



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Física

**Mecanismos de transducción de las condiciones
extracelulares en mediciones/representaciones
cuantitativas**

*Tesis presentada para optar por el título de Doctor de la
Universidad de Buenos Aires en el área Ciencias Físicas*

Autor:

Lic. Alan Matías Givré

Directora:

Dra. Silvina Ponce Dawson

Codirector:

Dr. Alejandro Colman-Lerner

Consejero de estudios:

Hernán Grecco

Lugar de trabajo: FCEN - IFIBA

Fecha de defensa: 06/05/2024

Resumen

Los seres vivos reaccionan y responden a cambios en su entorno. Para generar respuestas específicas la “información” contenida en los estímulos debe ser decodificada. Diversas observaciones han llevado a la conclusión de que, dependiendo de la función específica de cada sistema, la información se puede procesar y representar en dos amplias formas: las señales pueden estar codificadas en la amplitud de la concentración de algún intermediario de la transducción o bien en la frecuencia con la que algunas moléculas “señalizadoras” pasan de un estado activo a otro inactivo. Los mecanismos de señalización suelen involucrar reacciones químicas de las que participa un pequeño número de moléculas. El ruido ineludible en este tipo de procesos limita la capacidad de las vías de señalización. Sin embargo, en muchas ocasiones las respuestas son generadas con gran precisión. ¿De qué manera las células mejoran la relación señal-ruido para lograr esta precisión?

¿Cómo se comparan en este sentido las vías de señalización que codifican por amplitud con las que lo hacen por frecuencia? En esta Tesis se utilizó la teoría de la información de Shannon para intentar responder a estas preguntas.

En esta Tesis, se trabajó en tres aspectos relacionados entre sí que están desarrollados en los Capítulos 2–5. En primer lugar, se trabajó en la descripción mecanística del paso de codificación que va del estímulo externo a un intermediario cuando éste muestra un comportamiento pulsátil estocástico. Para tal fin el trabajo se basó en las numerosas observaciones y modelos en los que el intermediario es el calcio (Ca^{2+}) intracelular. Combinando modelos derivados de observaciones que muestran características comunes a las de sistemas excitables impulsados por ruido se obtuvieron resultados analíticos y numéricos que muestran cómo la información transmitida depende del valor medio del intervalo de tiempo entre pulsos y cómo esta información aumenta dependiendo del peso relativo entre las tasas de disparo (estocástico) y de recuperación de la inhibición luego de un disparo (determinista). En particular, se encontró que en presencia de inhibición, la información transmitida puede exceder en 1bit (y así duplicar) a la del caso puramente estocástico descrito por una estadística Poissoniana. Este resultado fue obtenido tanto estudiando un par de pulsos como en el límite de muchos pulsos.

En segundo lugar, se estudiaron las diferencias y similitudes del procesamiento de información cuando la intensidad de los estímulos extracelulares está codificada en la amplitud, la duración o la frecuencia entre pulsos de la fracción nuclear de un factor de transcripción (FT). Para ello, se analizó la información que puede transmitir el paso asociado a la transcripción usando un modelo sencillo basado en observaciones experimentales del FT Msn2 de la levadura *S. cerevisiae*, y calculando numéricamente la información mutua (MI) entre el FT y la cantidad de ARN mensajero producido, tomando a esta última como un indicador de la cantidad de proteína sintetizada. Explorando exhaustivamente

el espacio de parámetros, se determinó que la MI es máxima en las mismas condiciones para las tres formas de codificación (el umbral de activación del promotor debe ser alto y empujado y la dinámica de todo el proceso debe ser rápida) y que en todos los casos la MI fue de $\sim 1 - 2 \text{ bits}$. Se encontró, sin embargo, que no todas las formas de codificación tienen la misma sensibilidad frente a variaciones de los parámetros que caracterizan a la transcripción. Debido a esta diferente sensibilidad, es posible encontrar promotores relativamente “ciegos” a las entradas moduladas por frecuencia pero que actúan como “switches” ruidosos para las moduladas por duración o amplitud. La situación contraria no puede ser encontrada: los pulsos cortos activan selectivamente un subconjunto de promotores, mientras que pulsos prolongados de FT activan varias vías simultáneamente. Este resultado no solo ayuda a explicar observaciones experimentales en células de levadura sino que puede establecer límites para la “multiplexación” de información (el proceso que ocurre cuando se usa el mismo FT para inducir diferentes respuestas finales dependiendo de su dinámica).

El trabajo contenido en el Capítulo 4 combina el proceso dinámico “río arriba” del FT con el modelo de transcripción explorado en el Capítulo 3. El objetivo acá fue comparar diferencias y similitudes entre modulación por amplitud y por frecuencia cuando ambas etapas de la codificación (desde el estímulo externo hasta la transcripción) son tenidas en cuenta. Para tal fin se modeló el proceso río arriba suponiendo que en el caso de codificación por amplitud, la fracción nuclear del FT es una función de Hill del estímulo externo mientras que, en el caso de codificación por frecuencia, la frecuencia media entre pulsos sucesivos depende exponencialmente de dicho estímulo (como se hizo en el primer trabajo derivado de esta Tesis). Esta dependencia exponencial fue observada en pulsos intracelulares de Ca^{2+} mediados por canales-receptores de IP_3 , y fue también encontrada al analizar los pulsos de localización nuclear del FT, Crz1, en células de levadura. Como mostramos en el Capítulo 5, esta dependencia puede explicarse suponiendo que la dinámica deriva de un sistema excitable donde los pulsos se asocian a las largas excursiones en el espacio de fases que ocurren en estos sistemas cuando se supera un umbral. Utilizando una distribución de soporte compacto para la intensidad del estímulo externo, los resultados del Capítulo 4 muestran que la codificación por amplitud solo funciona de modo óptimo cuando la distribución del estímulo está centrada alrededor de valores para los que la derivada de la función de Hill es máxima. La transmisión por frecuencia, por otro lado, es mucho menos sensible a variaciones en la mediana de la distribución del estímulo externo y puede mejorar al aumentar las intensidades de dicho estímulo (dentro de ciertos límites). Podemos decir entonces que estos dos modos de codificación dotan a las células de una “lente” distinta con la que percibir su entorno. En particular, el hecho de que distintas estrategias de transducción de un mismo estímulo sean óptimas en distintos rangos de intensidades del estímulo podría explicar por qué algunas células, como las de levadura, son capaces de detectar gradientes de un efector externo (feromona en el caso de la levadura) para un rango muy amplio de concentraciones medias distintas.

El Capítulo 5 brinda el marco teórico para el uso de la dependencia expo-

nencial entre el tiempo medio entre pulsos del agente señalizador y la intensidad del estímulo externo cuyo uso en el Capítulo 4 fue principalmente justificado en términos de las numerosas observaciones experimentales existentes. Más específicamente, en este Capítulo usamos un modelo sencillo de la dinámica del Ca^{2+} intracelular promediado sobre toda la célula en el que el factor principal que impulsa dicha dinámica es la liberación de Ca^{2+} desde el retículo endoplasmático a través de canales-receptores de IP_3 . Englobando en un término de ruido aditivo los efectos de las inhomogeneidades espaciales y de los muchos procesos estocásticos que afectan la dinámica del Ca^{2+} intracelular, estudiamos el comportamiento del modelo para parámetros en los que el sistema es excitable. En los sistemas excitables el ruido puede provocar largas excursiones en el espacio de fases que, en el modelo acá analizado, corresponden a pulsos de Ca^{2+} . Usando el modelo con ruido determinamos entonces el tiempo medio entre los pulsos de Ca^{2+} como función de uno de los parámetros del modelo (la fracción β de receptores con IP_3 ligado pasibles de abrirse al ligar Ca^{2+}). Utilizando observaciones de señales espacio-temporales de Ca^{2+} intracelular y un modelo sencillo de los receptores de IP_3 , en este Capítulo pudimos cuantificar el ruido y vincular β con la concentración de IP_3 . Usando estimaciones de la relación entre esta concentración y la intensidad del estímulo externo para distintas células pudimos mostrar que el tiempo medio entre pulsos depende exponencialmente de dicha intensidad para un rango tal que el valor de los tiempos correspondientes varía sobre un intervalo igual o mayor a los observados experimentalmente en distintos tipos celulares. Si bien este resultado fue formalmente derivado para los pulsos intracelulares de Ca^{2+} , la generalidad y sencillez del modelo utilizado permite suponer que es posible aplicar un enfoque similar a otros casos, en particular, al comportamiento pulsátil de FTs, para el que existe abundante evidencia de que, en muchos casos, se rige por una dinámica excitable.

En conclusión, en esta Tesis, exploramos distintos aspectos en que la dinámica temporal en la señalización celular otorga a las células capacidades que en caso contrario no tendrían. Es importante notar que la capacidad de transmitir información es un fenotipo y, por tanto, la selección natural puede actuar sobre ella. En primer lugar, mostramos cómo una dinámica temporal con estadística no Poissoniana (mediante un Feedback Negativo Global) puede aumentar la información transmitida respecto del caso puramente Poissoniano. En segundo lugar, mostramos que la dinámica temporal de un Factor de Transcripción puede activar selectivamente promotores mediante el fenómeno de *multiplexación dinámica*. En tercer lugar, mostramos cómo la distinta forma en que los estímulos externos se codifican en los intermediarios de la señalización dependiendo de su dinámica expande el rango de intensidades del estímulo para los cuales puede transmitirse información de forma fidedigna. Este último resultado derivado inicialmente en base a observaciones experimentales fue finalmente validado a partir de un modelo teórico. Todos estos aspectos de la señalización se manifiestan en fenómenos biológicamente relevantes como se discute en la Tesis.

Palabras clave: *Señalización celular, teoría de la información, pulsos de Ca^{2+} , transcripción*

Transduction mechanisms of extracellular conditions into quantitative measurements/representations.

Living beings react and respond to changes in their environment. The “information” contained in the stimuli must be decoded to generate specific responses. Various observations have led to the conclusion that, depending on the particular function of each system, information can be processed and represented in two broad ways: signals can be encoded in the amplitude of the concentration of some transduction intermediate, or in the frequency with which some “signaling” molecules switch from an active to an inactive state. Signaling mechanisms frequently use chemical reactions involving a small number of molecules. The unavoidable noise in this type of process limits the capacity of the signaling pathways. However, on many occasions, the answers are generated with great precision. How do cells improve the signal-to-noise ratio to achieve this precision?

How do signaling pathways that use amplitude encoding compare with those that encode by frequency? In this Thesis Shannon’s information theory is used to try to answer these questions.

In this Thesis, we worked on three interrelated aspects that are developed in Chapters 2–5. First, we worked on the mechanistic description of the coding step that goes from the external stimulus to an intermediary that shows a stochastic pulsatile behavior. To this end, the work was based on numerous observations and models in which the intermediary is intracellular calcium (Ca^{2+}). By combining models derived from observations that show characteristics common to those of noise-driven excitable systems, analytical and numerical results were obtained that show how the information transmitted depends on the average value of the time interval between pulses and how this information increases depending on the relative weight between the rates of firing (stochastic) and the recovery from inhibition after firing (deterministic). In particular, it was found that in the presence of inhibition, the information transmitted can exceed by 1 bit (and thus double) that of the purely stochastic (Poissonian) case. This result was obtained both by studying a pair of pulses and in the limit of many pulses.

Second, we studied the differences and similarities of information processing when the intensity of extracellular stimuli is encoded in the amplitude, duration, or frequency between pulses of the nuclear fraction of a transcription factor (TF). To do this, we analyzed the information that the transcription-associated step can convey using a simple model based on experimental observations of the TF, Msn2 of the yeast *S. cerevisiae*, and numerically calculating the mutual information (MI) between the TF and the amount of messenger RNA produced, taking the latter as an indicator of the amount of protein synthesized. Exploring exhaustively the parameter space, it was determined that the MI is maximal under the same conditions for the three forms of coding (the promoter activation threshold must be high and steep and the dynamics of the whole process must be fast) and that in all cases $\text{MI} \sim 1 - 2 \text{ bits}$. It was found, however, that not

all coding forms have the same sensitivity to variations in the parameters that characterize transcription. Due to this different sensitivity, it is possible to find promoters that are relatively “blind” to frequency-modulated inputs but act as noisy switches for duration or amplitude-modulated ones. The reverse situation cannot be found: short pulses selectively activate a subset of promoters, while prolonged FT pulses activate several pathways simultaneously. This result not only helps to explain experimental observations in yeast cells but may set limits for information “multiplexing” (the process that occurs when the same FT is used to induce different final responses depending on its dynamics).

The work contained in Chapter 4 combines the process upstream the nuclear localization of the FT with the transcription model explored in Chapter 3. The objective here was to compare differences and similarities between modulation by amplitude and by frequency when both stages of the coding (from the external stimulus to the transcription) are taken into account. For this purpose, the upstream process was modeled assuming that in the case of amplitude coding, the nuclear fraction of the FT is a Hill function of the external stimulus while, in the case of frequency coding, the average frequency between successive pulses depends exponentially of said stimulus (as was done in the first work derived from this Thesis). This exponential dependence was observed for intracellular Ca^{2+} pulses. We also found it when analyzing the pulses of the FT, *Crz1*, in yeast cells. As shown in Chapter 5, this dependence can be explained assuming that the underlying dynamics is excitable and that the pulses are associated with the long excursions in phase space that occur in these systems when a threshold is surpassed. Using a compact support distribution for the intensity of the external stimulus, the results of Chapter 4 show that amplitude coding only works optimally when the stimulus distribution is centered around values for which the derivative of the Hill function is maximum. Transmission by frequency, on the other hand, is much less sensitive to variations in the median distribution of the external stimulus and improves with increasing intensities of said stimulus (within certain limits). We can then say that these two encoding modes provide cells with a different “lens” through which to perceive their environment. In particular, the fact that different transduction strategies for the same stimulus are optimal in different ranges of stimulus intensities could explain why some cells, such as yeast cells, are capable of detecting gradients of an external effector (pheromone in the case of yeast) for a very wide range of different mean concentrations.

Chapter 5 provides the theoretical framework for the use of the exponential dependence between the mean interpulse time of the signaling agent and the intensity of the external stimulus which use in Chapter 4 was mainly justified in terms of the numerous existing experimental observations. More specifically, in this Chapter we use a simple model of the intracellular Ca^{2+} dynamics averaged over the entire cell in which the main factor driving the dynamics is the release of Ca^{2+} from the endoplasmic reticulum through IP_3 -receptor-channels. Encompassing in a single additive noise term the effects of the spatial inhomogeneities and of the many stochastic processes that affect the dynamics of intracellular Ca^{2+} , we study the behavior of the model for parameters for which the system

is excitable. In excitable systems, noise can cause long excursions in phase space that, in the model analyzed here, correspond to pulses of Ca^{2+} . Using the noisy model we then determined the average time between Ca^{2+} pulses as a function of one of the model parameters (the fraction, β , of receptors with IP_3 bound ready to become open upon Ca^{2+} binding). Using spatio-temporal observations of intracellular Ca^{2+} and a simple kinetic model of the IP_3 receptors, in this Chapter we were able to quantify the noise and relate β to the IP_3 concentration. Using estimates of the relationship between this concentration and the intensity of the external stimulus for different cells we showed that the average time between pulses depends exponentially on said intensity for a range such that the value of the corresponding times varies over an interval equal to or greater than those observed experimentally in different cell types. Although this result was formally derived for intracellular Ca^{2+} pulses, the generality and simplicity of the model allows us to assume that it is possible to apply a similar approach to other cases, in particular, to the pulsatile behavior of FTs, for which there is abundant evidence that, in many cases, it is governed by excitable dynamics.

