distancia crece ya muy poco, también no importando el número de nodos blanco. Otra cosa relevante es que en algunos casos, la distancia de Hamming debida a cierta perturbación puede ser menor que la distancia obtenida debida a otra perturbación con menos nodos blanco, como para caso de $K = 5$ y $K = 12$ en la figura.

5.2. Efecto de múltiples perturbaciones

En los mapas de calor de la figura 5.6, se muestra la distancia de Hamming h_{ij} , cuando se aplicaron dos perturbaciones $A_1(10, C_1)$, $A_2(15, C_2)$ simultáneamente con distintas combinaciones

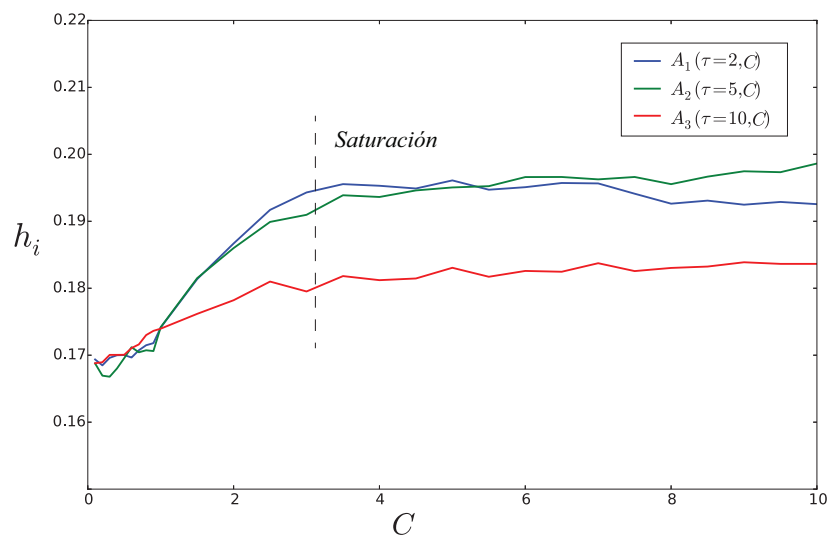


Figura 5.4: Distancia de Hamming promedio h_i correspondiente a la aplicación de tres perturbaciones distintas como función de la concentración C con la que se aplicaron a una red con $N = 1000$ y $K = 10$. El promedio de la distancia de Hamming en el tiempo se hace desde $T_{trans} = 500$ hasta $T = 1000$. La línea punteada vertical muestra que después de una concentración arriba de $C = 2$, h_i deja de aumentar, el sistema se *satura*. Por esa razón en los resultados que siguen siempre se usan concentraciones menores a $C = 2$.

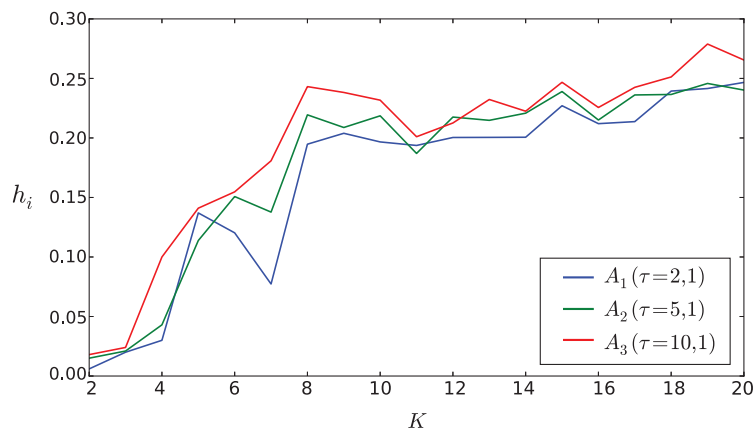


Figura 5.5: Distancia de Hamming calculada tras aplicar perturbaciones con distinto número de blancos, $A_1(2, 1)$, $A_2(5, 1)$ y $A_3(10, 1)$, como función de la conectividad de la red. La distancia de Hamming debida a estas perturbaciones aumenta con la conectividad asemejándose cualitativamente a la curva en que se observa la transición de fase (fig.4.5).

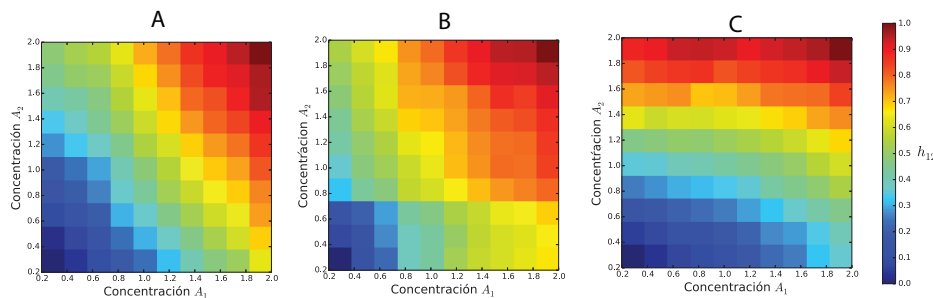


Figura 5.6: Mapas de calor de la distancia de Hamming h_{ij} correspondiente a la combinación de dos perturbaciones $A_1(10, C_1)$ y $A_2(15, C_2)$. El código de color corresponde al valor h_{12} , siendo el valor menor de color azul y el mayor en rojo. Los tres casos corresponden a un tipo de combinación distinto: aditiva en **A**, donde las perturbaciones se aplican en una sola red con $K = 13$; sinérgica en **B**, donde la red tiene $K = 10$, y en **C**, la combinación antagonónica en una red con $K = 19$. Los tres casos corresponden a una sola corrida de una red con $N = 1000$. h_{ij} está renormalizada entre 0 y 1 para facilitar la comparación de los mapas. Esta variedad de mapas tiene parecido a los mapas de tasas de crecimiento en presencia de fármacos (Fig. 3.1).

de concentraciones. Cada cuadro de color en un mapa representa la distancia de Hamming tras aplicar A_1 con concentración C_1 al mismo tiempo que se aplica A_2 con concentración C_2 . Se presentan tres casos en que la respuesta de la red es distinta. En **A**, una red con $K = 13$ tiene una respuesta aditiva. En **B**, la red tiene $K = 10$ y tiene una respuesta sinérgica. Y en **C**, la combinación es antagonónica en una red con $K = 19$. Para los tres casos los datos presentados están normalizados entre 0 y 1, para poder comparar los tres mapas.

La variedad en las perturbaciones, tanto en conexiones y número de blancos, así como la red a las que se aplican, permite tener una gran variedad de diferentes efectos en las combinaciones, tal y como en el caso de los antibióticos, como se observa en la figura 3.2.

En la figura 5.7 se compara el efecto h_i de dos perturbaciones, $A_1(5, C)$, $A_2(10, C)$ aplicadas individualmente y la distancia de Hamming h_{12} tras aplicar las perturbaciones simultáneamente, $A_{12} = A_1 A_2$, como función de la concentración C , manteniendo C igual para las dos perturbaciones. Esta gráfica permite observar más claramente que, en general, aplicar las perturbaciones en conjunto no da como resultado la simple suma de la distancia de Hamming de las perturbaciones individuales, sino que es mucho menos.

5.3. Predicción del efecto de perturbaciones

5.3.1. Predicción con ecuación de Isserlis

Ahora se quiere probar si la misma ecuación de Isserlis, 3.12 y 3.13, utilizada por Wood et al. en [1], para la predicción del efecto de combinaciones de fármacos, puede ser utilizada para predecir el efecto de la combinación de perturbaciones en las RBUs.