In conclusion, in this Thesis, we explored different aspects in which temporal dynamics in cell signaling give cells capabilities that they would not otherwise have. It is important to note that the ability to transmit information is a phenotype and, therefore, natural selection can act on it. First, we showed how a dynamics with non-Poissonian statistics (through Global Negative Feedback) can increase the transmitted information with respect to the purely Poissonian case. Secondly, we showed that the temporal dynamics of a Transcription Factor can selectively activate promoters through the phenomenon of dynamic multiplexing. Third, we showed how the different way in which external stimuli are encoded in signaling intermediaries depending on their dynamics expands the range of stimulus intensities for which information can be transmitted reliably. This last result initially derived based on experimental observations was finally validated based on a theoretical model. All these aspects of signaling are manifested in biologically relevant phenomena that are discussed in the Thesis.

Keywords: *Cell signaling, information theory, Ca^{2+} pulses, transcription*

Índice general

1. Introducción	12
1.1. Descripción general	12
1.2. Factores de transcripción y regulación de la expresión génica	16
1.3. Vías de señalización que codifican por amplitud	19
1.4. Codificación por frecuencia	21
1.4.1. Señales intracelulares de calcio.	21
1.4.2. Comportamiento pulsátil de factores de transcripción.	28
1.4.3. Excitabilidad y pulsos	30
1.5. Multiplexación: un mismo FT con distintas dinámicas para responder a distintos estímulos.	33
1.6. Un mismo estímulo codificado en amplitud y frecuencia	37
1.7. Teoría de la Información y su aplicación a problemas biológicos	38
1.7.1. La teoría de Shannon	39
1.7.2. Teoría de la información en sistemas biológicos: descripción general	40
1.7.3. Teoría de la información en sistemas biológicos: descripción detallada de un ejemplo.	42
1.7.4. Teoría de la información aplicada al análisis de la multiplexación en levaduras.	45
1.8. Organización de la Tesis y anticipo de los resultados principales.	46
2. Contenido de información en secuencias de pulsos estocásticos de mensajeros intracelulares.	50
2.1. Métodos	50
2.1.1. El modelo	50
2.1.2. Cálculos	51
2.1.3. Simulaciones Numéricas	52
2.2. Resultados	52
2.2.1. Información contenida en el tiempo entre pulsos.	52
2.2.2. Información contenida en el número de pulsos, N , que ocurren durante un intervalo.	53
2.3. Discusión y conclusiones del Capítulo.	54

3. La modulación de la dinámica de los Factores de Transcripción permite una transmisión versátil de la información.	56
3.1. Métodos	57
3.1.1. Modelo, entrada y salida	57
3.1.2. Simulaciones numéricas	58
3.1.3. Cómputo de la Información Mutua	59
3.1.4. Muestreo del espacio de parámetros, optimización y comparaciones.	60
3.2. Resultados	61
3.2.1. Información mutua máxima y parámetros del modelo de transcripción.	61
3.2.2. Dinámica del modelo y parámetros que maximizan MI.	65
3.2.3. Pares de parámetros e información mutua	65
3.2.4. Un factor de transcripción, dos genes	67
3.3. Discusión y conclusiones del Capítulo.	72
4. Las codificaciones por amplitud y frecuencia dan a las células una lente diferente para detectar el entorno.	77
4.1. Métodos	78
4.2. Resultados	80
4.3. Discusión y conclusiones del Capítulo.	82
5. Procesamiento de información celular mediante codificación de frecuencia y excitabilidad.	85
5.1. Métodos	86
5.2. Resultados	86
5.2.1. Análisis del modelo: estímulos externos y el parámetro β	86
5.2.2. Análisis del modelo: soluciones de equilibrio y nulclinas.	86
5.2.3. Descripción del modelo en términos de variables lentas y rápidas	88
5.2.4. Probabilidad por unidad de tiempo y período promedio entre pulsos sucesivos.	89
5.2.5. Tiempo medio entre pulsos e intensidad del estímulo externo	91
5.3. Discusión y conclusiones del Capítulo.	93
6. Discusión y conclusiones generales de la Tesis.	96
A. Cálculos suplementarios del Capítulo 2	100
A.1. Propiedades del modelo en los límites de Poisson y GNF	100
A.1.1. Relación entre la tasa de “disparo”, λ , y la concentración de ligando externo, C	101
A.2. Cuantificación de la información contenida en un pulso.	104
A.2.1. Detalles de Cálculos	104
A.2.2. Información contenida en $N \gg 1$ pulsos.	108

- A.2.3. Resultado final de la información contenida en el número de pulsos, N , que ocurren durante un intervalo. 111
- A.3. Más allá del Teorema Central del Límite. 111
- B. Apéndice al Capítulo 3 114**
- B.1. Modelo de transcripción: descripción mecanística 114
- B.2. Dinámica temporal en la modulación por amplitud y frecuencia . 118
- B.2.1. Simulación del modelo sin límite a la tasa de producción de ARNm 120

Agradecimientos

Esta tesis tiene 7 años en proceso, pero sus raíces tienen décadas atrás. Ya de por sí, son 13 años desde que empecé la carrera de Física. Pero me gustaría remontarme mucho antes.

Por ejemplo, podría remontarme a mis 12 años, fecha en que mi prima Melina Furman me regaló los libros de Fotosíntesis y El Código Genético de Asimov. Este fue mi primer contacto con la ciencia y con la biología en particular. Asimov me abrió los ojos de muchas maneras, y es su narrativa acerca de la ciencia que yo tomé, y mis estudios de Física y Sociología estuvieron inspiradas en su saga de la Fundación. Cuando fui al campamento científico que Melina organizaba a los 16 años, arraigó en mí el espíritu de que la vista estuviera atenta, la mente estuviera abierta, y la discusión estuviera activa. Ese fue el momento en terminé de consolidar la decisión de seguir física. Y muchos de mis mejores amigos provienen de ese campamento. Mi prima fue siempre mi inspiración de qué significa ser un científico. Entonces, mi primera gratitud es hacia mi prima Melina, y también a Asimov.

Pero también podría remontarme a mis 10 años, fecha en que comencé a asistir a las Olimpiadas de Matemáticas. Al primer examen no fui, por desidia de mi directora. Fue el impulso de mi padre, quien habló hasta con el director de la OMA para que me dejaran asistir a la siguiente instancia. Así como esa, la fuerza de mi padre en muchas instancias me abrió camino. Y así también fue quien me introdujo al valor, el disfrute y el juego del pensamiento formal y lógico, y a los juegos de mesa. Y en las Olimpiadas Ñandú pude abrirme al pensamiento matemático en su forma lúdica. Quiero agradecer así a la OMA. Y es de este juego lúdico del pensamiento matemático que me llevó a la física. Así, quiero agradecer a mi papá.

Quien me acompañó en el día a día de las frustraciones de esta tesis fue mi novia Adriana. Ella pudo ver lo que significaba la ciencia y sus frustraciones de cerca. Y ahora lo está viviendo en carne propia. Ella pudo ver que hay una esperanza en la Física, una calma y ecuanimidad en ver el mundo moviéndose por sí misma. Y pudo compartir conmigo mi visión del mundo biológico, poder ver los “cositos” moviéndose por sí mismos y jugando entre ellos.

Sobrellevar esta tesis requirió calma y paz emocional, y una y otra vez los consejos y sabiduría de mi mamá me ayudaron a que no perdiera por completo la cabeza. Así que esta es mi gratitud hacia ella. En mi hermana encontré mi mejor amiga, con quien siempre pude compartir con honestidad, pesares y

celebraciones, y me acompañó en todo momento.

A mi directora Silvina, realmente le quiero agradecer, la paciencia y la sabiduría. Frecuentemente, cuando yo no sabía cómo seguir, habiendo reflexionado sobre el problema por semanas, ella le pegaba un vistazo y encontraba la solución en segundos. A mi codirector, Alejandro Colman Lerner, le quiero agradecer recordarme que hay un mundo real más allá de las abstracciones de la física, no perderse en el mundo de la matemática.

Quiero agradecer a muchos de mis amigos. Para empezar, a la Vegana, su sabiduría y amor me acompañó en los momentos más difíciles. Años antes de que yo tuviera aceptación incondicional de amigos, ella me enseñó lo que eso significaba. Quiero agradecer a Diego por siempre estimular mi imaginación y nuestras ideas. Quiero agradecer a Ludmila por haber estado siempre ahí, en su transparencia y honestidad. Quiero agradecer a Lucas, por su practicidad y forma de bajarme a la tierra. Y a Gregorio por su expresividad y acompañamiento. Y a Felicitas por estar en mi momento más difícil. Y estos son sólo algunos amigos que me acompañaron estos años.

Capítulo 1

Introducción

1.1. Descripción general

La vida transcurre en contextos complejos. Para sobrevivir y reproducirse, los seres vivos necesitan protegerse de depredadores, nutrirse de la materia que los conforma (aminoácidos, lípidos, etc) y del agua que permite mantener sus componentes en solución coloidal. Necesitan energía para permanecer fuera del equilibrio. Para esto, a veces deben desplazarse o modificar su forma, o actuar de modo diferente dependiendo del momento del día o de la estación. Esto sucede para todo tipo de organismos, incluyendo bacterias, levaduras, plantas, y animales.

Desde el punto de vista de la física nos encontramos con un sistema en extremo complejo que involucra una pluralidad de moléculas distintas que interactúan entre sí. En el medio intracelular, cada tipo de molécula está en una cantidad limitada, de tal manera que las fluctuaciones y el ruido importan cuando se estudian procesos en esta escala. La misma célula no se va a comportar de la misma manera si está creciendo o si se está reproduciendo. No es fácil abordar la especificidad de los procesos biológicos, que son extremadamente diversos entre sí, y poder decir algo pasible de ser generalizado. Por otro lado, como los seres vivos están fuera del equilibrio, es necesario pensar en ellos de forma dinámica. El aspecto dinámico o temporal de los seres vivos enriquece su comportamiento y les permite responder de formas inaccesibles para un sistema en equilibrio, o incluso, un sistema estacionario.

Si nos proponemos buscar principios que sean compartidos por distintos sistemas, como plantea el investigador de la Universidad de Princeton, William Bialek, cuyo trabajo fue una gran inspiración para esta tesis [13], nos encontramos con aspectos de la termodinámica, como la criticalidad [75], o de la dinámica no lineal, como las bifurcaciones [67], o bien con las propiedades de la materia condensada blanda [56]. Si algo parece repetirse, es que la vida es sensible (responde a pequeñas diferencias de su entorno), diversa (sus fenotipos macroscópicos varían ampliamente) y flexible (el mismo sistema varía sus

aspectos macroscópicos temporalmente). A diferencia de una piedra, o un gas, los cuales tienen cierta homogeneidad espacial y temporal, los sistemas vivos se nos aparecen con una riqueza fractal, tanto espacial como temporalmente. Encontrar que la vida es sensible, flexible y diversa, no es un aspecto obvio o menor. Estos mismos aspectos que la hacen tan compleja nos dicen algo acerca de qué es la vida y cómo opera, ya que selecciona y limita la forma en que puede hacerlo.

Para que los seres vivos sean sensibles a su entorno, necesitan comunicarse con él, intercambiar información. En las células, la información proveniente de las condiciones externas e internas se combina con la del genoma para determinar las acciones futuras. Entre éstas se encuentran la muerte y la reproducción, es decir, los procesos básicos subyacentes a la vida. El mal funcionamiento de los sistemas de transmisión de información es causa de patologías como las enfermedades auto-inmunes y el cáncer. En particular, en el caso del cáncer la falla en los mecanismos de señalización resulta clara dado que se trata de una enfermedad que es consecuencia de una reproducción celular sin coordinación con el resto del organismo.

Existe una cantidad matemáticamente bien definida para describir cuánta información un sistema recibe de su entorno, la *entropía de Shannon* [86], y más específicamente la *información mutua* entre los sistemas que intentan comunicarse. La teoría de la información nos muestra algo sorprendente: el ruido y las fluctuaciones, que superficialmente estarían haciendo de la vida algo “imperfecto”, son los componentes que hacen que el concepto de información tenga sentido. La teoría de la información requiere una distribución de probabilidad del estímulo y de la respuesta [86]. De acuerdo a esta teoría el comportamiento de una célula en un contexto concreto debe ser visto como un caso particular del comportamiento que diversas células pueden desplegar en la pluralidad de entornos frente a los cuales se pueden encontrar.

Otro aspecto crucial de la vida es que las estructuras que constituyen la esencia del funcionamiento de las células están altamente conservadas entre todas las especies. De esta manera, los mismos caminos o “pathways” se repiten en organismos tan diversos como los animales y los hongos, o incluso las plantas. Siguen siendo los mismos elementos que hace centenares de millones de años. Esto permite utilizar modelos biológicos extremadamente sencillos, como la levadura *Saccharomyces cerevisiae* ya que, en ciertos aspectos, se comporta en forma similar a las células de seres vivos mucho más complejos.

A nivel celular, una de las formas de reaccionar y responder a estímulos es mediante la producción de proteínas que van a operar la transformación de la célula. Por ejemplo, si una célula de levadura de un cierto tipo encuentra feromonas de las que secreta otro tipo de levadura con la que puede aparearse, dependiendo de las otras señales que reciba en ese momento puede pasar a producir las proteínas y lípidos que le permitan generar una protrusión (“shmoo”) hacia la potencial pareja. La producción de estas proteínas requiere de una vía de señalización considerablemente compleja, como se ilustra en la Fig. 1.1.

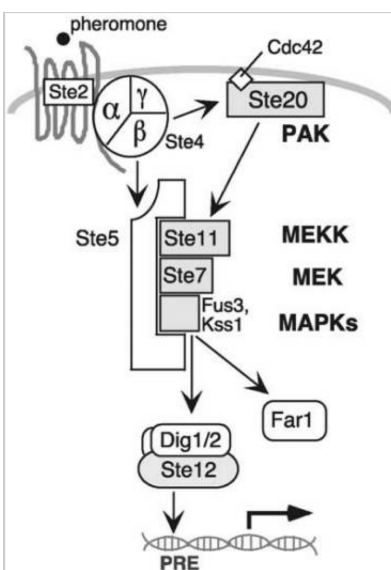


Figura 1.1: Esquema de algunos de los elementos que intervienen en la vía de señalización de respuesta a feromona en células de *S. cerevisiae*. La feromona se une a un receptor en la membrana plasmática (Ste2) del tipo GPCR, que lleva a que se active una cascada de proteínas kinasas (Ste11, Ste7, Fus3 y Kss1), que activan al factor de transcripción (FT) Ste12. Ste12 causa modificaciones en la expresión génica que se manifiestan, entre otras cosas, en cambios en la forma celular. Figura tomada de [4].

La Fig. 1.1 muestra que la célula recibe la información externa a través de receptores localizados en su membrana a los que se une la feromona induciendo cambios en el citosol. En el caso de células eucariotas (como en la figura), mediante una *cascada de señalización* que va activando distintos mensajeros intermedios, la información es eventualmente transmitida al núcleo. En particular, esto último algunas veces involucra que *Factores de Transcripción* (FTs) ingresen al núcleo para interactuar con el ADN en regiones *río arriba* del sitio del inicio de la transcripción en el promotor de los genes blanco. Así, el FT causa que se inicie la transcripción del ARNm que luego se traducirá en la producción de alguna de las proteínas involucradas en la respuesta al estímulo. Este proceso se puede cuantificar a través de la información mutua entre una “entrada” o “input” y una “salida” o “output”. Estudios previos usaron esta estrategia para estimar, bajo ciertas aproximaciones, la relación que maximiza la información mutua entre la concentración del FT y la expresión del gen correspondiente [105]. En esta Tesis usamos esta estrategia para comparar modos de codificar la información externa y generar respuestas como se describe más adelante.

Cada uno de los pasos que van desde el estímulo externo hasta la producción de proteínas tiene una cierta eficiencia, y la información que se pierde por baja eficiencia no puede recuperarse *a posteriori* como consecuencia de la *desigualdad del procesamiento de datos* (o *data processing inequality* en inglés) [8].

Se ha encontrado que la información procesada por algunos sistemas biológicos puede llegar a ser extremadamente precisa, hasta los límites del ruido físico [13]. El estudio de las interneuronas del ojo compuesto de la mosca de Laughlin [59] es pionero en este sentido. Este autor encontró que la función que relacionaba la respuesta con la intensidad del estímulo externo era aproximadamente igual a la distribución de probabilidad acumulada de los contrastes en el paisaje donde se movía la mosca. Esto brindaba una codificación óptima ya que permitía que todos los estados de respuesta posibles fueran “usados” con igual probabilidad por la mosca en su entorno natural. Este trabajo pionero dio lugar a numerosos análisis similares en otros sistemas que permitieron encontrar resultados análogos. Casos de optimalidad en la transmisión de información se pueden ver también, por ejemplo, en contextos de quimiotaxis [9]. La alta precisión es necesaria también en el desarrollo embrionario (en la respuesta a gradientes de morfógenos) [26, 43, 104, 105], ya que es necesario traducir correctamente la información genética en una respuesta fenotípica para la viabilidad del embrión. Otro ejemplo involucra la percepción que los animales tienen de su entorno, mediante los órganos receptores, y las delicadas vías de señalización eléctricas y químicas del sistema nervioso [12]. Cuando la respuesta involucra cambios en la expresión génica mediados por FTs, bajo ciertas aproximaciones es posible determinar, teóricamente, la distribución de concentraciones de FT que maximiza la información mutua entre el FT y la proteína cuya producción regula. En el caso de algunos pasos en el desarrollo de los embriones de la mosca de la fruta, la comparación de la distribución teórica con las observaciones experimentales permitieron inferir la “optimalidad” del proceso.

Como en el ejemplo de la levadura y la feromona de la Fig. 1.1, en las células los cambios en el entorno, que se manifiestan en cambios en la concentración de sustancias efectoras externas, se traducen en cambios en la concentración de sustancias dentro de la célula, incluyendo, eventualmente, cambios en la fracción nuclear de FTs y, consecuentemente, en la expresión génica. Acá hay dos aspectos vinculados al cambio en el entorno y la generación de respuestas que la célula debe “atender”. Por un lado, la identificación de la sustancia efectora cuya concentración cambió. Por el otro, la capacidad de distinguir entre distintas concentraciones de una misma sustancia, es decir, entre intensidades del estímulo externo. Esto último es particularmente importante para la detección de gradientes como sucede en quimiotaxis o durante distintos procesos de crecimiento (como en el ejemplo de la levadura y la feromona). Las células típicamente usan dos modos de “codificar” las distintas intensidades del estímulo externo: por *Amplitud* o por *Frecuencia* [61]. En el primer caso, intensidades crecientes del estímulo generan concentraciones crecientes de los intermediarios de la señalización, incluyendo, eventualmente [110, 112], los FTs. En otros casos, aparecen comportamientos pulsátiles en algunos de los intermediarios y/o los FTs en los que la frecuencia entre pulsos aumenta al aumentar la intensidad del estímulo [17, 46]. Algunas respuestas incluso usan información redundante codificada en amplitud [21] o en frecuencia [19]. En esta Tesis utilizamos herramientas de la Teoría de la Información para analizar las diferencias y similitudes entre ambos tipos de codificación y en qué casos uno u otro modo de representar

los cambios en el entorno puede ser más adecuado.

Es particularmente importante dilucidar la relevancia de la señalización codificada en frecuencia ya que el mantenimiento de un sistema que tenga una regulación temporal implica un mayor gasto de energía y de recursos materiales. Por ejemplo, en el caso de las neuronas y señales de Ca^{2+} es necesario mantener un gradiente de los iones a través de la membrana manteniéndolos fuera del equilibrio lo que requiere una gran inversión de energía. Que las células utilicen la codificación por frecuencia en situaciones muy distintas significa que debe haber otros beneficios.

La posibilidad de utilizar información modulada por la dinámica temporal, que puede estar codificada en la frecuencia o en la duración de los pulsos, enriquece significativamente la comunicación celular [1, 17, 48, 61]. En particular, permite el uso de la *multiplexación dinámica*. Esto es, el uso de distintas dinámicas de la misma especie (ya sea un FT o un ion) para señalar distintos procesos, permitiendo la codificación en amplitud, duración o frecuencia como respuesta a distintos tipos de estímulos. Mostramos más adelante resultados experimentales de otros grupos ilustrando este comportamiento.

De esta manera, el objetivo general de esta tesis es exponer cómo el acceso a una dinámica temporal le permite a una célula mayor sensibilidad y flexibilidad ante su entorno, así como expresar con claridad el rol que tienen el ruido y el azar en la respuesta a estímulos. De esta manera, se exponen algunas de las muchas ventajas que un comportamiento dinámico -lo cual requiere estar fuera del equilibrio- provee informacionalmente a las células.

1.2. Factores de transcripción y regulación de la expresión génica

La información necesaria para producir proteínas está contenida en los genes. En esta tesis cuando hablamos de genes nos limitaremos a aquellas secuencias de nucleótidos en las moléculas de ADN (ácido desoxirribonucleico) de las células que pueden ser transcritas a ARN (ácido ribonucleico) mensajero (ARNm) y luego traducidas a proteínas, y a sus regiones regulatorias adyacentes. La *transcripción* constituye el primer paso hacia la producción de proteínas y consiste en copiar la información genética en otro tipo de ácido nucleico: el ARN. Posteriormente ocurre la *traducción* que permite pasar de la información contenida en el ARN a proteínas. La Fig. 1.2 ilustra esquemáticamente estos pasos que constituyen el llamado “dogma central de la biología molecular” [2].

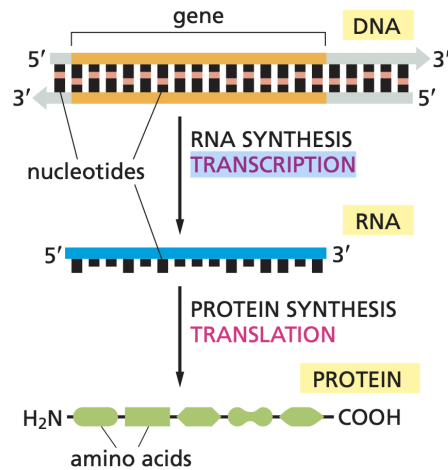


Figura 1.2: Esquema mostrando el flujo de la información genética desde el ADN hasta las proteínas. Figura tomada de [2]

Para que la transcripción pueda ocurrir, el ADN debe estar “disponible para la lectura”. Para ello, sus dos hebras deben poder separarse para que la información genética que contienen sea leída. En este proceso intervienen las ARN polimerasas que utilizan una de las hebras de la doble hélice como molde y, siguiendo las reglas de apareamiento de Watson y Crick, catalizan la polimerización de ARNm. Las polimerasas se mueven paso a paso a lo largo del ADN desenrollando la hélice y exponiendo nuevas regiones para su lectura. El inicio de la transcripción es particularmente crítico ya que es el que determina qué ARNm y, por lo tanto, qué proteínas se producirán. Para comenzar la transcripción, la ARN polimerasa debe reconocer dónde comienza el gen a ser leído para unirse firmemente allí al ADN. Esto sucede en una región del ADN llamada *promotor* que está inmediatamente “aguas arriba” del sitio de comienzo de la transcripción. La forma que tienen las células eucariotas (como las levaduras y las células animales) de lograr la expresión génica es mediante moléculas accesorias que se pegan al ADN y “ayudan” a que la ARN polimerasa encuentre el promotor, son los llamados Factores de Transcripción *basales* (o *generales*), y son necesarios esencialmente para la transcripción de todos los genes que codifican para proteínas (Fig. 1.3) [79].

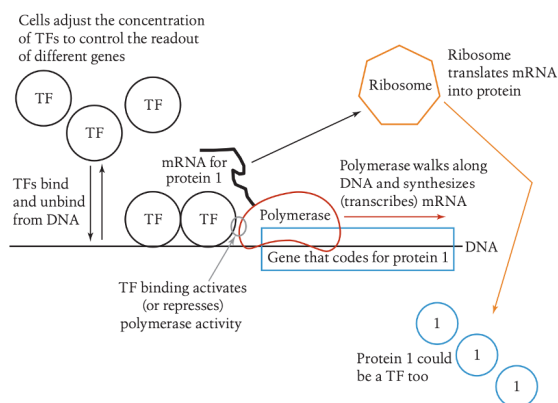


Figura 1.3: Figura esquemática que ilustra la regulación de la expresión génica mediante factores de transcripción (TF). Un componente importante de esta regulación está dada por la unión de FTs al ADN en sitios específicos. En el esquema se muestra una “entrada” (la concentración del FT) y una “salida” (la concentración de la proteína). En las células eucariotas el ADN está en el núcleo y el ARNm se transporta al citoplasma, donde se produce la traducción. Figura tomada de [13]

No basta con que la ARN polimerasa se una al promotor vía los FT basales. Para iniciar la transcripción, las ARN polimerasas requieren la ayuda de un conjunto amplio de proteínas accesorias. Estas proteínas reciben señales de activación o represión de un segundo tipo de FTs, conocidos como FTs *regulatorios*. A diferencia de los basales, los FT regulatorios se unen a secuencias de ADN más alejadas del inicio de la transcripción y solo lo hacen en los genes que van a controlar su expresión, ya que solo ellos tienen las secuencias en el ADN reconocidas por estos FTs. Este tipo de FT, una vez unido al ADN, puede interactuar con la maquinaria de FT basales para inducir la transcripción del gen. Un subgrupo de estos FTs son los responsables de identificar al gen que corresponda dependiendo de la cascada de señalización que haya sido activada por el estímulo externo. La arquitectura del ADN permite que un solo gen sea controlado por una variedad extensa de FTs regulatorios como se ilustra en la Fig. 1.4. En el resto de la Tesis, cuando nos referimos a FTs, hablamos específicamente de esta clase, los regulatorios.

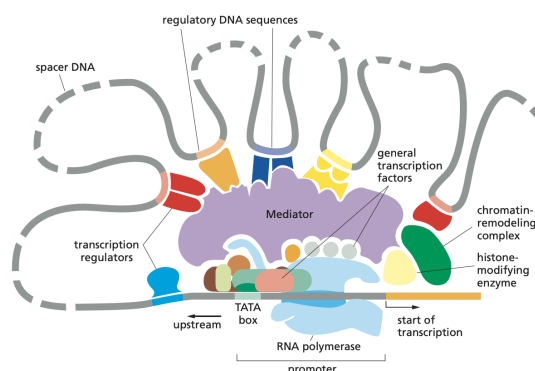


Figura 1.4: La regulación de la transcripción en células eucariotas involucra muchos componentes que funcionan en forma coordinada. Los reguladores y las proteínas modificadoras de la cromatina son ensamblados en el promotor por un mediador combinando así los múltiples factores que determinan la tasa final de inicio de la transcripción. Figura tomada de [2]

1.3. Vías de señalización que codifican por amplitud

Un ejemplo muy estudiado de señalización en células donde el estímulo externo está codificado por amplitud es el de la respuesta canónica a la feromona en células de levadura ilustrada en la Fig. 1.1. El ciclo de vida de la levadura tiene una fase diploide (con 16 pares de cromosomas). Las células diploides, en condiciones de hambre, pueden esporular, un proceso que involucra realizar meiosis, formando 4 esporas haploides (16 cromosomas) por cada célula diploide, de 2 sexos llamados $MATa$ y $MAT\alpha$. Estas esporas germinan y secretan feromonas, factor a o α , respectivamente, por la que células del sexo opuesto se sienten atraídas. Esto induce un crecimiento celular polarizado y eventualmente puede conducir a la fusión celular y la formación nuevamente de una célula diploide. Los experimentos suelen utilizar células $MATa$ a las que se somete a distintas cantidades de factor α en el baño. Expresando algunas de las proteínas involucradas en la vía fusionadas a una proteína fluorescente es posible observar y cuantificar los cambios que se inducen en el interior celular. Como se ilustra en la Fig. 1.1 el primer paso es la unión de la feromona a un receptor (GPCR). Luego se activan varias proteínas que inducen cambios en la fracción activa de FT de la vía, Ste12. Algunas de las proteínas expresadas como consecuencia de este proceso llevan al crecimiento polarizado hacia la potencial pareja. Para ello no solo es necesario detectar la presencia de la feromona, sino también inferir la dirección en la que su concentración es mayor. En la respuesta canónica, concentraciones crecientes de feromona conducen a un aumento en la concentración o en la activación de diversos intermediarios de la cascada, incluidos los FTs que regulan la expresión génica. Esto se ilustra en la Fig. 1.5 tomada de [110]. Se

puede observar que la dependencia entre las distintas cantidades observadas y la concentración del factor α queda bien descrita por una función de Hill [50]:

$$r = \frac{X^n}{EC_{50}^n + X^n}, \quad (1.1)$$

donde X , en el ejemplo, corresponde a la concentración de factor α y r a la respuesta observada (normalizada como en la figura).

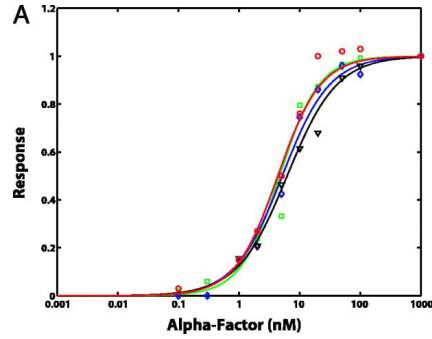


Figura 1.5: Figura tomada de [110] donde se muestra la superposición de diferentes curvas dosis-respuesta en la respuesta a feromona en *S. cerevisiae*. En todos los casos las células fueron expuestas a concentraciones de la feromona, factor α , entre 0.1 nM y $1\mu\text{M}$ y se midió luego la respuesta observando distintos componentes de la vía de señalización. En la figura se muestran los datos normalizados, en cada caso, a la respuesta a $1\mu\text{M}$ de: (i) la afinidad del receptor de membrana (triángulos negros), (ii) la activación de la proteína G (círculos rojos), (iii) la inducción transcripcional de la proteína PFUS1-GFP (cuadrados verdes) y (iv) la detención de ciclo celular (rombos azules). Las curvas llenas corresponden a curvas de Hill de la forma de la Ec. (1.1) que ajustan a los datos.