En el caso de los fármacos, la medida de su efecto está dada por g_i , variable que cuantifica el crecimiento del patógeno una vez administrado el fármaco D_i . En el caso de las redes, esta cantidad g_i es remplazada por h_i , la distancia de Hamming entre los estados de la red antes y después de aplicada la perturbación A_i . De modo que la ecuación de interés para analizar si permite predecir la respuesta de una red tras haber aplicado combinaciones de tres y de cuatro perturbaciones es:

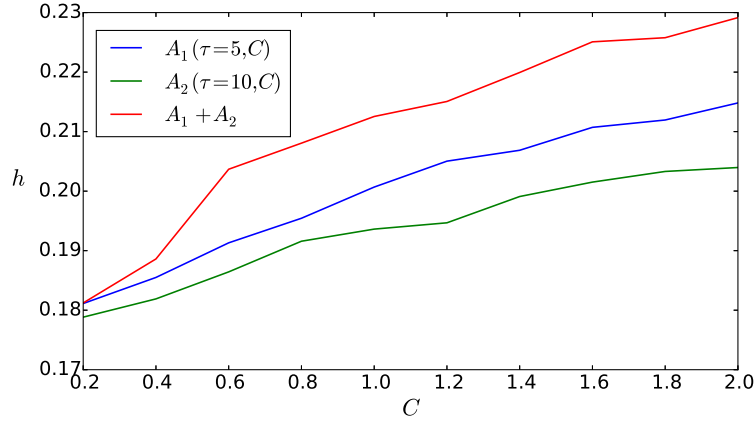


Figura 5.7: Respuesta de una red tras modificarla con dos perturbaciones distintas, $A_1(5, C)$ y $A_2(10, C)$, independientemente y en combinación. La red tiene $N = 1000$ y $K = 10$. Se observa que el efecto h_{12} de la combinación de las perturbaciones no corresponde a la suma aritmética de los efectos h_1 y h_2 de las perturbaciones aplicadas individualmente.

$$\hat{h}_{ijk} = h_i h_{jk} + h_j h_{ik} + h_k h_{ji} - 2h_i h_j h_k, \quad (5.1)$$

$$\hat{\hat{h}}_{ijkl} = h_{ij} h_{kl} + h_{ik} h_{jl} + h_{il} h_{jk} - 2h_i h_j h_k h_l, \quad (5.2)$$

donde los términos de pares h_{ij} se refieren, como antes, a la aplicación de las perturbaciones $A_i(\tau_i, C_i)$ y $A_j(\tau_j, C_j)$ simultáneamente.

Puesto que es posible calcular numéricamente los valores reales h_{ijk} y h_{ijkl} , se puede también evaluar qué tan buena es la predicción de la ecuación de Isserlis. Para esto utilicé una red con $N = 1000$ nodos y conectividad $K = 10$. A esta red apliqué conjuntamente tres perturbaciones $A_1(5, C_1)$, $A_2(15, C_2)$ y $A_3(20, C_3)$, en combinaciones de concentraciones de un rango de 0,2 a 2 en pasos de 0,2, de modo que se tienen 120 combinaciones distintas para la aplicación de la triada de perturbaciones. Cada una de estas combinaciones corresponde a un valor h_{ijk} .

Una vez calculados el conjunto de todos los puntos $\{h_{ijk}\}$ (de la simulación) y $\{\hat{h}_{ijk}\}$ (de la ecuación de Isserlis), normalizo los datos de la manera:

$$\begin{aligned} x_l^* &= \frac{x_l - x^{\min}}{x^{\max} - x^{\min}}, \\ y_l^* &= \frac{y_l - y^{\min}}{y^{\max} - y^{\min}}, \end{aligned} \quad (5.3)$$

donde con x_l me refiero al elemento l del conjunto $\{h_{ijk}\}$ y con y_l al elemento l de $\{\hat{h}_{ijk}\}$; x^{\min} y x^{\max} es el valor menor y mayor del conjunto $\{h_{ijk}\}$ de cada corrida, similarmente y^{\min} y y^{\max} son los valores menor y mayor de $\{\hat{h}_{ijk}\}$, también de cada corrida, por lo que x^{\min} , x^{\max} , y^{\min} y y^{\max} también variables aleatorias. Este rescalamiento es análogo al que se hace en las tasas de crecimiento de las colonias de bacterias. En efecto, la tasa de crecimiento g_i en presencia del antibiótico D_i se reporta como una fracción de la tasa de crecimiento máxima g_0 que se observa en ausencia de antibióticos.

En la figura 5.8, está graficado el valor real h_{ijk} contra el valor predicho \hat{h}_{ijk} . Del lado izquierdo muestro los datos crudos obtenidos del cálculo directamente y del lado derecho los datos normalizados con 5.3. El código de colores corresponde a la suma de las concentraciones $C_i + C_j + C_k$ correspondiendo el color azul a las concentraciones más bajas y el rojo a las más altas, pasando antes por color cyan, verde, amarillo y naranja, para las otras combinaciones de concentraciones.

De ser \hat{h}_{ijk} un buen ajuste de h_{ijk} , los puntos caerían sobre la línea identidad, que está coloreada de azul en la gráfica. La línea de color verde corresponde a cuando en la ecuación de Isserlis, 5.1, todos los términos son iguales, $h_i = h_j = \dots = h_{ik} = h$, es decir si todas las perturbaciones tuvieran el mismo efecto, independientemente del número de blancos τ y la concentración:

$$P(h) = 3h^2 - 2h^2. \quad (5.4)$$

En la figura 5.8a se muestra la comparación para una red con conectividad $K = 2$. Para este caso es evidente que la predicción de la ecuación de Isserlis no es buena, ya que los puntos de la gráfica no se ajustan bien ni siquiera a una función lineal. Para $K = 5$ tampoco se obtiene un buen ajuste lineal, el conjunto de puntos están dispersados por toda la gráfica. Para conectividad $K = 15$ el ajuste a la línea identidad es mucho mejor.

Para medir qué tan buena es la aproximación \hat{h}_{ijk} al valor real h_{ijk} , utilizamos el Coeficiente de Concordancia ρ_c que evalúa el grado en que pares de variables caen sobre la recta identidad [45]. Tiene el valor $\rho_c = 1$ cuando los dos conjuntos de puntos son idénticos y de $\rho_c = -1$ si están anti-correlacionados. En 5.9 se muestra el coeficiente de concordancia para las parejas de puntos (h_{ijk}, \hat{h}_{ijk}) como función de la conectividad de la red a la cual se están aplicando las perturbaciones. Se aprecia que, efectivamente, la ecuación de Isserlis provee de una buena predicción del efecto de múltiples perturbaciones cuando la conectividad de la red en la que se aplican es alta.

En el caso de una combinación de cuatro perturbaciones la ecuación (5.2), que está basada en la ecuación de Isserlis para el cuarto momentos, también provee de una buena aproximación para h_{ijkl} . Esto se muestra en la figura 5.10, donde se gráfica la distancia de Hamming h_{ijkl} asociada a las perturbaciones $A_1(5, C_1)$, $A_2(10, C_2)$, $A_3(15, C_3)$ y $A_4(20, C_4)$ contra el valor \hat{h}_{ijkl} calculado con ecuación de Isserlis. Cada punto de la gráfica corresponde a una una combinación de concentraciones para cada uno de las perturbaciones, las cuales toman los valores desde $C_i = 0,2$ hasta $C_i = 2$. El color de los puntos corresponde a la suma de la combinación de concentraciones, mismo código de colores que el utilizado en la figura 5.8. El coeficiente de concordancia de estos datos es de $\rho_c = 0,959$.

5.3.2. Predicción utilizando promedios

Hasta el momento, se ha mostrado que para redes con una conectividad a partir aproximadamente de $K = 10$ la ecuación de Isserlis provee de una buena predicción del efecto de tres y cuatro perturbaciones aplicadas simultáneamente, sin embargo no se ha dado una explicación para esto.

Como una prueba de que la ecuación de Isserlis funciona no por razones triviales, sino por algo más profundo en la respuesta de la red a las perturbaciones, utilicé el simple promedio de las distancias de Hamming para hacer la predicción. En otras palabras, calculo el promedio de la distancia de Hamming correspondiente a la aplicación de las perturbaciones individuales:

$$h_{ijk}^I = \frac{1}{3}(h_i + h_j + h_k). \quad (5.5)$$

A esta ecuación la llamo en adelante *ecuación de momentos individuales*, haciendo la referencia a la analogía de la distancia de Hamming h_i con el primer momento $\langle x_i \rangle$ de una distribución gaussiana.