La *ecuación de Hill*, Ec. (1.1), caracteriza una curva típica Estímulo-Respuesta. Se obtiene, en particular, al plantear el equilibrio de una reacción química entre un ligando, L , y otra molécula (*e.g.*, un receptor), R , en la que n moléculas de L se pueden pegar/despegar a/de una molécula de R de acuerdo a:



Suponiendo que la reacción ocurre en un ámbito donde las sustancias están uniformemente distribuidas en el espacio y del que no salen ni entran moléculas, se deduce de (1.2) que $R_T \equiv [R] + [R']$ es constante y que, en el equilibrio la fracción de moléculas de R con n ligandos pegados, $r \equiv [R']/R_T$, y la concentración de ligando, $X \equiv [L]$, satisfacen la Ec. (1.1) con $EC_{50}^n = k_{off}/k_{on}$. Si $n > 1$, la ecuación de Hill es una aproximación simplificada ya que supone que n moléculas se pegan simultáneamente, lo que no suele ocurrir. En realidad, esta simplificación corresponde a suponer que es muy poco probable encontrar a R con entre 1 y $n - 1$ moléculas de X pegadas. La ecuación tiene dos parámetros.

El primero, EC_{50} , representa el valor que toma la concentración de ligando para que un 50% de las moléculas de R tengan n moléculas de ligando pegado. El segundo, el número de Hill n representa el grado de cooperatividad en el pegado de las moléculas de ligando. Si $n = 1$, esto representa falta de cooperatividad. Es decir, la probabilidad de que se pegue un ligando a R es independiente de si ya hay otro ligando pegado o no. Si $n > 1$ existe *cooperatividad positiva*. En este caso la probabilidad de pegar un ligando aumenta si ya hay otros ligandos pegados. Las curvas Estímulo-Respuesta como las de la Fig. 1.5 son estáticas: la respuesta es representada por una función instantánea del estímulo.

1.4. Codificación por frecuencia

Un mensajero universal secundario que presenta comportamiento pulsátil y que codifica los estímulos externos en la frecuencia entre sus pulsos es el calcio (Ca^{2+}). El Ca^{2+} es crucial en la señalización tanto durante la vida de las células -para indicar, por ejemplo, la necesidad de crecer- como en la muerte -apoptosis- [10]. También tiene un rol clave en la señalización neuronal y en la contracción muscular. Las oscilaciones de Ca^{2+} intracelular [28] regulan la actividad enzimática [23, 29] y pueden incrementar la expresión génica [25]. Existen también numerosos ejemplos de vías de señalización que resultan en fracciones nucleares pulsátiles de FTs. En particular, esto fue observado en los factores Msn2 y Crz1 en *S. cerevisiae* cuando las células son sometidas a estrés y en la activación del camino ERK (quinasas reguladas por señales extracelulares) / MAPK (proteína kinasa activada por mitógenos) [1, 17, 48]. Otro FT que presenta comportamiento pulsátil, pero en células de mamífero, es el p53, el que cumple un rol fundamental para impedir el desarrollo de varios tipos de cáncer [5, 7, 64, 108].

1.4.1. Señales intracelulares de calcio.

En condiciones basales, la concentración de Ca^{2+} libre en el citosol es muy baja. Las señales intracelulares de Ca^{2+} se construyen a partir del ingreso de los iones desde el medio extracelular o desde reservorios internos de la célula, generalmente, a través de canales específicos que cambian de conformación frente a distintos tipos de estímulos [10]. Las respuestas que el Ca^{2+} puede inducir dependen de la distribución espacio-temporal de su concentración en el citosol. En particular, la liberación de Ca^{2+} desde el retículo endoplasmático (RE) a través de canales receptores de IP_3 , IP_3Rs , dan a las células las herramientas para producir diferentes patrones espacio-temporales. Los IP_3Rs necesitan ligar IP_3 y Ca^{2+} para poder abrirse. Por lo tanto, la apertura de uno de estos canales puede conducir a la apertura de otros vecinos a través del mecanismo de *liberación de Ca^{2+} inducida Ca^{2+}* (CICR) [34]. La Fig. 1.6 (a) ilustra algunos de los varios procesos que modulan las señales intracelulares de Ca^{2+} cuando la liberación de los iones desde el RE a través de IP_3Rs está involucrada. Los IP_3Rs suelen estar organizados en cúmulos que contienen $\sim 10 - 20$ canales. Por

otro lado, las altas concentraciones de Ca^{2+} inhiben a los IP_3Rs . Así, las señales pueden variar desde *blips* y *puffs* donde la liberación de Ca^{2+} ocurre a través de un único canal o de un cúmulo de IP_3Rs , hasta ondas que se propagan por toda la célula, dependiendo del nivel de IP_3 [94, 99, 109] y de la interacción entre la activación e inhibición por Ca^{2+} . El acoplamiento entre cúmulos mediado por Ca^{2+} es fundamental para la propagación de las ondas, como se puede concluir a partir de experimentos en los que se impide la propagación mediante el uso de “buffers” lentos de Ca^{2+} como el EGTA [18, 80, 82].

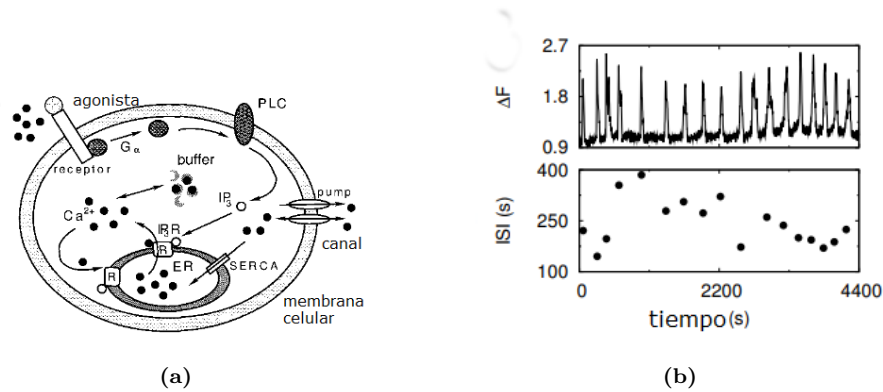


Figura 1.6: (a) Esquema ilustrando algunos de los varios mecanismos que modulan la dinámica del Ca^{2+} intracelular en un caso en que hay liberación desde el retículo endoplasmático a través de IP_3Rs . Una estimulación externa activa la proteína G induciendo la producción de IP_3 a partir de precursores presentes en la membrana. La unión de IP_3 y Ca^{2+} al IP_3R induce la apertura del canal permitiendo el ingreso de Ca^{2+} desde el RE al citosol. Estos iones difunden permitiendo la apertura de IP_3Rs vecinos. A su vez, muchos mecanismos comienzan actuar para bajar la $[\text{Ca}^{2+}]$ citosólica. Entre ellos se encuentran el atrapamiento por parte de *buffers* de Ca^{2+} móviles e inmóviles, la acción de la bomba SERCA (Sarco-Endoplasmic Reticulum Calcium ATP-ase) que reingresa los iones al RE y de otras bombas e intercambiadores que sacan el Ca^{2+} hacia el medio extracelular. (b) Ejemplo de una observación típica obtenida mediante experimentos de microscopía en los que se monitorean señales de Ca^{2+} sin discriminar entre distintos puntos dentro de la célula. En estos experimentos las señales se observan usando fluoróforos que aumentan la intensidad de su fluorescencia al ligar Ca^{2+} . El ingreso de (suficiente) Ca^{2+} al citosol se ve entonces como un incremento en la fluorescencia, ΔF , por encima del nivel basal como se muestra en la parte superior de la figura. Abajo aparecen graficados los tiempos entre pulsos *ISI*. Figura tomada de Skupin & Falcke, 2007 [90].

Observación experimental de pulsos y oscilaciones de Ca^{2+} .

En muchos tipos de células diferentes, la estimulación constante con ligandos extracelulares produce secuencias de estas ondas que se visualizan con técnicas de microscopía de fluorescencia como secuencias de pulsos de Ca^{2+}

cuando el interior de la célula no se resuelve [28, 103]. Esto se ilustra en la Fig. 1.6 (b). Generalmente cada onda de Ca^{2+} (un pulso en la Fig. 1.6 (b)) está precedida por señales espacialmente localizadas que comienzan a ocurrir con más frecuencia hasta que se produce la onda [69]. Esto se ilustra en la Fig. 1.7 donde se muestran tanto imágenes de fluorescencia bidimensionales como registros al estilo del de la Fig. 1.6 (b) donde las señales más localizadas se observan como pulsos pequeños y las ondas como pulsos de mucha mayor amplitud. Este comportamiento es típico de los sistemas excitables espacialmente extendidos en los que los eventos locales aleatorios (*puffs* de Ca^{2+}) se amplifican a través del espacio eventualmente desencadenando un evento “extremo” (un pulso global de Ca^{2+}) [65]. Recordamos acá que los sistemas excitables están caracterizados por un estado estacionario estable y un umbral tal que, si es superado debido a perturbaciones, ocurre una larga excursión en el espacio de fases antes de que el sistema regrese al estado estacionario. Un nuevo pulso puede ser evocado luego de un período refractario [53].

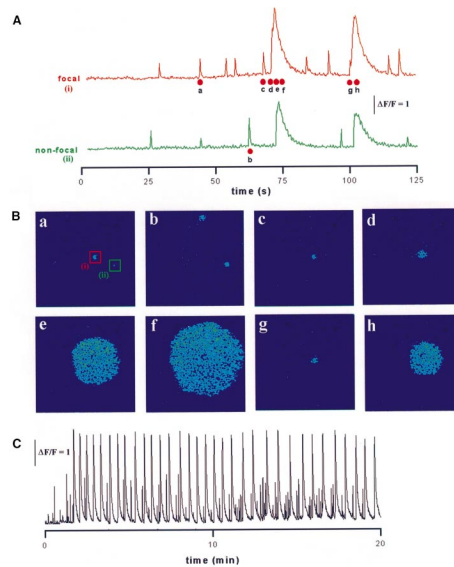


Figura 1.7: Figura tomada de Marchant *et al* [69] donde se pueden observar *puffs* y ondas de Ca^{2+} que ocurren en ovocitos de *Xenopus laevis* frente a la fotoliberación sostenida IP_3 . En (A) se muestra la fluorescencia integrada en regiones de $8\mu\text{m} \times 8\mu\text{m}$ alrededor de dos sitios identificados de liberación de Ca^{2+} (los que se identifican con las letras (i) y (ii) en (B) (a)). En (B) se muestran imágenes bidimensionales de fluorescencia sobre toda la región observada (de $65\mu\text{m} \times 65\mu\text{m}$) en los instantes indicados con puntos rojos en (A). En (C) se muestra el registro de fluorescencia recogido en la región de $8\mu\text{m} \times 8\mu\text{m}$ alrededor del sitio (i) en una escala de tiempo más lenta donde se pueden ver los pulsos pequeños correspondientes a señales localizadas y los más grandes correspondientes a ondas de Ca^{2+} ocurridos durante los 20 minutos que duró la fotólisis sostenida de IP_3 .

M. Falcke *et al* [90, 92, 102, 103] estudiaron sistemáticamente las propiedades de los pulsos de Ca^{2+} intracelular evocados por concentraciones constantes de ligandos externos en distintos tipos de células. Algunas de las observaciones se ilustran en la Fig. 1.8 tomada de [103]. Los primeros trabajos se enfocaron en responder si los pulsos ocurrían “estocásticamente o como un reloj” [92] concluyendo que tenían un componente azaroso importante. Más específicamente, determinaron que el intervalo de tiempo entre dos pulsos sucesivos (ISI por Inter Spike Interval, o T_{IS}) tiene dos componentes. El primer componente es determinista, pero variable entre células distintas, T_{min} . Este mínimo valor que puede tomar el tiempo entre dos pulsos sucesivos está determinado por la duración del pulso y un tiempo “refractario” durante el cual no se pueden evocar nuevos pulsos. El otro componente, t_{stoch} , es estocástico y tiene la particularidad de que su valor medio y su varianza cambian de célula a célula. Sin embargo, la relación entre este valor medio y el desvío estándar es una relación invariante lineal:

$$T_{IS} = T_{min} + t_{stoch} \quad (1.3)$$

$$\langle t_{stoch} \rangle = k\sigma_{ts} \quad (1.4)$$

con σ_{ts} el desvío estándar de t_{stoch} .

En [92] se mostró evidencia de este comportamiento en astrocitos y microglía de ratones, en células HEK-293 (provenientes de riñón de embrión humano) y en células PLA (células grasas provenientes de lipoaspiraciones). En [103] se confirmó este comportamiento para las células de la Fig. 1.8 como se muestra en la Fig. 1.9.

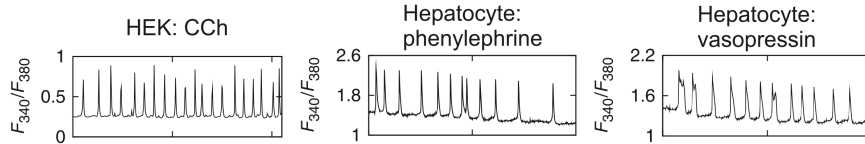


Figura 1.8: Pulsos de Ca^{2+} observados en células HEK293 en presencia de $30\mu\text{M}$ de carbachol, en hepatocitos en presencia de $1\mu\text{M}$ de fenilefrina y nuevamente en hepatocitos en presencia de 10nM de vasopresina. Estas señales fueron observadas usando el indicador, Fura-2, por lo que el aumento en la concentración de Ca^{2+} se reporta dando el cociente de la fluorescencia en dos rangos de longitudes de onda (F340/F380). Figura tomada de Thurley *et al* [103].

La Fig. 1.9 muestra que, en los experimentos, el valor medio de la parte estocástica del tiempo entre pulsos y su desvío standard están relacionados por la Ec. (1.4) pero con $k < 1$. Es decir, la secuencia de pulsos no constituye un proceso de Poisson con un tiempo entre pulsos distribuido exponencialmente. Este factor es una consecuencia de que la dinámica implícita tiene un *Feedback Negativo Global*, o GNF, el cual es modelado por Skupin y Falcke [91] mediante una distribución de Gompertz-Makeham como describimos más adelante.

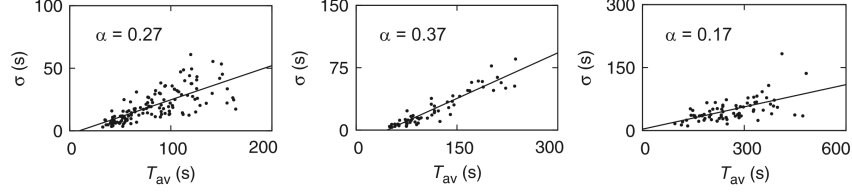


Figura 1.9: Desvío estándar *vs* promedio de la parte estocástica del tiempo entre pulsos, acá llamado T_{av} , calculados a partir de los experimentos ilustrados en la Fig. 1.8. Cada punto en la figura corresponde a una célula individual. Figura tomada de Thurley *et al* [103].

El análisis de los experimentos de [103], ilustrados en las Figs. 1.8-1.9 mostró, por otro lado, que la parte estocástica del tiempo medio entre pulsos decrecía exponencialmente con la concentración del ligando externo, $C = [L_{ext}]$, tanto cuando se hacía la estadística en células individuales, $T_{av} \approx \langle T_{IS} \rangle$, como cuando se consideraba el valor medio sobre una población de células del mismo tipo, T_{pop} :

$$T_{av,pop} = A_{av,pop} \exp(-\gamma C) + T_{min}, \quad (1.5)$$

con aproximadamente el mismo valor de γ para cada célula de un tipo y para la población de células de ese mismo tipo. La Fig. 1.10 (tomada de [103]) muestra gráficos de T_{pop} como función de la concentración de los ligandos externos correspondientes para los experimentos ilustrados en las Figs. 1.8-1.9 y para experimentos adicionales realizados en células de glándula salival. En todos los casos las curvas de línea llena están dadas por la Ec. (1.5) con parámetros como se indica en el epígrafe de la figura.

Modelos de pulsos y oscilaciones de Ca^{2+} .

Desde hace varias décadas se considera que hay información codificada en la frecuencia de las oscilaciones de Ca^{2+} intracelular, en particular, de las mediadas por IP_3Rs [11]. Los primeros modelos matemáticos surgidos para describir estas oscilaciones fueron deterministas. Acá describiremos en detalle solo uno de ellos que utilizaremos más adelante, el modelo de un solo “pool” de Dupont y Goldbeter [27].

Las ecuaciones dinámicas del modelo son:

$$\begin{aligned} \dot{C} &= v_0 + v_1\beta - v_{M2} \frac{C^n}{K_2^n + C^n} + \beta v_{M3} \frac{C^p}{K_A^p + C^p} \frac{Y^m}{K_R^m + Y^m} \\ &\quad + k_f Y - kC, \end{aligned} \quad (1.6)$$

$$\dot{Y} = v_{M2} \frac{C^n}{K_2^n + C^n} - \beta v_{M3} \frac{C^p}{K_A^p + C^p} \frac{Y^m}{K_R^m + Y^m} - k_f Y, \quad (1.7)$$

donde C es la concentración de Ca^{2+} libre en el citosol y se puede suponer que Y es proporcional a la $[\text{Ca}^{2+}]$ libre en el lumen del retículo endoplásmico

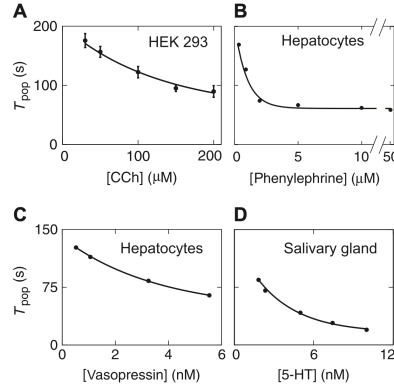


Figura 1.10: T_{pop} como función de la concentración de ligando para los experimentos de la Fig. 1.8 y para un experimento adicional realizado con glándulas salivarias de insecto en presencia de hormona 5-hidroxitriptamina (5HT). Las curvas de línea llena están dadas por la Ec. (1.5). En (A) se grafica el promedio y desvío estándar de los valores, T_{av} , obtenidos para células de HEK293 individuales. En este caso los parámetros que definen la línea llena, $\gamma = 7,84mM^{-1}$ y $T_{min} = 57s$, fueron derivados a partir de un análisis separado no mostrado en esta Tesis. En (B), (C) y (D) los valores de T_{pop} derivan de datos previamente publicados y los que definen las curvas de línea llena ($\gamma = 1,059\mu M^{-1}$, $T_{min} = 61s$ y $\gamma = 0,279\mu M^{-1}$, $T_{min} = 44s$ para hepatocitos en presencia de fenilefrina y de vasopresina, respectivamente y $\gamma = 0,319nM^{-1}$, $T_{min} = 16s$ para glándula salival) fueron obtenidos fiteando estos datos con una expresión de la forma (1.5).

(RE) donde la constante de proporcionalidad tiene en cuenta la relación entre los volúmenes luminal y citosólico y las capacidades “buffer” en ambos reservorios [66]; v_0 representa el flujo entrante de Ca^{2+} desde el medio extracelular; los dos términos proporcionales a β representan la liberación IP_3 -dependiente de Ca^{2+} desde el RE con $0 \leq \beta \leq 1$, la fracción de IP_3Rs con IP_3 ligado pasibles de abrirse al ligar Ca^{2+} ; el término proporcional a v_{M2} representa el bombeo de Ca^{2+} hacia el RE, el término $k_f Y$ representa la fuga de Ca^{2+} desde el RE al citosol y el último término en la Ec. (1.6), la remoción del Ca^{2+} citosólico libre por diversos medios. Mostramos en la Tabla 1.1 los parámetros utilizados en [27] que, para $\beta = 0,4$ dan lugar a oscilaciones como se muestra en la Fig. 1.11.

Ahora bien, como demostraron los análisis de Falcke *et al* descritos anteriormente, los pulsos intracelulares de Ca^{2+} que involucran la liberación desde el RE a través de IP_3Rs tienen una componente estocástica importante como se refleja en la Ec. (1.4). Por lo tanto los pulsos no corresponden a oscilaciones periódicas como las de la Fig. 1.11. Por otro lado, el hecho de que $k \neq 1$ en la Ec. (1.4) significa que el “disparo” de los pulsos no es un proceso de Poisson que ocurre a una tasa constante. Los experimentos de [69] ilustrados en la Fig. 1.7 muestran que los pulsos son, en realidad, ondas que se propagan por toda la célula que son precedidas por señales espacialmente localizadas (*puffs*).

Parámetro	Valor
v_0	$3,4\mu Mmin^{-1}$
v_1	$3,4\mu Mmin^{-1}$
v_{M2}	$50\mu Mmin^{-1}$
v_{M3}	$650\mu Mmin^{-1}$
K_A	$0,9\mu M$
K_2	$1\mu M$
K_R	$2\mu M$
k_f	$1min^{-1}$
k	$10min^{-1}$
p	4
m	2
n	2

Tabla 1.1: Parámetros del modelo definido por las Ecs. (1.6)–(1.7) usados en [27].

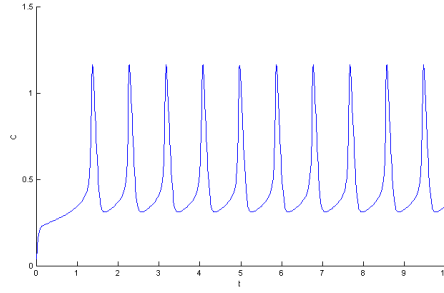


Figura 1.11: Oscilaciones de C en el modelo de Dupont-Goldbeter definido por las Ecs. (1.6)–(1.7).

La transición de señales localizadas a globales se puede modelar como un proceso de percolación [96] en donde el Ca^{2+} citosólico acumulado baja el umbral de “disparo” de nuevas señales [65]. Así, el citosol parece actuar como un sistema excitable modulado por Ca^{2+} en el que los pulsos son el resultado de un proceso de nucleación [90]. Por otro lado, una vez liberado el Ca^{2+} desde el RE durante un evento global, muchos IP_3Rs pasan a un estado inhibido debido a la alta $[Ca^{2+}]$ en el citosol. Por lo tanto, para que la célula esté en condiciones de producir un nuevo evento de liberación global, el Ca^{2+} debe reingresar al RE y los IP_3Rs deben salir de la inhibición. Ahora bien, no es necesario que el RE esté totalmente cargado o que todos los IP_3Rs hayan salido de la inhibición para que ocurra un nuevo pulso. Los pulsos pueden ocurrir, aunque con menor probabilidad, si una fracción de los IP_3Rs sigue inhibida. Estos aspectos fueron considerados en el modelo sencillo de disparo de pulsos propuesto por Skupin y Falcke en [90]. En este modelo se supone que, luego de un pulso, no solo debe pasar un tiempo determinista, T_{min} , desde el pulso anterior, sino que

después hay una tasa de recuperación que modifica la probabilidad de disparo. Más específicamente, en el modelo la probabilidad por unidad de tiempo de que se dispare un nuevo pulso una vez transcurrido un tiempo $T_{IS} = T_{min} + t_{stoch}$ desde el disparo del último pulso está dada por:

$$p(t_{stoch}) = \lambda(1 - e^{-\rho t_{stoch}}) \exp\left(-\int_0^{t_{stoch}} \lambda(1 - e^{-\rho t'}) dt'\right), \quad (1.8)$$

donde ρ es la tasa de recuperación y λ la tasa de disparo para $\rho = 0$. La tasa de disparo resultante, una vez transcurrido T_{min} , está dada por $\Lambda(t) = \lambda(1 - \exp(-\rho t))$. La densidad de probabilidad de t_{stoch} corresponde a una distribución de Gompertz-Makeham. Se muestra en la Fig. 1.12a esta distribución para distintos valores de ρ y λ .

La evolución temporal de una distribución así se puede ver en 1.12a y los datos para los distintos tipos de células se pueden ver en 1.12b. En otras palabras, la tasa de nucleación -la cual de ocurrir dispararía un pulso- no es una constante λ como en el caso de un proceso de Poisson, sino una cantidad que crece en el tiempo, hasta converger a una tasa máxima.

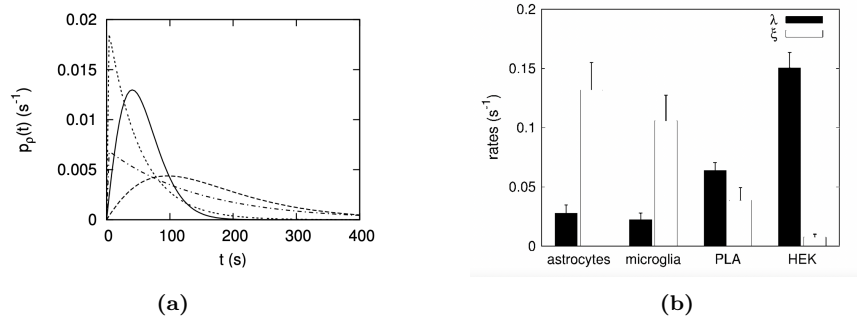


Figura 1.12: (a) Densidad de probabilidad de t_{stoch} (t en la figura) dada por la Ec. (1.8). Para valores altos de $\rho(t)$, se fusiona en un proceso de Poisson puro como se muestra para $\rho = 1$ y $\lambda = 0,02$ (línea de puntos) y $\lambda = 0,007$ (línea de puntos). Con ρ decreciente el máximo de la distribución, se desplaza a valores más altos de t . Para ρ fijo = 0,01, el ancho aumenta al disminuir λ como para el ejemplos $\lambda = 0,06$ (línea continua) y $\lambda = 0,01$ (línea discontinua). Figura tomada de [90]. (b) Valores de λ y ρ (ψ en la figura) para distintos tipos de células. Figura tomada de Skupin& Falcke[92].

1.4.2. Comportamiento pulsátil de factores de transcripción.

Como mencionamos en la introducción a esta Sección, existen numerosas situaciones en las que un estímulo externo constante provoca pulsos de acumulación nuclear del FT involucrado en la respuesta [61]. El examen exhaustivo del comportamiento pulsátil de varios FTs en levaduras realizado por Dalal *et al* [22] mostró que el 8 % de los FTs analizados (es decir, 10) se caracterizaban

por pulsos de localización nuclear, como se puede ver en la Fig. 1.13. En particular, el FT Crz1 responde a un aumento del Ca^{2+} extracelular con pulsos estereotipados de acumulación nuclear, cuya frecuencia aumenta con la concentración de Ca^{2+} [17]. Es interesante notar que la dependencia de la frecuencia con la $[\text{Ca}^{2+}]$ extracelular puede ajustarse bien por una exponencial como se ilustra en la Fig. 1.14 realizada con datos extraídos de la Fig. 2.a de [17]. Esto es indicativo de que la dependencia exponencial puede ocurrir en varios sistemas distintos. Como describimos más adelante y exploramos en el Capítulo 5 pensamos que se debe a la excitabilidad de la dinámica subyacente.

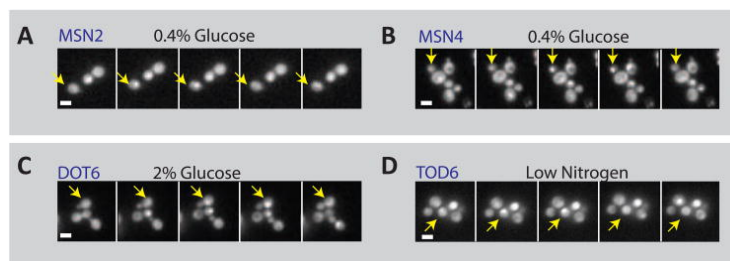


Figura 1.13: Secuencias de imágenes que revelan pulsos en la localización nuclear de algunos FTs en células de levadura. Tomado de Dalal *et al* [22].

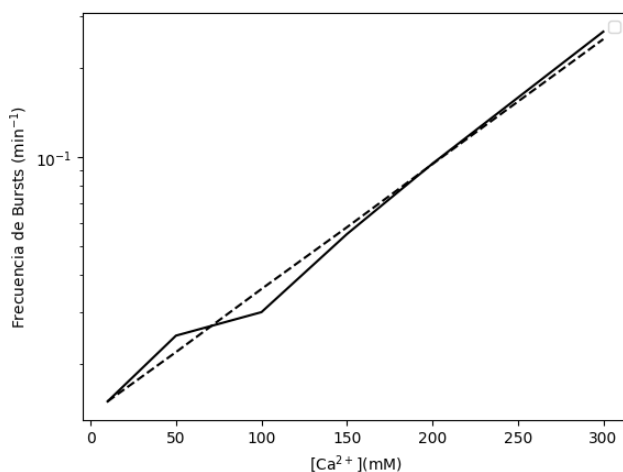


Figura 1.14: Frecuencia de las ráfagas de localización nuclear del FT Crz1 en células de levadura como función de la concentración de Ca^{2+} externo. La curva llena fue hecha con datos tomados de la Fig.2.a de [17]. La curva en líneas quebradas corresponde a una función exponencial entre frecuencia y $[\text{Ca}^{2+}]$.

En células de mamíferos, el FT $\text{NF-}\kappa\text{B}$ responde a una variedad de estímulos con pulsos de acumulación nuclear. En este caso, la intensidad del estímulo no afecta la frecuencia sino la amplitud de los pulsos nucleares [52]. El caso del FT

de tipo dedos de zinc, Msn2 [55, 78], que actúa en células de levadura frente a distintas condiciones de estrés es quizás más sorprendente. Como se muestra en la Sec. 1.5, la dinámica de su traslocación nuclear depende del tipo de estímulo.