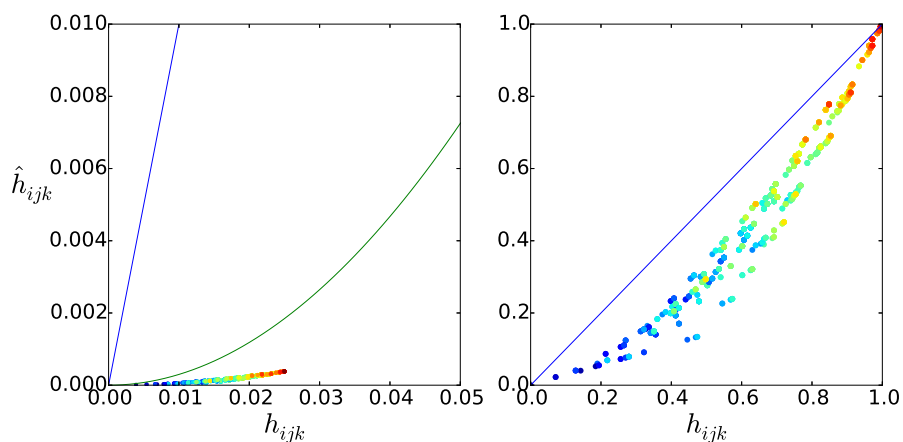
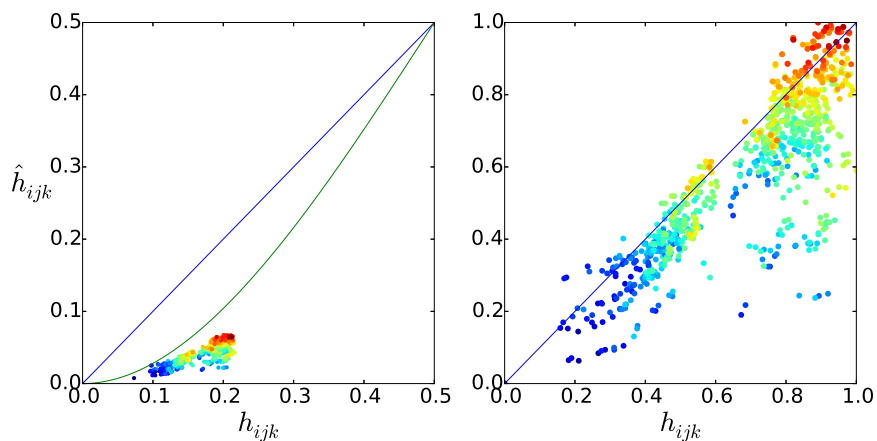
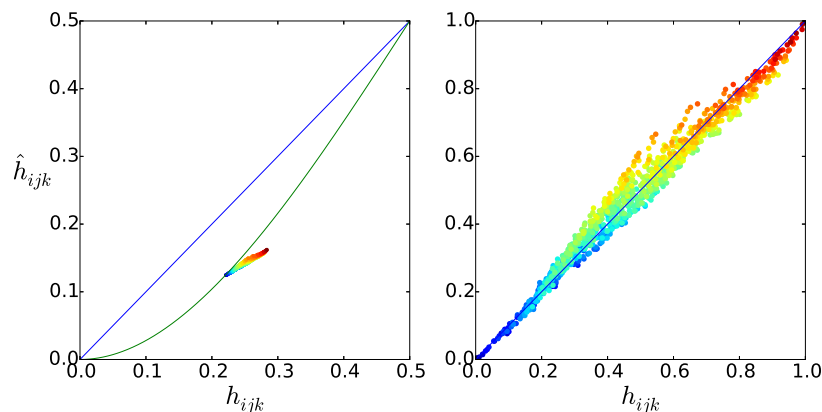
(a) $K = 2$ (b) $K = 5$ (c) $K = 15$

Figura 5.8: Comparación de h_{ijk} y \hat{h}_{ijk} para distintas conectividades: (a) $K = 2$, (b) $K = 5$, (c) $K = 15$. Cada punto corresponde a las posibles concentraciones de entre 0,2 a 2, de la triada de perturbaciones, que tienen un número de nodos blanco $\tau = 5, 10$ y 20 en los tres casos. El código de color corresponde a la suma de concentraciones de las tres perturbaciones, (suma menor en azul y mayor en rojo). Para $K = 2$ (red ordenada) y $K = 5$ (red crítica), la predicción \hat{h}_{ijk} no es una buena aproximación para h_{ijk} . Para $K = 15$ se observa que los puntos caen sobre la recta identidad (línea azul) una vez que están normalizados los datos. Para los tres casos se observa que, sin normalizar los datos, los puntos de concentración menor están cerca del polinomio (5.4).

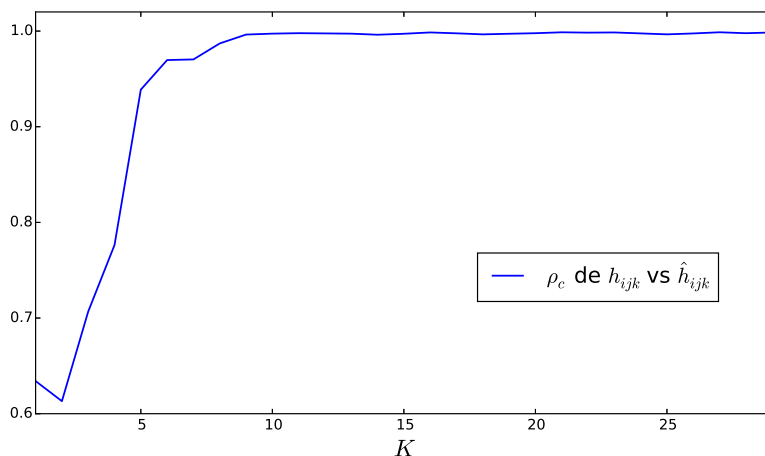


Figura 5.9: Coeficiente de concordancia ρ_c para los puntos (h_{ijk}, \hat{h}_{ijk}) como función de la conectividad K . Cada punto de la gráfica es el promedio para diez realizaciones distintas de redes. Para redes con baja conectividad, $K < 10$ el coeficiente es bajo, lo cual indica que la aproximación de Isserlis con \hat{h}_{ijk} no es buena. Se observa que cuando la conectividad aumenta, también el coeficiente aumenta, para redes en el régimen caótico $K > 5$ la ecuación \hat{h}_{ijk} si es una muy buena aproximación para la distancia real h_{ijk} . Los datos utilizados para calcular el coeficiente están normalizados con las relaciones 5.3.

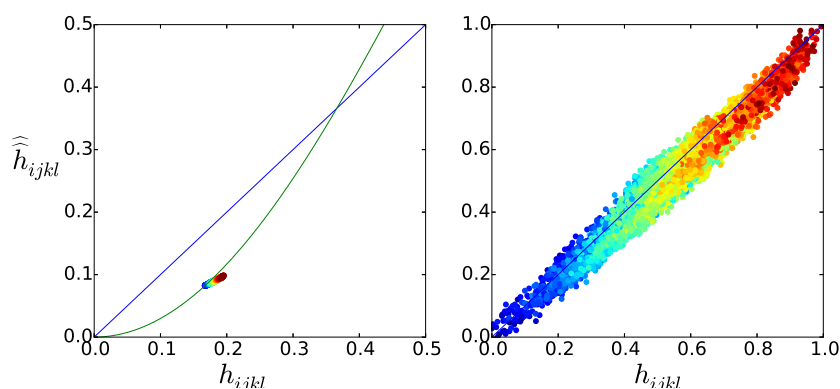


Figura 5.10: Comparación de la distancia de Hamming h_{ijkl} correspondiente a cuatro perturbaciones contra la predicción hecha con la ecuación de Isserlis \hat{h}_{ijkl} . Cada punto corresponde al efecto de aplicar simultáneamente $A_1(5, C_1)$, $A_2(10, C_2)$, $A_3(15, C_3)$ y $A_4(20, C_4)$ con las diferentes combinaciones de concentraciones que van de $C = 0,2$ a 2 . La red tiene $N = 1000$ nodos y conectividad $K = 13$. En la izquierda se muestran los datos originales y en la derecha los datos escalados, los cuales muestran que \hat{h}_{ijkl} da una buena aproximación a la distancia real, con un coeficiente de concordancia de $\rho_c = 0,959$.

Ahora calculo el promedio donde, además de incluir los momentos individuales, también incluyo los términos correspondientes a las perturbaciones aplicadas por pares, y a este promedio le llamo *ecuación de momentos mixtos*:

$$h_{ijk}^M = \frac{1}{6}(h_i + h_j + h_k + h_{ij} + h_{ik} + h_{jk}). \quad (5.6)$$

En la figura 5.11 se muestra, para la misma red y conjunto de perturbaciones, la comparación entre la distancia de Hamming real h_{ijk} y la predicción usando: (a) la ecuación de Isserlis \hat{h}_{ijk} , (b) el promedio con términos individuales h_{ijk}^I y (c) el promedio con los término mixtos h_{ijk}^M . Del lado derecho están los datos sin normalizar y del izquierdo ya normalizados. Los colores de los puntos corresponden de nuevo a la suma de la concentración de las triadas de perturbaciones, ahora yendo de morado a amarillo para concentraciones de bajas a altas.

De las imágenes de los datos normalizados se concluye que las tres predicciones son igualmente buenas, al menos con estos parámetros. Además la distribución de puntos en las tres gráficas es muy parecida también. Algo diferente entre las tres comparaciones es la posición de los puntos cuando los datos no están normalizados. Para el caso de los promedios, la concentración más baja comienza muy cerca de la recta identidad, mientras que para el caso de Isserlis los puntos de concentración baja están cerca al polinomio de la ecuación (5.4), que corresponde al caso en que todos los momentos son iguales. Es de notar que la ecuación de Isserlis es de tercer orden, mientras que los promedios son ecuaciones lineales y aún así las tres aproximaciones son buenas una vez que los datos han sido normalizados.

La Fig. 5.12 muestra la gráfica del coeficiente ρ_c como función de la conectividad correspondiente a las tres predicciones \hat{h}_{ijk} , h_{ijk}^I , y h_{ijk}^M . Esta imagen revela que los simples promedios también son una buena aproximación a la distancia real h_{ijk} , e incluso son mejores que la predicción de Isserlis en el rango en que ésta no funciona, que es para conectividades bajas. Es interesante además que en el punto crítico las tres funciones fallan para predecir la distancia de Hamming h_{ijk} .

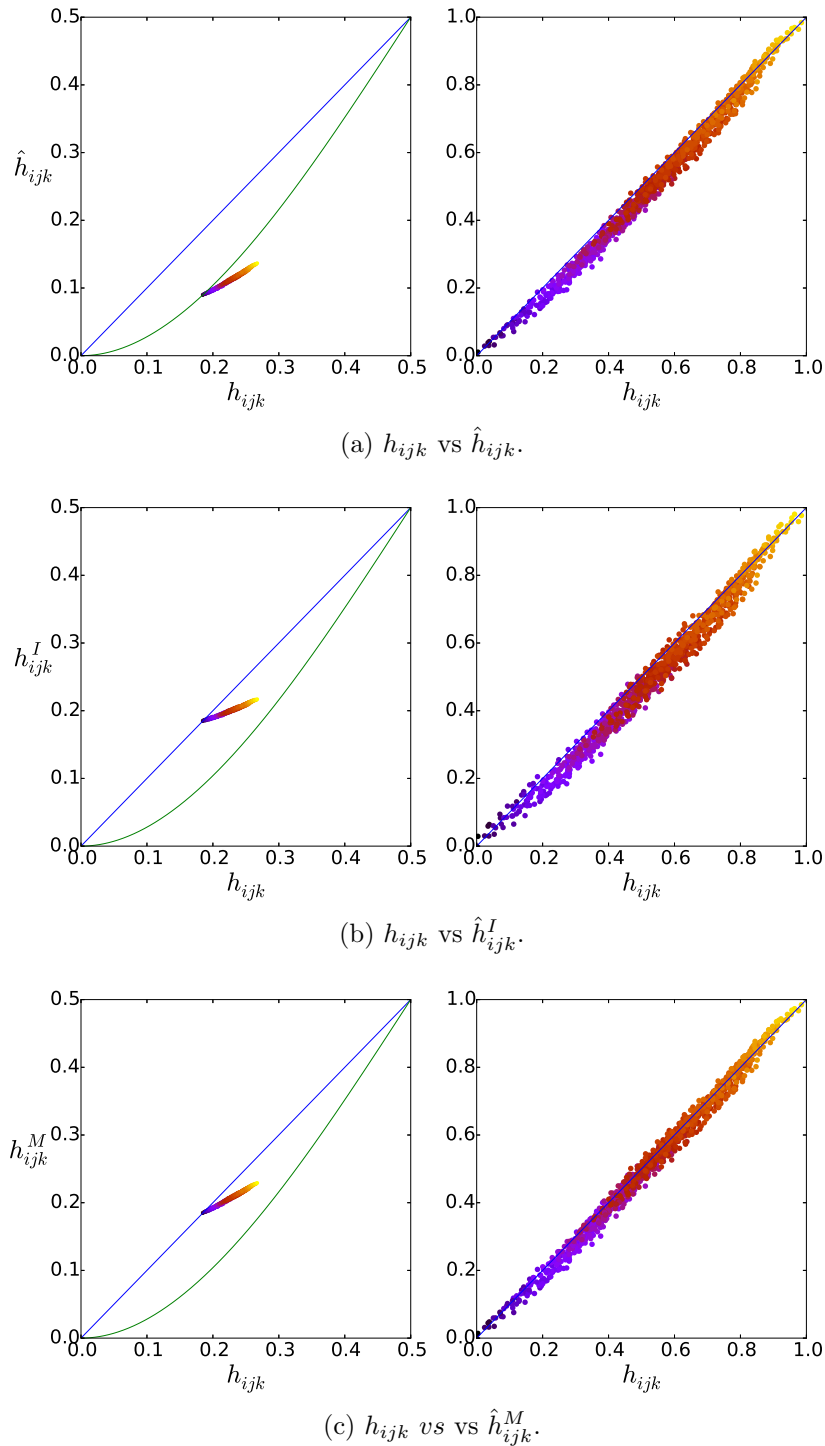


Figura 5.11: Comparación del valor real de la distancia de Hamming h_{ijk} con su predicción utilizando: (a) la ecuación de Isserlis (ecn. (3.12)), (b) el promedio de los términos individuales (ecn. (5.5)) y (c) promedio de momentos mixtos (ecn. (5.6)). Las perturbaciones tienen $\tau = 5, 10$ y 20 y son aplicadas en combinación de concentraciones entre $C = 0,2$ a 2 a una red con $N = 1000$ y $K = 12$. En el panel izquierdo están los datos crudos directamente de la simulación y en el panel derecho los datos escalados. Las tres relaciones proveen una buena predicción. La diferencia notable es que para bajas concentraciones los datos sin escalar están cerca de la línea identidad para el caso de los promedios, mientras que para \hat{h}_{ijk} están cerca del polinomio (5.4).

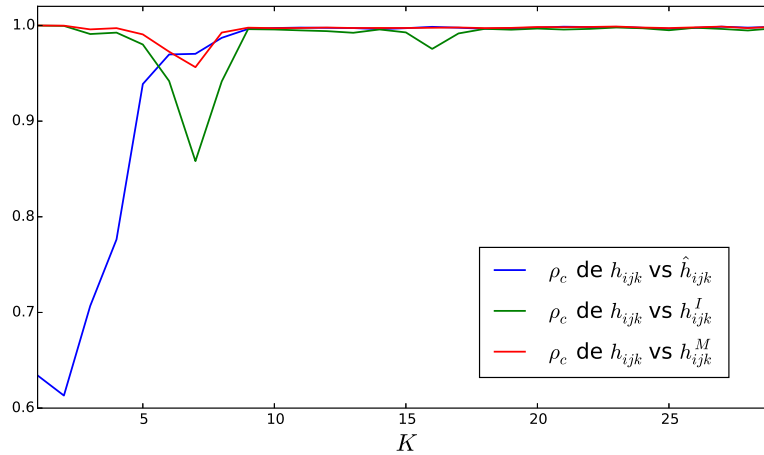


Figura 5.12: Coeficiente de concordancia ρ_c entre los puntos h_{ijk} y la predicción utilizando las ecuaciones de Isserlis (azul), de momentos individuales (verde) y momentos mixtos (rojo), como función de la conectividad K . Para cada conectividad se muestra el promedio sobre 10 realizaciones en las que se aplicaron triadas de perturbaciones tomadas de un conjunto de 10 perturbaciones distintas, todas con $\tau = 5$. Para las redes con $K < 5$ las ecuaciones h^I_{ijk} y h^M_{ijk} presentan una mejor aproximación a h_{ijk} que la predicha por Isserlis. Para redes con $K > 10$ las tres ecuaciones son buenas para la predicción, mientras que alrededor de punto $K = 7$, ninguna de las tres aproximaciones es buena, cercana al punto crítico $K_c = 5$.

5.4. Hipótesis de aditividad en las perturbaciones

El hecho de que los simples promedios h^I_{ijk} y h^M_{ijk} sean buenas aproximaciones a la distancia real h_{ijk} sugiere que la red está teniendo una respuesta aditiva a las perturbaciones. En esta sección estudiamos la validez de esta hipótesis.

5.4.1. Hipótesis de aditividad en ecuación de Isserlis

Definamos h_0 como el valor asintótico de la distancia de Hamming cuando la misma red comienza desde dos condiciones iniciales diferentes, tal y como en la sección 4.2. En este caso no se aplicó ninguna perturbación en pesos. h_0 como función de la conectividad se mostró en la imagen 4.5. Consideremos ahora la distancia de Hamming h_i calculada después de aplicar una perturbación $A_i(\tau_i, C_i)$.

Ahora en la imagen 5.13 se presentan las dos curvas h_0 y h_i , para hacer notar que el efecto de aplicar la perturbación A_i es el de introducir una pequeña fluctuación δ_i entre h_0 y h_i tal que $h_i = h_0 + \delta_i$.