1.4.3. Excitabilidad y pulsos

Un sistema excitable es un sistema dinámico que tiene un único punto fijo como atractor, pero cuya respuesta varía dependiendo de si se le aplica una perturbación pequeña, lineal, o una perturbación grande. Si la perturbación es tal que el sistema supera un umbral, o *threshold*, éste realiza entonces una larga excursión en el espacio de parámetros antes de volver al punto fijo, creando lo que se conoce como *spike* (*espiga* en castellano). El retorno al punto fijo involucra un tiempo de recuperación que debe transcurrir para que una perturbación suficientemente grande pueda generar un nuevo *spike*. Uno de los modelos paradigmáticos de sistemas excitables es el de FitzHugh-Nagumo [37], que se basa en el modelo más general de Hodgkin-Huxley [51] desarrollado para describir el potencial de acción en el axón gigante del calamar. El modelo de Hodgkin-Huxley está dado por:

$$C_m \frac{dV_m}{dt} = -g_K n^4 (V_m - V_k) - g_{Na} m^3 h (V_m - V_{Na}) - g_l (V_m - V_l) + I_{ext}, \quad (1.9a)$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1 - n) - \beta_n(V_m)n, \quad (1.9b)$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1 - m) - \beta_m(V_m)m, \quad (1.9c)$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1 - h) - \beta_h(V_m)h. \quad (1.9d)$$

y el de FitzHugh-Nagumo por:

$$\dot{v} = v - \frac{v^3}{3} - w + RI_{ext}, \quad (1.10a)$$

$$\tau \dot{w} = v + a - bw. \quad (1.10b)$$

Las variables del modelo de Hodgkin-Huxley son el potencial de membrana, V_m , y tres variables relacionadas con los dos canales principales que determinan la dinámica durante el potencial de acción: los de Na^+ (h y m) y los de K^+ (n). El potencial de acción se observa como un *spike* en V_m que se produce frente a la aplicación de un pulso de corriente externa, I_{ext} , que hace que V_m se aleje suficientemente del estado estacionario para, luego de un período refractario, regresar a él. En este modelo las variables V_m y m son rápidas mientras que las otras dos, que, de acuerdo a observaciones, aproximadamente satisfacen $h + n = 0,8$, son más lentas [57](pp. 195-216). En el modelo de FitzHugh-Nagumo la variable v es rápida y representa el potencial de membrana mientras que w es una variable de recuperación más lenta ($\tau \gg 1$) que puede asociarse a los canales

de K^+ involucrados en dicha recuperación. La dinámica de las Ecs. (1.10) para distintos parámetros puede entenderse en términos de las nulclinas (las curvas en el plano (v, w) que satisfacen, respectivamente, $\dot{v} = 0$ y $\dot{w} = 0$) que en la Fig. 1.15 están graficadas en color magenta y amarillo-verdoso. En esta figura se grafica también una trayectoria del sistema (curva violeta) que muestra que, para los parámetros utilizados, el atractor es un ciclo límite. Debido a la separación de escalas temporales entre v y w se trata de un oscilador de relajación que, por tramos, se “pega” a las ramas estables de la nulclina cúbica (las que tienen pendiente negativa en la Fig. 1.15) y que en el medio realiza rápidas excursiones entre ambas ramas. La separación de escalas permite describir esta dinámica en términos de una u otra de las Ecs. (1.10): la Ec. (1.10b) con $w \approx \text{constante}$ para las excursiones rápidas entre ramas y $\dot{w} \approx 0$ para el trayecto sobre cada una de las ramas estables de la nulclina cúbica. La Ec. (1.10b) con $w \approx \text{constante}$ corresponde a un sistema con dos puntos fijos estables ubicados sobre cada una de las ramas estables de la nulclina cúbica separados por un “saddle” ubicado sobre la rama inestable (de ahí el nombre de ramas “estables” e “inestable”). El sistema oscila entonces entre las dos ramas estables y, al moverse a lo largo de una de ellas ($\dot{w} \approx 0$), se acerca a una bifurcación *Saddle Node* del sistema reducido en la que la rama desaparece y debe saltar a la otra. Esta bifurcación del sistema reducido es consecuencia del *Feedback* Negativo presente en el modelo. Este tipo de dinámica fue observada en muchos sistemas biológicos, por ejemplo, en la competencia para la captación de ADN del *Bacillus subtilis* [98]. El ejemplo típico de comportamiento excitable ocurre cuando las nulclinas se intersectan en la rama estable de la nulclina cúbica que está a la izquierda en la Fig. 1.15. En este caso, el punto fijo es estable, pero si la perturbación es tal que lleva al sistema a la derecha de la rama inestable, la trayectoria se acerca primero a la rama estable a la derecha de la Fig. 1.15, se mueve luego a lo largo de ella hasta que encuentra la bifurcación *Saddle Node* por lo que se desplaza a la rama estable izquierda y pasa a moverse a lo largo de esta última hasta que finalmente alcanza el punto fijo. Cuando un sistema excitable es sometido a perturbaciones continuas este comportamiento puede repetirse a lo largo del tiempo. Esto se manifiesta en secuencias de pulsos en la variable rápida que corresponden a las largas excursiones en el espacio de fases causadas por las perturbaciones. Es de interés entonces determinar los tiempos característicos que separan a los pulsos para distintas perturbaciones.

El estudio teórico de los sistemas excitables en presencia de ruido [77] en combinación con distintas observaciones experimentales, como la de que las neuronas presentan períodos irregulares de *spikes* [95], llevaron a la descripción de la *resonancia coherente* [44, 63] y de la *resonancia estocástica auto-inducida* [60]. La resonancia coherente ocurre cuando el sistema dinámico está cerca de una bifurcación homoclínica. En esa situación, el “salto” que cruza el umbral de la bifurcación actúa como una barrera de energía modulada por la perturbación externa [31, 63]. En el caso de la resonancia estocástica auto-inducida [60] un ruido aditivo en la variable rápida permite superar el umbral que separa las ramas estable e inestable de la nulclina rápida. De esta manera, la derivada temporal de la variable rápida se puede describir como si el sistema tuviera que

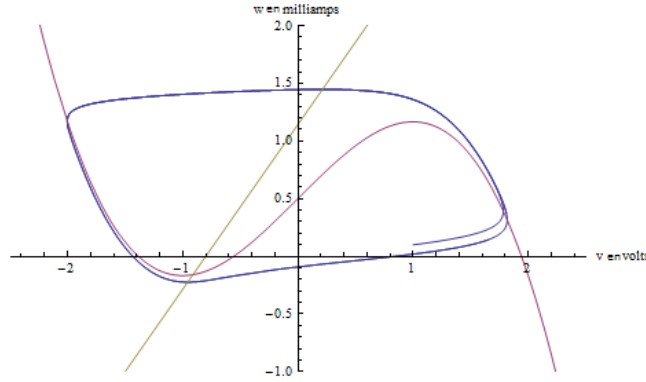


Figura 1.15: Nulclinas del modelo de FitzHugh-Nagumo ($\dot{v} = 0$ en magenta y $\dot{w} = 0$ en amarillo-verdoso). Los parámetros usados son $I = 0,5$, $a = 0,7$, $b = 0,8$ y $\tau = 12,5$. En este caso la nulclina lineal cruza a la cúbica en su rama inestable y el atractor del sistema es un ciclo límite como lo ilustra la trayectoria graficada en violeta. Tomado de wikipedia.

atravesar una barrera de potencial. Esto implica que la probabilidad con la que ocurren los pulsos está regulada por la Ley de Kramers como si se tratara de una reacción química [45]. En la teoría de Kramers las reacciones son modeladas como un paseo al azar (debido a fluctuaciones térmicas) a lo largo de una coordenada de reacción unidimensional, x , en presencia de un potencial, $U(x)$, como se ilustra en la Fig. 1.16. En esta figura los mínimos en x_a y x_c corresponden a estados metaestables del conjunto de reactivos involucrados y el máximo, x_b , al estado de transición. Kramers derivó una relación entre la forma de este potencial y la probabilidad de transición por unidad de tiempo entre los estados metaestables. La teoría se basa en aproximar el paisaje energético a segundo orden alrededor de sus extremos. En particular, se obtiene que la probabilidad de transición por unidad de tiempo de A a C está dada por:

$$k^+ = \frac{\left(U''(x_A) |U''(x_B)| \right)^{1/2}}{2\pi} \exp(-\beta E_b^+), \quad (1.11)$$

donde E_b^+ es la energía activación, $\beta = 1/(k_B T)$ con T la temperatura y k_B la constante de Boltzmann y $U'' = d^2U/dx^2$; con una ecuación análoga para k^- . Este cálculo se basa en una separación de escalas temporales que se satisface si $\beta E_b^{+,-} \ll 1$. La aplicación de esta teoría a otras situaciones como la de los sistemas excitables modulados por ruido requiere establecer una analogía entre la temperatura de las reacciones químicas y el ruido. En el Capítulo 4 establecemos esta analogía para estudiar el tiempo medio entre pulsos sucesivos de Ca^{2+} intracelular [41].

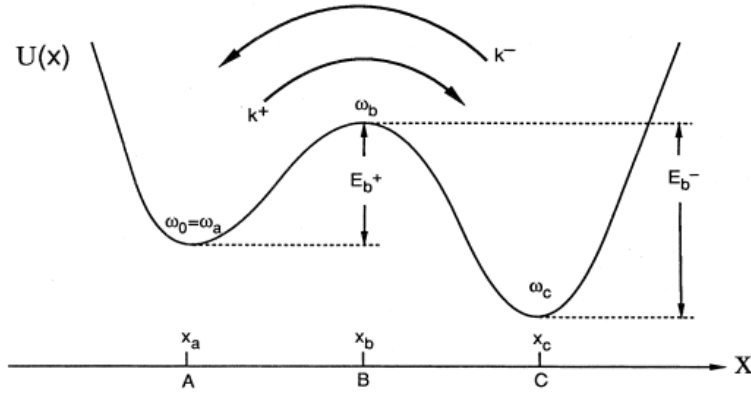


Figura 1.16: Se muestra una curva de energía, la cual es el contexto de donde surge la ecuación de Kramers. Tomado de Hänggi *et al.* [45]

Si bien los potenciales de acción en neuronas son ejemplos paradigmáticos de comportamiento excitable, este tipo de dinámica es relevante en otros sistemas de interés biológico. Los pulsos de Ca^{2+} intracelular, en particular, se han modelado a través de distintas ecuaciones, como las de DeYoung-Keizer [24], Li-Rinzel [62], Tang-Othmer -modelo que presenta resonancia coherente [100]-o Dupont-Goldbeter [27], que presentan comportamiento excitable para ciertos rangos de parámetros. Como ya mencionamos, suponemos que la dependencia exponencial del tiempo medio entre pulsos de Ca^{2+} mediados por IP_3Rs y la intensidad del *input* externo observada en muchos tipos celulares distintos [103] es una consecuencia de la dinámica excitable que subyace a estos pulsos. En particular, como se muestra en el Capítulo 5, es posible derivar esta dependencia aplicando la ley de Kramers al modelo sencillo de la dinámica de Ca^{2+} intracelular [27] dado por las Ecs. (1.6)–(1.7) [41]. Teniendo en cuenta que esta derivación se basa principalmente de la excitabilidad del sistema subyacente suponemos que la dependencia exponencial entre el tiempo medio entre pulsos y la intensidad del estímulo externos debe ocurrir para otros pulsos de relevancia biológica, en particular, los de localización nuclear de FTs. Refuerza esta hipótesis el comportamiento ilustrado en la Fig. 1.14 y el hecho de que se han observado dependencias en forma de función creciente convexa en varias situaciones diversas [22, 48, 111].

1.5. Multiplexación: un mismo FT con distintas dinámicas para responder a distintos estímulos.

Los estudios de [46–48] demostraron que las células pueden usar el mismo FT para modular la expresión de diferentes genes dependiendo de la dinámica

de traslocación nuclear del FT. La dinámica entonces puede servir para *multiplexar* la transmisión de información [71]. En este caso es relevante si diferentes estímulos externos pueden codificarse de manera confiable en diferentes dinámicas de las fracciones nucleares del FT que permitan su identificación [47]. En particular, los experimentos demostraron que el comportamiento de la fracción nuclear del FT Msn2 en células de *S. cerevisiae* depende del estímulo [48]. Por un lado, responde a la limitación de glucosa con pulsos de acumulación nuclear de frecuencia creciente cuanto más baja sea la concentración de glucosa. Responde a su vez al estrés osmótico con un pulso inicial de localización nuclear cuya duración depende de la intensidad del shock osmótico. Finalmente, responde al estrés oxidativo con una prolongada localización nuclear cuya amplitud (la fracción de Msn2 nuclear) crece con la magnitud del estrés. Esto se ilustra en la Fig. 1.17 tomada de [48].

El supresor de tumores de mamíferos p53 es otro FT que responde con distintas dinámicas según el estímulo. Mientras que la exposición a los rayos UV provoca un pulso único de amplitud y duración crecientes con dosis de UV crecientes, aparecen secuencias de pulsos cuando se rompe la doble hebra del ADN a causa de la radiación γ y el número de pulsos aumenta con el nivel de daño [108]. Estos ejemplos muestran que diferentes estímulos pueden codificarse en distintos “aspectos” de la fracción nuclear de un mismo FT [49, 87, 88]. Por otro lado, dinámicas distintas de localización nuclear producen respuestas distintas. Por ejemplo, el comportamiento pulsátil del p53 induce la expresión de genes de reparación del ADN mientras que un solo pulso sostenido de p53 conduce a la expresión de genes involucrados en gatillar la senescencia [108]. Algo similar ocurre con los FTs, NF- κ B, también en células de mamíferos [3, 38, 101], Crz1 en la levadura *S. cerevisiae* [17] y otros genes. El hecho de que un mismo FT sea usado para activar distintos genes en un mismo tipo celular permite suponer que el rol de la señalización pulsátil sea la de dotar a las células de mayor flexibilidad para la generación de respuestas.