Diferentes perturbaciones A_i , A_j y A_k producirán diferentes desviaciones δ_i , δ_j y δ_k , respectivamente. Análogamente, aplicar dos perturbaciones simultáneamente A_i y A_j producirá la desviación δ_{ij} de h_0 , mientras que aplicar simultáneamente A_i, A_j y A_k producirá la desviación δ_{ijk} .

Ahora supongamos que las desviaciones δ_i , δ_j y δ_k son aditivas en el sentido de que la desviación δ_{ij} obtenida tras aplicar A_i y A_j juntos, puede ser escrita como $\delta_{ij} = \delta_i + \delta_j$, o sea, como la suma de las desviaciones cuando A_i y A_j son aplicadas individualmente. Análogamente aplicar A_i , A_j y A_k simultáneamente produce la desviación $\delta_{ijk} = \delta_i + \delta_j + \delta_k$. Si esta hipótesis de aditividad fuera

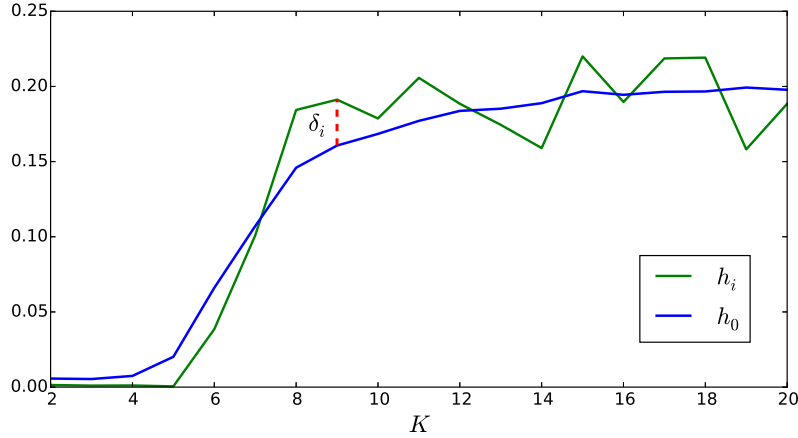


Figura 5.13: Distancia de Hamming promedio h_0 (curva morada) debida a la perturbación en las condiciones iniciales, y la distancia h_i (curva verde) debida a una perturbación en pesos $A_i(10, 1)$. Es de notar que el efecto de la perturbación A_i es desviar la distancia h_0 por una cantidad δ_i .

verdad, la ecuación de Isserlis (5.1) podría ser escrita en términos de las desviaciones δ_i , δ_j y δ_k como:

$$\begin{aligned}
 \hat{h}_{ijk} &= h_i h_j k + h_i h_j k + h_i h_j k - 2h_i h_j h_k \\
 &= (h_0 + \delta_i)(h_0 + \delta_j + \delta_k) + (h_0 + \delta_j)(h_0 + \delta_i + \delta_k) \\
 &\quad + (h_0 + \delta_k)(h_0 + \delta_i + \delta_j) - 2(h_0 + \delta_i)(h_0 + \delta_j)(h_0 + \delta_k) \\
 &= 3h_0^2 - 2h_0^3 + (3h_0 - 2h_0^2)(\delta_i + \delta_j + \delta_k) - 2h_0(\delta_i \delta_j + \delta_j \delta_k + \delta_i \delta_k).
 \end{aligned} \tag{5.7}$$

La ecuación de arriba la podemos escribir como:

$$\hat{h}_{ijk} = P(h_0) \left[1 + \frac{1}{h_0} (\delta_i + \delta_j + \delta_k) \right] + O(\delta^2) \tag{5.8}$$

donde $P(h_0) = 3h_0^2 - 2h_0^3$, es el polinomio de la ecuación (5.4), y $O(\delta^2)$ son los términos de segundo y más alto órdenes en las desviaciones δ_i, δ_j y δ_k .

Por otro lado, de acuerdo a la hipótesis de aditividad, la distancia de Hamming real cuando se aplican tres perturbaciones A_i , A_j y A_k simultáneamente sería:

$$h_{ijk} = h_0 + (\delta_i + \delta_j + \delta_k) \tag{5.9}$$

La importancia de las ecuaciones (5.8) y (5.9) es que ambas son linealmente proporcionales a la suma $\delta_i + \delta_j + \delta_k$ de las desviaciones producidas por las perturbaciones A_i, A_j y A_k individualmente. Si esto fuera verdad, la gráfica \hat{h}_{ijk} contra h_{ijk} sería casi una línea recta (ignorando los términos de orden $O(\delta^2)$) que tendría una pendiente $P(h_0)/h_0$ e interceptaría al polinomio $P(h_0)$.

Para probar esta hipótesis de aditividad generamos δ_i , δ_j y δ_k al azar de una distribución de probabilidad uniforme en un intervalo $(0, a)$ con $a = 0,05$. Introducimos estos valores en las ecuaciones (5.8) y (5.9), y el valor h_0 lo obtenemos de la curva graficada en la Fig. 4.5. El resultado de este análisis se reporta en la figura 5.14, que muestra los datos crudos (sin normalizar) utilizando

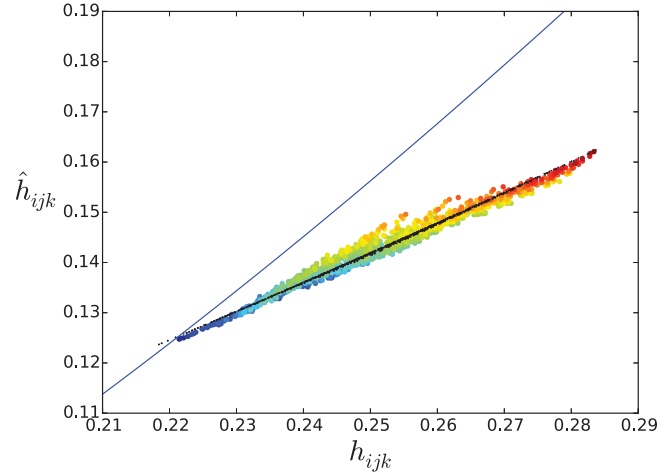


Figura 5.14: Los puntos de colores son los mismos datos que en la figura 5.8c, mientras que los puntos en negro fueron calculados usando las ecuaciones (5.8) y (5.9), que corresponden a la hipótesis de aditividad. El notable traslape de estos dos conjuntos de datos indica que las desviaciones δ_i , δ_j y δ_k son en efecto aditivas. La línea en morado es la gráfica del polinomio $P(h_0)$ definido en la ecuación 5.8.

la predicción de Isserlis (puntos de colores, mismos datos que en 5.8c), contra los datos predichos por la hipótesis de aditividad (puntos negros, ecuaciones (5.8) y (5.9)). El traslape entre estos dos conjuntos de datos claramente muestra que el efecto de diferentes perturbaciones en la dinámica de la red es aditivo, y que la validez de la descomposición en pares de múltiples perturbaciones, implícita en la ecuación de Isserlis, es solamente consecuencia de la aditividad de las desviaciones δ_i individuales producidas por cada perturbación.

5.4.2. Hipótesis de aditividad en pares de perturbaciones

Ahora que la prueba de aditividad ha sido mostrada utilizando tres perturbaciones simultáneas y la ecuación de Isserlis, me remito a mostrar otra prueba, pero ahora utilizando solamente la aditividad en pares de perturbaciones.

Si aplico dos perturbaciones A_i y A_j simultáneamente el efecto de éstas queda determinado por h_{ij} . Este efecto lo puedo comparar con la suma de las perturbaciones individuales $h_i + h_j$, usando la hipótesis de aditividad, como se vio antes:

$$\begin{aligned}
 h_{ij} &= h_0 + \delta_i + \delta_j \\
 &= h_0 + (h_i - h_0) + (h_j - h_0) \\
 &= h_i + h_j - h_0
 \end{aligned} \tag{5.10}$$

En la figura 5.15 comparo h_{ij} y $h_i + h_j$ para tres redes con distinta conectividad a las que se aplicaron dos perturbaciones $A_i(10, C_i)$ y $A_j(20, C_j)$, en conjunto e individualmente. Los paneles izquierdos muestran la gráfica de h_{ij} contra $h_i + h_j$, que si la hipótesis de aditividad fuese válida, debería poder ajustarse con una línea recta con pendiente 1 y con intersección h_0 , como se concluye de la ecuación (5.10). Los paneles centrales muestran mapas de calor para h_{ij} mientras que los paneles de la derecha son mapas de calor para $h_i + h_j$. Para $K = 2$ y $K = 15$ los mapas de calor de h_{ij} y $h_i + h_j$ son prácticamente idénticos (excepto por la escala), y la gráfica de h_{ij} versus $h_i + h_j$

es efectivamente una línea recta con pendiente 1 y desplazada una distancia h_0 , como se esperaría de la hipótesis de aditividad. Sin embargo, para $K = 5$, es decir justo en la transición de fase, es claro que la hipótesis de aditividad no se cumple.

Para analizar la hipótesis de aditividad en los distintos regímenes en los que opera la red, en la figura 5.16 está graficada la pendiente e intersección correspondiente a la extrapolación lineal de los puntos de la comparación h_{ij} contra $h_i + h_j$, que como ya se mencionó, de ser aditivas las desviaciones, la pendiente debe ser uno y la intersección igual a h_0 . Para cada conectividad se hizo un promedio sobre 10 realizaciones distintas, en las que las redes tienen $N = 1000$ nodos y a las que se aplicaron combinaciones de 10 perturbaciones distintas con $\tau = 10$ número de blancos, a concentraciones iguales entre $C = 0,2$ a 2.

La gráfica de la pendiente 5.16a muestra que esta sí está alrededor de uno, excepto para conectividades alrededor de $K = 5$, que corresponde a las redes con dinámica crítica. En el caso del intercepto como función de la conectividad, figura 5.16b, se recupera la curva h_0 de la figura 5.13. Estas dos gráficas en conjunto demuestran que la hipótesis de aditividad se cumple para redes en régimen dinámico ordenado y caótico, pero falla alrededor del punto crítico $K = 5$.

5.4.3. Conclusiones

Por un lado, todas las pruebas que fueron presentadas indican que la descomposición por parejas $A_{ij} = A_i A_j$ de perturbaciones compuestas $A_{ijk\dots l} = A_i A_j A_k \dots A_l$ sugerida por la validez de la fórmula de Isserlis es consecuencia de la aditividad de las desviaciones δ_i , δ_j y δ_k producida por cada perturbación actuando individualmente.

Por otro lado se concluye que las RBUs con perturbaciones actuando sobre sus pesos no son sistemas adecuados para modelar células y el efecto de combinaciones de fármacos, dado que este sistema, pese a que se creía ser altamente no lineal, en realidad resulta tener un comportamiento lineal ante perturbaciones, mientras que la combinación de fármacos sabemos que no es lineal. La linealidad en la respuesta de las RBUs es un resultado importante ya que, como he mencionado antes, las redes con umbrales han sido utilizadas ampliamente para describir regulación genética sin saber que de hecho son sistemas con respuesta lineal, a excepción de en el régimen crítico, mientras que los sistemas genéticos sabemos que son altamente no lineales. Por lo que solamente redes RBU que correspondan al régimen crítico serían convenientes para modelar redes de regulación genética. Sin embargo, en ese régimen no se obtuvo, con las ecuaciones que exploramos en este trabajo, una buena aproximación de las perturbaciones, que era el propósito inicial del trabajo.

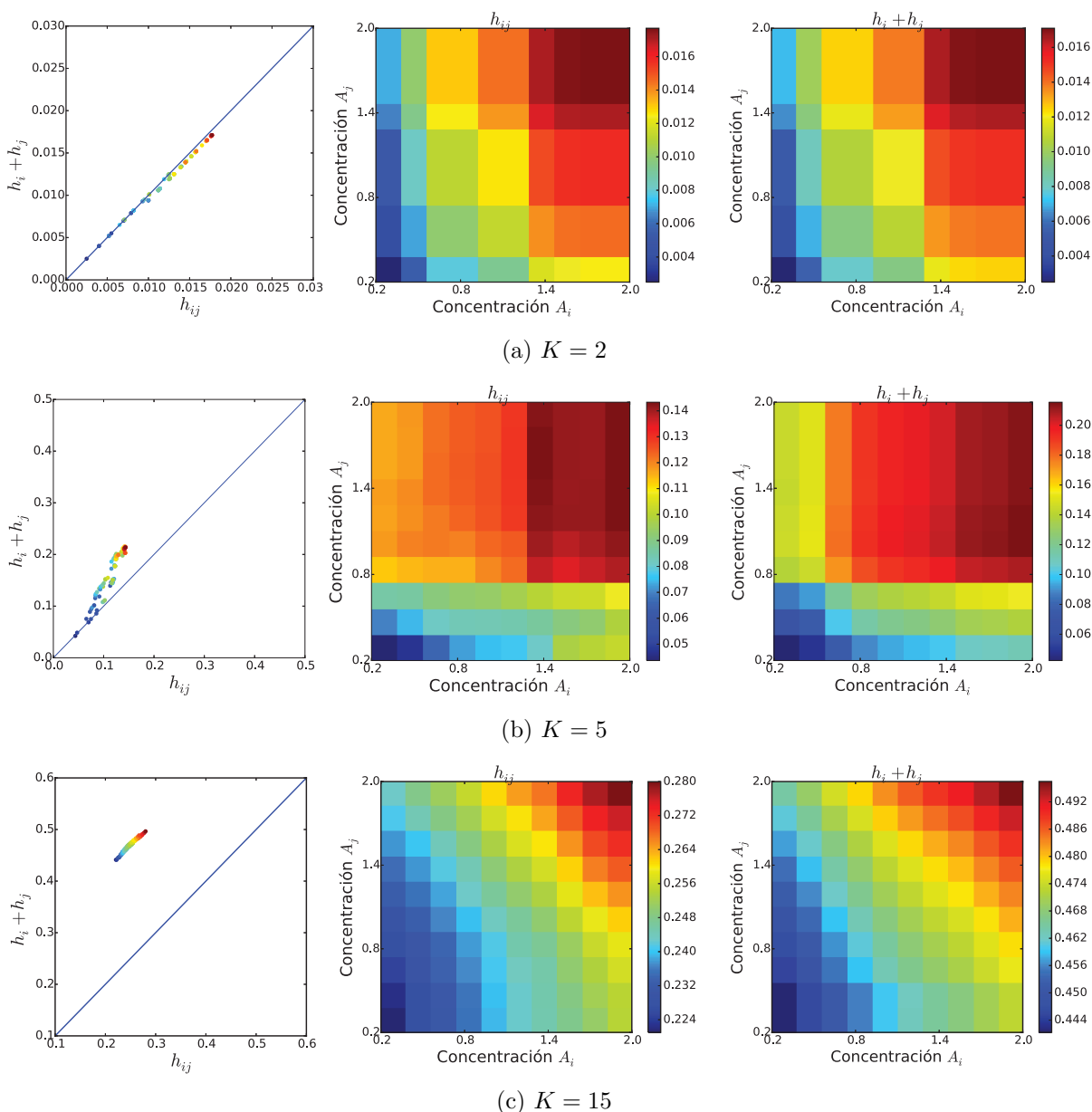


Figura 5.15: Comparación distancia de Hamming debida a perturbaciones simultaneas, h_{ij} , y de la suma de las perturbaciones individuales $h_i + h_j$. La red que se perturba tiene $N = 1000$ nodos y conectividad: (a) $K = 2$, (b) $K = 5$ y (c) $K = 15$. Izquierda: comparación h_{ij} contra $h_i + h_j$. En el centro: mapa de calor h_{ij} . Derecha: mapa de calor de la suma $h_i + h_j$. Para las tres redes las perturbaciones aplicadas son $A_i(10, C_i)$ y $A_j(20, C_j)$. Es de notar que a pesar que los mapas de calor no muestran ser aditivos en el mismo sentido de las combinaciones de fármacos, sí hay una similitud entre h_{ij} y $h_i + h_j$, esto es debido a que las desviaciones δ_i y δ_j son aditivas.