Como mencionamos en la Sec. 1.2, los FTs regulan la transcripción génica ligándose a regiones específicas del ADN, comúnmente río arriba de los promotores. El hecho de que un mismo FT pueda generar distintas respuestas dependiendo de su dinámica implica entonces que esta dinámica afecta diferencialmente el pegado/despegado del FT y la consiguiente producción del ARNm asociado a las proteínas involucradas en cada tipo de respuesta. Hansen y O’Shea [46] estudiaron en qué diferían las propiedades de la interacción entre el FT Msn2 y los promotores asociados a las distintas respuestas ilustradas en la Fig. 1.17. Para tal fin analizaron 12 modelos numéricos distintos de la transcripción y eligieron

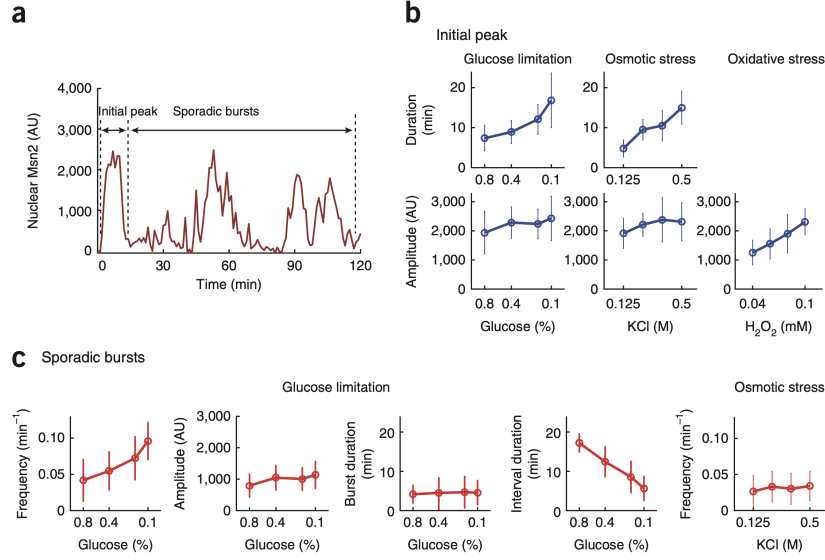
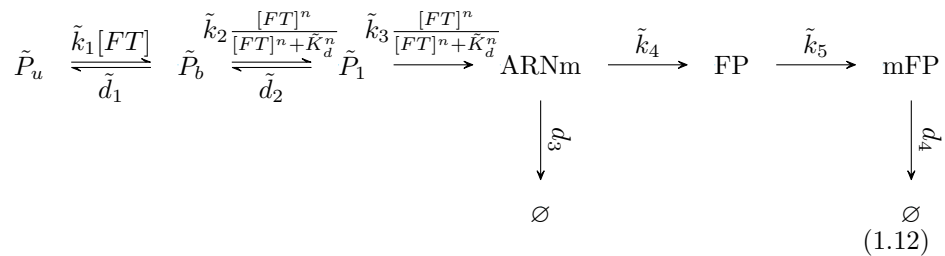


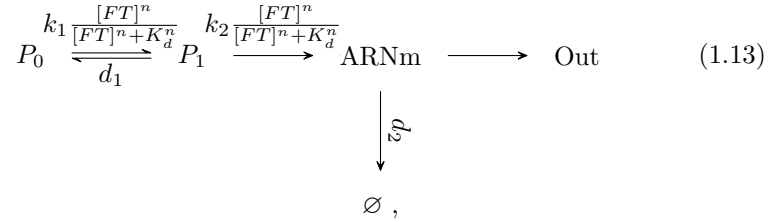
Figura 1.17: En la figura tomada de [48] se muestran las trazas de traslocación nuclear del Msn2 en células únicas de levadura frente a distintos estímulos: deprivación de glucosa, estrés osmótico (inducido introduciendo KCl en el baño) y estrés oxidativo (inducido introduciendo H_2O_2 en el baño). (a) El esquema define el pico de traslocación inicial y las ráfagas esporádicas posteriores sobre las que se hace el análisis presentado en las siguientes subfiguras. (b) Cuantificación de la duración (fila superior) y de la amplitud (fila central) del pico inicial para las condiciones de estrés indicadas (círculos abiertos: valor medio de células individuales; barras de error: desvío estándar calculado sobre ~ 60 células, en al menos dos experimentos independientes). La duración no se cuantifica para el tratamiento con H_2O_2 (estrés oxidativo) porque en esta condición solo se observó un evento de traslocación sostenida. (c) Frecuencia, amplitud y duración de las ráfagas esporádicas y del intervalo en el que se observan ráfagas en respuesta a la limitación de glucosa. Se muestra también la cuantificación de la frecuencia de ráfagas esporádicas bajo estrés osmótico. Figura tomada de Hao&O'Shea [48].

el que mejor ajustaba a los datos. El mismo está dado por:



donde las varias cantidades, \tilde{P} , corresponden a la probabilidad de que el pro-

motor esté en sus distintos estados posibles: \tilde{P}_u inactivo y sin FT ligado, \tilde{P}_b inactivo con FT pegado y \tilde{P}_1 activo y donde la flecha que va de \tilde{P}_1 a ARNm representa la producción de este último (es decir, la transcripción) a una tasa $\tilde{P}_1 \tilde{k}_3 [FT]^n / ([FT]^n + \tilde{K}_d^n)$ y no una transformación de \tilde{P}_1 a ARNm. En este modelo, solo cuando el promotor está en el estado activo el ARNm es transcripto. En los experimentos de [46], el ARNm produce una proteína fluorescente (FP), la cual debe madurar (y pasar a ser mFP) para ser observada. Esta proteína constituye la “salida” del proceso (*Out*) en los experimentos. A su vez, tanto el ARNm como la proteína son degradados a tasas d_3 y d_4 , respectivamente. En [46] los autores proponen también una versión simplificada del modelo:



donde como antes la flecha que apunta al ARNm representa la transcripción. Como discutimos en el Capítulo 3 (ver también [39]), donde presentamos un modelo mecánico a partir del cual se puede derivar el de la Ec. (1.13), P_0 corresponde a la probabilidad de que el promotor esté inactivo mientras que P_1 corresponde a la probabilidad de que esté activo. Solo cuando el promotor está en el estado activo y con FT pegado el ARNm correspondiente es transcripto, es por ello que la tasa de transcripción depende de P_1 y de la concentración de FT. En esta Tesis usamos este modelo simplificado.

Usando el modelo completo 1.12, o su versión simplificada, 1.13 Hansen y O’Shea analizaron sus resultados experimentales. Así pudieron mostrar que los genes activados por la modulación por duración tienen un umbral de activación alto (\tilde{K}_d relativamente grande en el modelo) y un tiempo de reacción lento ($\frac{k_1 k_2}{d_1}$) (que puede deberse a un mayor nivel de ocupación del ADN por nucleosomas), mientras que los genes activados por la modulación por frecuencia tienen un umbral bajo y un tiempo de reacción rápido. Por esta razón los llamaron HS y LF, respectivamente. Estos resultados, que se ilustran en la Fig. 1.18, pueden interpretarse del siguiente modo. El tiempo de reacción rápido hace a los genes de tipo LF aptos para ser modulados por frecuencia, mientras que los de tipo HS filtran las señales pulsátiles. El rol del umbral en favorecer la señalización modulada por duración es menos clara y será explorada en profundidad en esta Tesis. En principio, un umbral alto permite filtrar señales de baja amplitud (incluyendo ruido aditivo), mientras que un umbral bajo permite mayor sensibilidad para pulsos de baja amplitud.

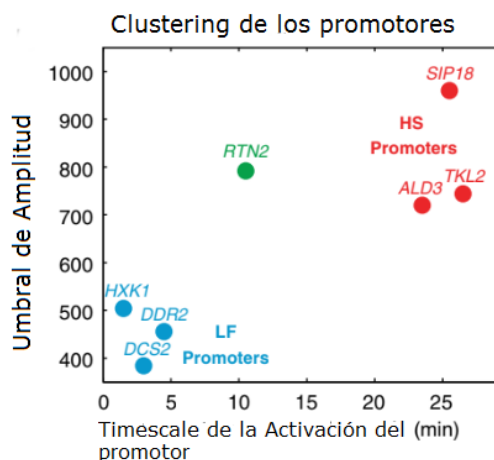


Figura 1.18: Tiempos y umbrales de activación estimados en base al modelo (1.12) para los genes analizados en [46]. Estos quedan separados por HS (alto umbral de activación y lenta respuesta) y LF (bajo umbral de activación y rápida respuesta). Tomado de Hansen&O’Shea [46].

1.6. Un mismo estímulo codificado en amplitud y frecuencia

Así como hay situaciones en las que las células usan dinámicas diferentes de un mismo FT para generar respuestas diferenciadas a distintos estímulos (ver Sec. 1.5), hay otras en las que un mismo estímulo puede codificarse simultáneamente en la amplitud o en la frecuencia de algunos de los componentes involucrados en la generación de la respuesta. Esto no es de extrañar dado que, frente a un estímulo que requiere una respuesta compleja, pueden “encenderse” distintas vías cada una de ellas encargada de inducir un subconjunto de las modificaciones involucradas en la respuesta. Esto ocurre, en particular, en la respuesta a feromona en células de levadura. Como mostramos en la Fig. 1.5, al aumentar la cantidad de feromona factor α , la respuesta canónica involucra la producción de cantidades cada vez mayores de diversos intermediarios de la cascada, incluidos algunos de los FTs que regulan la expresión génica. Es decir, la respuesta canónica codifica por amplitud. La presencia de feromona, sin embargo, evoca también la generación de pulsos de Ca^{2+} en estas células con mayor probabilidad que en ausencia de feromona. Por otro lado, la frecuencia de estos pulsos aumenta con la concentración de factor α , como se ilustra en la Fig. 1.19 tomada de [19]. Si bien el rol de los pulsos de Ca^{2+} en la vía de respuesta a la feromona no fue dilucidado aún (aparentemente son necesarios para la integridad celular), es posible que, como ocurre en la respuesta al estrés por Ca^{2+} en *S. cerevisiae* (ver Fig. 1.14), estén asociados a la traslocación pulsátil de algún FT. Independientemente de esto, surge la pregunta de por qué las células usarían

distintas estrategias para responder a un mismo estímulo. El trabajo descrito en el Capítulo 4 intenta, entre otras cosas, dar una respuesta a esta pregunta.

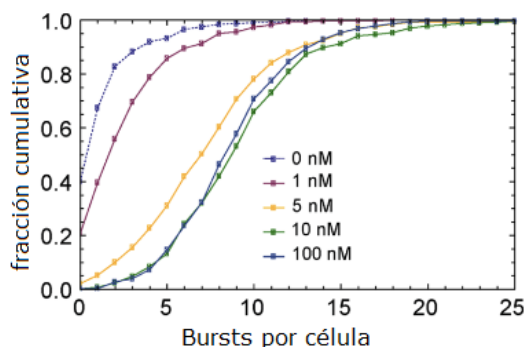


Figura 1.19: Pulsos de Ca^{2+} en células de *S. cerevisiae* observados para distintas concentraciones de feromona. La figura, tomada de Carbó *et al* [19], muestra la fracción acumulada del número de pulsos por célula para las concentraciones de factor α indicadas en la leyenda. Se observa que, en ausencia de feromona, los pulsos ocurren solo en un 60 % de las células (mitóticamente activas). En presencia de feromona aumenta la fracción de células que muestran pulsos (aproximadamente un 80 % en presencia de 1nM y 100 % para concentraciones de 10nM o más). Por otro lado, el número de pulsos por célula y, por la tanto, la frecuencia de los pulsos, aumenta con la concentración de factor α .

1.7. Teoría de la Información y su aplicación a problemas biológicos

Como enfatiza W. Bialek en su libro [13], existe la idea de que “lo importante en biología es el flujo de información”. Si bien la palabra información se usa a menudo en el marco de los sistemas biológicos, no hay tanto acuerdo sobre si las versiones matemáticas formales de este concepto son útiles en biología o no. W. Bialek se encargó de explorar la utilidad y aplicabilidad de la teoría introducida por Shannon a distintos problemas concretos de la biología. En particular, gran parte de su trabajo se basó en probar que los principios de optimización aplicados a distintas cantidades definidas por la teoría de la información permiten abordar problemas que van desde la señalización bioquímica en células individuales hasta el aprendizaje y la predicción en el cerebro en un marco matemático más o menos unificado. En esta Sección introducimos brevemente la teoría de Shannon y después describimos algunas aplicaciones a distintos problemas biológicos.

1.7.1. La teoría de Shannon

La teoría de Shannon se basa en definir la función “información” [86]. Esto es, qué tanta información subjetiva nos provee la experiencia de que ocurra un evento cuando éste tenía una probabilidad *a priori* de ocurrir p . Shannon define a la función información, \tilde{I} , como la única función que cumple:

- $\tilde{I}(p)$ decrece monótonamente con p (es decir, descubrir un evento nos provee menos información si era probable que ocurriera).
- $\tilde{I}(1) = 0$, es decir, la experiencia de un evento seguro no nos provee información.
- $\tilde{I}(p_1 p_2) = \tilde{I}(p_1) + \tilde{I}(p_2)$, es decir, la información que provee la experiencia de dos procesos independientes es la suma de las informaciones individuales.

La función información así definida sólo tiene una posibilidad matemática: $\tilde{I}(p) = -k \log(p)$, donde k define las unidades utilizadas (o bien la base del logaritmo elegida). Esta definición sirve para estimar cuánta información se gana al observar un evento dado o, equivalentemente, al medir una variable aleatoria que puede tomar un número discreto, n , de valores posibles y se obtiene el valor que ocurre con probabilidad p . Ahora bien, es de interés saber, en promedio, *a priori*, cuánta información se puede ganar midiendo dicha variable. En ese caso se define la entropía, H , como el valor esperado de la información que se gana al medir la variable, es decir:

$$H(\{p\}) = \langle I(p) \rangle = -k \sum_i p_i \log(p_i), \quad (1.14)$$

donde p_1, p_2, \dots, p_n , son las probabilidades con las que la variable puede tomar cada uno de sus valores posibles.

Para los problemas de comunicación es de interés definir una entrada (el mensaje que envía el emisor) y una salida (el mensaje que recibe el receptor) y estimar la información que puede recibir el receptor para un dado canal de comunicación. Supongamos que la entrada y la salida están representadas, respectivamente, por dos variables aleatorias discretas, i y o , cuya probabilidad conjunta es $p(i, o)$. Se define la entropía conjunta, $H(i, o)$ como:

$$H(i, o) = - \sum p(i, o) \log(p(i, o)), \quad (1.15)$$

donde la suma es sobre todos los valores que pueden tomar i y o . Se define, por otro lado, la entropía condicional, $H(o|i)$, entre la salida y la entrada como menos la información de la salida dado que se conoce el valor de la entrada. Esta entropía depende de la probabilidad condicional, $p(o|i)$, que satisface: $p(o|i) = \frac{p(o, i)}{p(i)}$, y que, en nuestro problema de interés, está definida por el canal de comunicación. La entropía condicional está dada por:

$$H(o|i) = - \sum_i p(i) \sum_o p(o|i) \log(p(o|i)) = - \sum_{i, o} p(i, o) \log \left(\frac{p(o, i)}{p(i)} \right), \quad (1.16)$$

donde las sumas se hacen sobre todos los valores que pueden tomar las variables i y o . Finalmente, se define la *información mutua* como la información que se “gana”, en promedio, cuando se mide la salida conociendo previamente la entrada respecto de cuando no se la conoce. Si la información total que se obtiene al medir la variable o sin conocer el valor de i es $-H(o)$ y la información que se obtiene conociendo i es $-H(o|i)$, la información mutua resulta:

$$MI(i, o) = -H(o|i) + H(o), \quad (1.17)$$

El concepto de información mutua se ilustra en la Fig. 1.20, en donde MI está denotada como I.

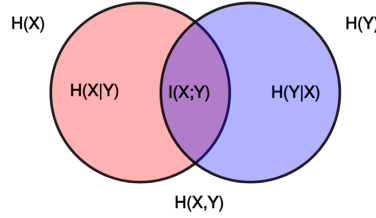


Figura 1.20: Se muestra la relación entre las entropías $H(X)$ y $H(Y)$, la entropía conjunta $H(X,Y)$, las entropías condicionales $H(X|Y)$ y $H(Y|X)$ y la información mutua $I(X;Y)$. Tomado de D Well, Wikipedia.