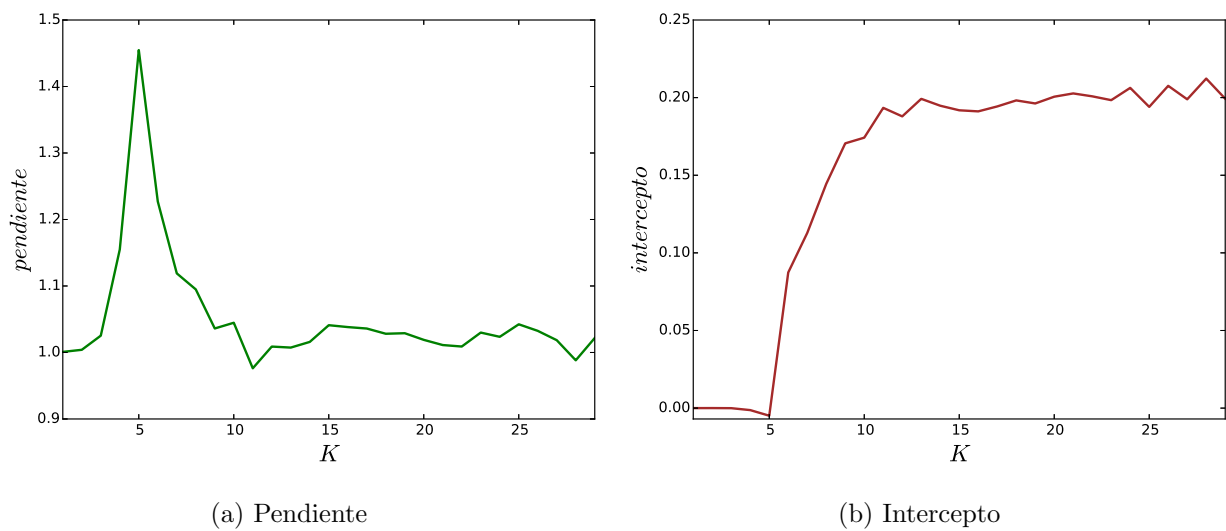


Figura 5.16: Pendiente e intercepto del ajuste lineal de los puntos de la gráfica de comparación de h_{ij} contra $h_i + h_j$, calculado para redes con distintas conectividades. La pendiente para la mayoría de las conectividades tiene valor alrededor de uno, que es lo esperado por la hipótesis de aditividad, a excepción de redes con conectividades cercanas a $K_c = 5$. La gráfica del intercepto recupera la curva h_0 de la figura 5.8. Esto demuestra la validez de la aditividad de las desviaciones para redes en régimen ordenado y caótico, pero no para redes cercanas al punto crítico.

6. Discusión

Las ecuaciones para predecir el efecto de múltiples perturbaciones basadas en las ecuaciones de momentos de Isserlis mostraron ser buenas para predecir el efecto real de las perturbaciones. Sin embargo encontramos que la razón para tal éxito proviene de que el sistema para el que las utilizamos, redes booleanas con umbrales, sorprendentemente tienen una respuesta lineal a perturbaciones que afectan los pesos de las conexiones. Por lo tanto la aseveración de que el efecto de las perturbaciones, medido con la distancia de Hamming entre la red perturbada y la red sin perturbar, se comporten como los momentos de una distribución de probabilidad cuya entropía es máxima, no es necesariamente cierta.

Valdría la pena hacer un análisis de si en los trabajos en que hacen descomposición por pares utilizando el principio de máxima entropía, los de Bialek para las correlaciones en el cerebro y en las parvadas de pájaros, y en el trabajo de Wood et. al. para el efecto de cocteles de medicamentos, sucede lo mismo que en las RBU, es decir que en esos sistemas la hipótesis de aditividad también es válida de alguna manera.

Por otro lado es importante el resultado de las RBU tienen una respuesta aditiva ante perturbaciones que afectan los pesos de sus conexiones, ya que estas redes son comúnmente utilizadas para modelar sistemas de regulación genética, y aunque han tenido éxito describiendo y prediciendo patrones de expresión genética observados experimentalmente, el hecho de que respondan de manera lineal ante perturbaciones no es algo que se haya tomado en cuenta. Sin embargo sí se debería tomar en cuenta, ya que sin duda la regulación genética es un fenómeno no lineal y si las redes con que se está modelando no están en un régimen crítico, se ha estado modelando con un sistema lineal.

Aún queda por analizar la relación de la aditividad de las desviaciones debidas a perturbaciones en pesos y el régimen dinámico de las redes. Se mostró que las redes con conectividad baja, en régimen ordenado, la aditividad se cumple, y al aumentar la conectividad, en el punto crítico la aditividad se pierde. Esto puede sonar plausible si se considera que al aumentar las conexiones, las perturbaciones se propagan de manera no trivial y por lo tanto no de manera lineal. Sin embargo cuando la conectividad es aún más alta, en régimen caótico, la aditividad de nuevo se restituye, haciendo inválido, en principio, el argumento anterior. Una explicación posible sería que en el punto crítico la conectividad funcional sea mayor que en el régimen caótico. A pesar de que la información mutua no apunta hacia ese camino, es una opción para analizar en un trabajo futuro.

A. Análisis semi-analítico de transición de fase

Considero dos configuraciones distintas para el conjunto de nodos de una red con N número de nodos y conectividad K : $\Sigma_t = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ y $\tilde{\Sigma}_t = \{\tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_N\}$. Estas dos configuraciones son dos estados que difieren en pocos estados en ellas.

Para analizar el régimen dinámico de la red se busca saber cuál es la diferencia en un tiempo posterior entre la configuración de la red al haber inicializado la evolución con estas dos configuraciones distintas.

Si x_t es la probabilidad de que en un tiempo t $\tilde{\sigma}_i = \sigma_i$, la probabilidad x_{t+1} dependerá de la regla dinámica 4.1. El estado de $\sigma_i(t+1)$ y de $\tilde{\sigma}_i(t+1)$ depende de la suma respectiva:

$$S_i = \sum_{j=1}^K \omega_{ii_j} \sigma_{i_j}(t) = \sum_{j=1}^m \omega_{ii_j} \sigma_{i_j}(t) + \sum_{j=m+1}^K \omega_{ii_j} \sigma_{i_j}(t), \quad (\text{A.1})$$

$$\tilde{S}_i = \sum_{j=1}^K \omega_{ii_j} \tilde{\sigma}_{i_j}(t) = \sum_{j=1}^m \omega_{ii_j} \sigma_{i_j}(t) + \sum_{j=m+1}^K \omega_{ii_j} \tilde{\sigma}_{i_j}(t) \quad (\text{A.2})$$

donde m es el número de nodos del conjunto de reguladores del nodo i que tienen el mismo estado en un tiempo t en ambas configuraciones.

Entonces la probabilidad de que $\tilde{\sigma}_i(t+1) = \sigma_i(t+1)$ dependerá justo de este número m , ya que si $m = K$ las sumas S_i y \tilde{S}_i son iguales y por lo tanto los estados serán o los dos 1, o ambos 0. Si $m \neq K$ las diversas opciones para que $\tilde{\sigma}_i(t+1) = \sigma_i(t+1)$ son:

- Ambas sumas, S_i y \tilde{S}_i , sean mayor a 0 o ambas menores a 0.
- $S_i = \tilde{S}_i = 0$ y también los estados $\sigma_i(t) = \tilde{\sigma}_i(t)$.
- La suma $S_i = 0$ con $\tilde{S}_i > 0$ y $\sigma_i(t) = 1$ o $\tilde{S}_i < 0$ y $\sigma_i(t) = 0$.
- La suma $\tilde{S}_i = 0$ con $S_i > 0$ y $\tilde{\sigma}_i(t) = 1$ o $S_i < 0$ y $\tilde{\sigma}_i(t) = 0$.

La suma de las probabilidades de que cada uno de estos casos suceda la denoto P_K^m .

La probabilidad x_{t+1} es entonces:

$$x_{t+1} = \sum_{m=0}^K \binom{K}{m} (x_t)^m (1-x_t)^{K-m} P_K^m \quad (\text{A.3})$$

La probabilidad P_K^m la calculo computacionalmente, contando el número de veces que suceden cada uno de los casos mencionados antes. La probabilidad de que los estados sean distintos es $h_t = 1 - x_t$, esa es la que en el caso de la simulación es la distancia de Hamming. La gráfica de h_t como función de la conectividad K es la figura 4.7.

Bibliografía

- [1] Kevin Wood, Satoshi Nishida, Eduardo D. Sontag y Philippe Cluzel. “Mechanism-independent method for predicting response to multidrug combinations in bacteria.” En: *PNAS* 109.30 (2012), págs. 12254-12259.
- [2] Michael A. Fischbach. “Combination therapies for combating antimicrobial resistance.” En: *Current Opinion in Microbiology* 14.5 (2011), págs. 519-523.
- [3] François Nosten y Philippe Brasseur. “Combination therapy for malaria: the way forward?” En: *Drugs* 62.9 (2002), págs. 1315-1329.
- [4] David M. Scollard, L. B. Adams, T. P. Gillis, J. L. Krahenbuhl, R. W. Truman y D. L. Williams. “The continuing challenges of leprosy.” En: *Clinical Microbiology Reviews* 19.2 (2006), págs. 338-381.
- [5] Cesare Saltini. “Chemotherapy and diagnosis of tuberculosis.” En: *Respiratory Medicine* 100.12 (2006), págs. 2085-2097.
- [6] Robert Moellering Jr. “Rationale for use of antimicrobial combinations.” En: *The American Journal of Medicine* 75.2A (1983), págs. 4-8.
- [7] S. Loewe. *The problem of synergism and antagonism of combined drugs*. 1953.
- [8] Remy Chait, Allison Craney y Roy Kishony. “Antibiotic interactions that select against resistance.” En: *Nature* 446 (2007), págs. 668-671.
- [9] Rafael Pena-Miller, David Laehnemann, Gunther Jansen, Ayari Fuentes-Hernandez, Philip Rosenstiel, Hinrich Schulenburg y Robert Beardmore. “When the Most Potent Combination of Antibiotics Selects for the Greatest Bacterial Load: The Smile-Frown Transition.” En: *PLoS Biology* 11.4 (2013), e1001540.
- [10] Joseph Peter Torella, Remy Chait y Roy Kishony. “Optimal drug synergy in Antimicrobial Treatments.” En: *PLoS Computational Biology* 6.6 (2010), págs. 1-9.
- [11] Ayari Fuentes-Hernandez, Jessica Plucain, Fabio Gori, Rafael Pena-Miller, Carlos Reding, Gunther Jansen, Hinrich Schulenburg, Ivana Gudelj y Robert Beardmore. “Using a Sequential Regimen to Eliminate Bacteria at Sublethal Antibiotic Dosages.” En: *PLOS Biology* 13.4 (2015), e1002104.
- [12] Ting-Chao Chou. “Theoretical Basis and Experimental Design and and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies.” En: *Pharmacology Reviews* 53.3 (2007), págs. 621-681.
- [13] Elad Schneidman, Michael J. Berry, Ronen Segev y William Bialek. “Weak pairwise correlations imply strongly correlated network states in a neural population.” En: *Nature* 440.7087 (2006), págs. 1007-1012.

- [14] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale y Aleksandra M. Walczak. “Statistical mechanics for natural flocks of birds.” En: *PNAS* 109.13 (2012), págs. 4786-4791.
- [15] Fred Rieke, David Warland, Rob de Ruyter van Steveninck y William Bialek. *Spikes: exploring the neural code*. MIT Press. 1997.
- [16] William P. Russ, Drew M. Lowery, Prashant Mishra, Michael B. Yaffe y Rama Ranganathan. “Natural-like function in artificial WW domains.” En: *Nature* 437 (2005), págs. 579-583.
- [17] Michael Socolich, Steve W. Lockless, William P. Russ, Heather Lee, Kevin H. Gardner y Rama Ranganathan. “Evolutionary information for specifying a protein fold.” En: *Nature* 437.7058 (2005), págs. 512-518.
- [18] Yasser Roudi, Sheila Nirenberg y Peter E. Latham. “Pairwise maximum entropy models for studying large biological systems: When they can work and when they can’t”. En: *PLoS Computational Biology* 5.5 (2009).
- [19] Edwin T. Jaynes. “Information Theory and Statistical Mechanics.” En: *Physical Review* 106.4 (1957), págs. 620-630.
- [20] Claude E. Shannon. “The Mathematical Theory of Communication.” En: *The Bell System Technical Journal* 27.3 (1948), 379–423.
- [21] Robert M. Mnatsakanov. “Hausdorff moment problem : Reconstruction of probability.” En: *Physical Review Letters* 78.212 (2008), págs. 1869-1877.
- [22] Thierry Mora, Aleksandra M. Walczak, William Bialek y Curtis G. Callan. “Maximum entropy models for antibody diversity.” En: *PNAS* 107.12 (2010), págs. 5405-5410.
- [23] Aonan Tang, David Jackson, Jon Hobbs, Wei Chen, Jodi L. Smith, Hema Patel, Anita Prieto, Dumitru Petrusca, Matthew I. Grivich, Alexander Sher, Pawel Hottowy, Wladyslaw Dabrowski, Alan M. Litke y John M. Beggs. “A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro.” En: *The Journal of Neuroscience* 28.2 (2008), págs. 505-518. ISSN: 1529-2401.
- [24] Jonathon Shlens, Greg D. Field, Jeffrey L. Gauthier, Matthew I. Grivich, Dumitru Petrusca, Alexander Sher, Alan M. Litke y E. J. Chichilnisky. “The structure of multi-neuron firing patterns in primate retina.” En: *The Journal of Neuroscience* 26.32 (2006), págs. 8254-8266.
- [25] S. Kauffman. “Metabolic Stability and Epigenesis in Randomly Constructed Nets.” En: *Journal of Theoretical Biology* 22.3 (1969), págs. 437-467.
- [26] Luis Mendoza, Denis Thieffry y Elena R. Alvarez-Buylla. “Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis.” En: *Bioinformatics* 15.7 (1999), págs. 593-606.
- [27] Carlos Espinosa-Soto, Pablo Padilla-Longoria y Elena R. Alvarez-Buylla. “A gene regulatory network model for cell-fate determination during *Arabidopsis Thaliana* flower development that is robust and recovers experimental gene expression profiles.” En: *Plant Cell* 16.11 (2004), págs. 2923-2939.
- [28] Reka Albert y Hans Othmer. “The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*.” En: *Journal of Theoretical Biology* 223 (2003), págs. 1-18.
- [29] Fangting Li, Tao Long, Ying Lu, Qi Ouyang y Chao Tang. “The yeast cell-cycle network is robustly designed.” En: *PNAS* 101.14 (2004), 4781–4786.

- [30] Maria I. Davidich y Stefan Bornholdt. “Boolean Network Model Predicts Cell Cycle Sequence of Fission Yeast.” En: *PLoS ONE* 3.2 (2008).
- [31] Claudius Gros. *Complex and Adaptive Dynamical Systems A Primer*. Springer H. 2008.
- [32] Bertrand Berche y Christophe Chatelain. “Phase Transitions in Two-Dimensional Random Potts Models.” En: *Order and Disorder and Criticality. Advanced Problems of Phase Transition Theory*. Ed. por Yuriy Holovatch. Vol. 1. Ukraine Academy of Science y Ukraine: World Scientific Publishing Co. Pte. Ltd., 2004. Cap. 4, págs. 147-195.
- [33] Prof. Richard J. Gonsalves. *Course PHY 411-506 Computational Physics 2 Statistical Mechanics and Phase Transitions and the Ising Model*. University at Buffalo and The State University of New York. 2009. URL: <http://www.physics.buffalo.edu/phy411-506-2009/lectures.html>.
- [34] Per Bak y Dimitris Stassinopoulos. “Democratic reinforcement: A principle for brain function.” En: *Physical Review E* 51.5 (1995).
- [35] Werner y Gerhard. “Metastability and criticality and phase transitions in brain and its models.” En: *Biosystems* 90.2 (2007), págs. 496-508.
- [36] Anna Levina, J. Michael Herrmann y Theo Geisel. “Dynamical synapses causing self-organized criticality in neural networks.” En: *Nature Physics* 3 (2007), págs. 857-860.
- [37] Moritz Heimpel. “Critical behaviour and the evolution of fault strength during earthquake cycles.” En: *Nature* 388 (1997), págs. 865-868.
- [38] Ole Peters y J. David Neelin. “Critical phenomena in atmospheric precipitation.” En: *Nature Physics* 2 (2006), págs. 393-396.
- [39] Matti Nykter, Nathan D. Price, Maximino Aldana, Stephen A. Ramsey, Stuart Kauffman, Leroy E. Hood, Olli Yli-harja e Ilya Shmulevich. “Gene expression dynamics in the macrophage exhibit criticality.” En: *PNAS* 105.6 (2007), págs. 1897-1900.
- [40] Christian Torres, Sui Huang y Maximino Aldana. “Criticality Is an Emergent Property of Genetic Networks that Exhibit Evolvability.” En: *PLoS Computational Biology* 8.9 (2012), e1002669.
- [41] Joanna Masel y Meredith V. Trotter. “Robustness and Evolvability.” En: *Trends in Genetics* 26.9 (2010), págs. 406-414.
- [42] Günter P. Wagner, Joshua B. Plotkin, Jeremy A. Draghi y Todd L. Parsons. “Mutational robustness can facilitate adaptation.” En: *Nature* 463 (2010), págs. 353-355.
- [43] Matti Nykter, Nathan D. Price, Antti Larjo, Tommi Aho y Stuart Kauffman. “Critical Networks Exhibit Maximal Information Diversity in Structure-Dynamics Relationships.” En: *Physical Review Letters* 100.5 (2008).
- [44] Alireza Goudarzi, Christof Teuscher, Natali Gulbahce y Thimo Rohlf. “Emergent criticality through adaptive information processing in Boolean networks.” En: *Physical Review Letters* 108.12 (2012).
- [45] Lawrence I kuei Lin. “A Concordance Correlation Coefficient to Evaluate Reproducibility.” En: *Biometrics* 45.1 (1989), págs. 255-268.