Crucialmente, la información mutua es una cantidad simétrica entre i y o (la información que la salida contiene sobre la entrada es cuantitativamente igual a la que la entrada contiene de la salida):

$$\begin{aligned} MI(i, o) &= H(o) - H(o|i) = -\sum p(i, o) \log(p(o)) + \sum p(i, o) \log\left(\frac{p(o, i)}{p(i)}\right) = \\ &= \sum p(i, o) \log\left(\frac{p(o, i)}{p(i)p(o)}\right), \end{aligned} \quad (1.18)$$

donde, como antes, las sumas se hacen sobre todos los valores que pueden tomar las variables o e i .

La información mutua va a ser la cantidad crucial que vamos a analizar a lo largo de esta Tesis.

1.7.2. Teoría de la información en sistemas biológicos: descripción general

Gran parte de las ideas usadas en esta Tesis provienen de las sugerencias originadas en el trabajo de Laughlin sobre la visión en moscas [59], y profundizadas en los trabajos de Bialek aplicados tanto al procesamiento en neuronas como

a la embriogénesis y a las vías metabólicas [13]. En el trabajo de Laughlin, se encuentra que en los ojos de las moscardas (*Calliphora Stygia*) la respuesta ante los contrastes de luz tiene una forma sigmoidea como función de la intensidad de luz. Esta respuesta coincide experimentalmente con la que se supondría si la información transmitida por el sistema estuviera optimizada a la distribución de entradas (de contrastes) externa, en el medioambiente en el cual esta moscarda se encuentra. Una de las mayores manifestaciones de esto es que la curva se encuentra más empinada (y, por tanto, el sistema es máximamente sensible) a los contrastes más frecuentes en ese ambiente. En cambio, la curva de respuesta es más plana y menos empinada para las entradas infrecuentes. Esto se puede ver en la Fig. 1.21.

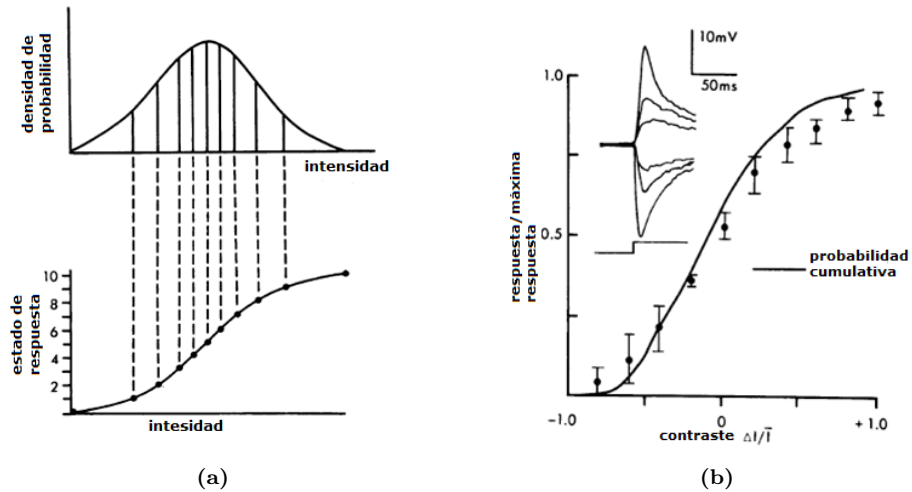


Figura 1.21: Se muestra la distribución de probabilidad de la entrada (proveniente del entorno externo) para que la curva dosis-respuesta mostrada sea la ideal para transmitir información. 1.21a. El resultado experimental muestra que la respuesta está dentro de la predicción teórica. 1.21b Tomado de Laughlin [59].

En general, Bialek analiza que para un sistema de señalización estático con un input con ruido aditivo pequeño (en régimen lineal) Gaussiano y homogéneo y un canal de amplitud con una curva dosis-respuesta, la información total transmitida es máxima si el *output no condicionado al input* tiene una distribución de probabilidad uniforme. De esta manera, (para ruido Gaussiano no homogéneo $\sigma_o(i)$) la curva dosis respuesta queda determinada por la ecuación: $\frac{d\bar{o}}{di} = Z\sigma_o(i)p(i)$, y si el ruido es homogéneo, la derivada es proporcional a la probabilidad del input, $\frac{d\bar{o}}{di} \propto p(i)$. Por otro lado, la desigualdad del procesamiento de datos (*data processing inequality*) [8] determina que en una cadena de Markov $X \rightarrow Y \rightarrow Z$, la Información Mutua entre el input y el output ($MI(X, Z)$) no es mayor que la información entre la entrada y el proceso intermedio. Teniendo en cuenta este principio se concluye que la información transmitida en un sistema

compuesto por varios pasos va a estar limitada por la capacidad de transmisión de los subsistemas río arriba y río abajo.

1.7.3. Teoría de la información en sistemas biológicos: descripción detallada de un ejemplo.

Las etapas tempranas en el desarrollo de embriones de la mosca de la fruta, *Drosophila melanogaster* constituye un ejemplo interesante estudiado en detalle por Bialek y colaboradores donde aplicaron la teoría de la información [26, 104, 106]. La pregunta motivadora es cómo distintas partes del embrión, casi no diferenciadas entre sí, se transforman eventualmente en las distintas partes del animal adulto. Esto sucede porque los FTs involucrados en la regulación de los genes que eventualmente van a dar lugar a la formación de uno u otro tipo de célula no están uniformemente distribuidos en el embrión. La posición está codificada en la concentración de estos FTs. Considerando a la posición como la “entrada” y a la concentración del FT a lo largo del embrión como la “salida” la pregunta que surge es cuál es la información mutua entre ambas cantidades. El ejemplo analizado en [26] es el de la proteína Hunchback (Hb), un FT que se produce a tasas no uniformes a lo largo del eje anteroposterior del embrión durante las etapas tempranas de desarrollo. Bialek y colaboradores observaron muchos embriones y midieron la concentración de Hb como función de la posición relativa (respecto del largo total) en cada uno (ver Fig. 1.22). Si existiera una relación funcional perfecta sin ruido, entonces la codificación de la información posicional en [Hb] estaría limitada por el número de muestras tomadas a lo largo de la posición. Asociando [Hb] al nivel de expresión, g , del gen que codifica a la proteína Hb es posible estimar la probabilidad condicional $P(g|x)$ a partir de los datos experimentales. Considerando que todas las posiciones, x , a lo largo del embrión son igualmente probables, se puede calcular $P(g)$ a partir de $P(g|x)$ y luego, la información mutua entre x y g . Procediendo de esta manera, los autores de [26] determinaron que el nivel de expresión de Hb proporciona $2,26 \pm 0,04$ bits de información sobre la posición a lo largo del eje anteroposterior del embrión, un número superior al supuesto *a priori* (~ 1 bit) dada la distribución tipo escalón de [Hb] a lo largo del embrión .

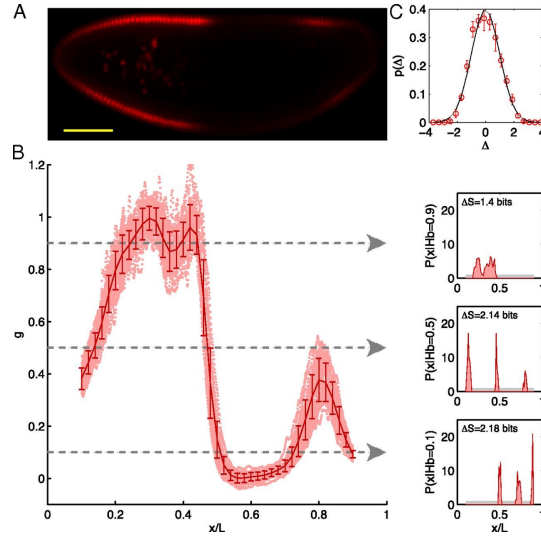


Figura 1.22: Figura tomada de [26] donde se ilustra la información posicional acarreada por el FT Hunchback (Hb) en embriones de mosca. (A) Plano medio sagital de un embrión de *Drosophila melanogaster* con la proteína Hb marcada fluorescentemente. (B) Perfiles normalizados de intensidad de fluorescencia a lo largo del embrión que se identifican con el nivel de expresión, g , de Hb. Los datos provienen de 24 embriones y la posición a lo largo del eje está normalizada por la longitud L del embrión. Las medias, $\bar{g}(x)$, y los desvíos, $\sigma_g(x)$, para cada posición, x , se representan en rojo más oscuro. Considerando los puntos con $g = 0, 1, 0,5$ o $0,9$ (izquierda), se obtienen las distribuciones condicionales de la derecha, $P(x|g)$. Se indica también la reducción de la incertidumbre en x ($\Delta S = S[P_x(x)] - S[P(x|g)]$) que se obtiene “leyendo” cada uno de estos valores de g . (C) Variaciones en el nivel de expresión alrededor de la media en cada posición, estimadas por la distribución de la expresión relativa normalizada, dada por $\Delta = (g - \bar{g}(x))/\sigma_g(x)$ (círculos rojos con SEM). La línea continua es una Gaussiana de media cero/varianza unitaria.

Bialek y colaboradores también se preguntaron sobre la capacidad de un ser vivo de maximizar la información mutua, $MI(i, o)$, entre una entrada externa de distribución $p_i(i)$ sobre la que el ser vivo no tiene control, y una respuesta, o , cuya densidad de probabilidad condicional, $p(o|i)$, puede ser seleccionada [104, 106]:

$$MI(i, o) = \int di p_i(i) \int do p(o|i) \log \left(\frac{p(o|i)}{\int di p(o|i)p_i(i)} \right) \quad (1.19)$$

El objetivo planteado es encontrar la distribución de probabilidad $p(o|i)$ que maximiza MI dejando fijo $p(i)$. Para facilitar los cálculos y teniendo en cuenta observaciones experimentales como las de la Fig. 1.22 que muestran que la variable estocástica $o(i)$ puede aproximarse por la suma de su valor medio, $\bar{o}(i)$, y un ruido Gaussiano de varianza, $\sigma_o^2(i)$, Bialek *et al* aproximan la probabilidad

condicional por:

$$p(o|i) = \frac{1}{\sqrt{2\pi\sigma_o^2(i)}} e^{-\frac{(o(i)-\bar{o}(i))^2}{2\sigma_o^2(i)}} \quad (1.20)$$

De esta manera, dando un valor del ruido $\sigma_o^2(i)$ en la respuesta, o , como función del estímulo, i , lo único que queda a determinar es la función $\bar{o}(i)$, que es precisamente la curva estímulo respuesta. El sistema tiene una solución analítica bajo la aproximación de ruido pequeño, mientras que para ruido más grande, la solución se obtiene numéricamente. De esta manera, la pregunta pasa a ser cuál es la curva estímulo respuesta que optimiza el flujo de información dada una distribución de probabilidad del estímulo, $p(i)$, fija.

La solución analítica para ruido pequeño es tal que la probabilidad de la respuesta no condicionada, $p_o(o) = \int p(o|i)p_i(i) di$, satisface $p_o(\bar{o}) = \frac{1}{Z} \frac{1}{\sigma_o(\bar{o})}$ (con Z una constante de normalización). En otras palabras, la densidad de probabilidad es mayor para valores de o en los cuales el sistema tiene ruido más bajo. Si el ruido fuera homogéneo, el resultado es un fenómeno que se llama *whitening*, que significa que la respuesta es equiprobable en un intervalo dado entre una respuesta mínima y una respuesta máxima. Si el ruido es pequeño, $p_o(\bar{o}) \sim p_o(o)$, con lo cual, sabiendo que $p_o(\bar{o}) = \frac{d\bar{o}}{di} p_i(i)$, el resultado es que $\frac{d\bar{o}}{di} = Z\sigma_o(i)p_i(i)$. Lo que esto significa es que dada la distribución de probabilidad de $p_i(i)$ y el valor del ruido $\sigma_o(i)$, queda determinada la curva estímulo-respuesta que maximiza $MI(i, o)$.

En [104] se analizaron los datos experimentales de distribución espacial de dos proteínas, Hb y Bicoid (Bcd), en embriones de mosca en etapas tempranas de su desarrollo. Si bien ambas proteínas son FTs, Bcd es la que controla la expresión de Hb. Por lo tanto, los autores consideraron que la concentración de Bcd era la entrada, i , y la de Hb, la salida, o . A partir de los datos experimentales pudieron determinar las varias densidades de probabilidad como se explicó anteriormente y la relación estímulo respuesta entre Bcd y Hb. La comparación de esta última con la que, según la teoría maximiza MI, da un muy buen acuerdo como se muestra en la Fig. 1.23.

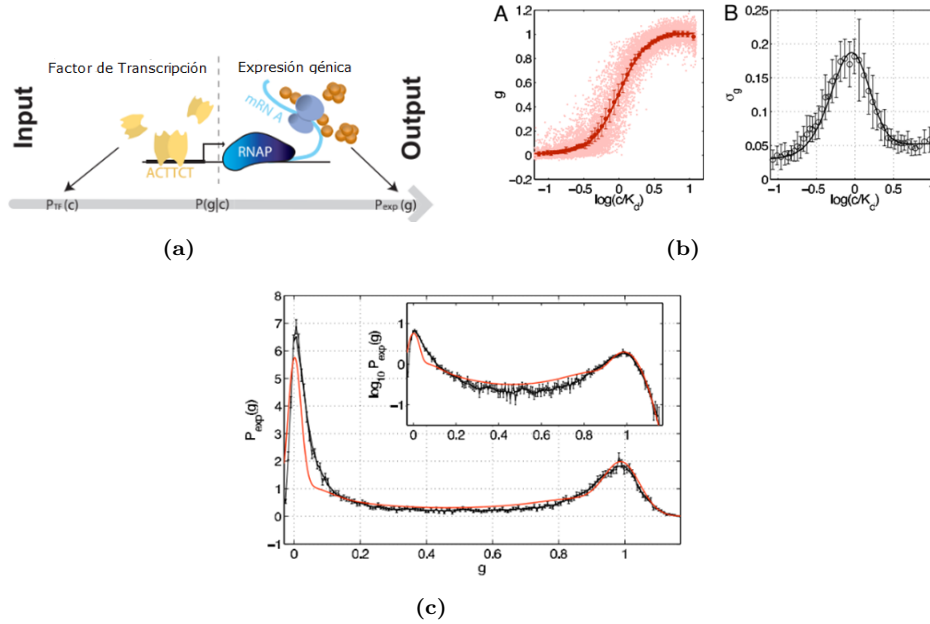


Figura 1.23: (a) El modelo analizado en [106]. El FT es la entrada, i , y tiene una distribución de probabilidad, $P_{FT}(c)$ ($p_i(i)$ en la notación del cuerpo principal de la Tesis). La expresión génica, g , es la salida, o , y tiene una distribución, $P_{exp}(g)$ ($p_o(o)$ en la notación del cuerpo principal de la Tesis). $P(g|c)$ es la probabilidad condicional ($p(o|i)$). (b) Comparación entre la curva estímulo-respuesta ($\bar{o}(i)$ que maximiza $MI(i, o)$ según la teoría y los resultados experimentales (de concentración de Hb vs concentración de Bcd). Se observa un muy buen acuerdo. (c) En este gráfico se muestra que la distribución de probabilidad del output coincide con la predicha teóricamente. Tomado de Tkacik *et al* [106].

Se destacan dos aspectos clave de esta teoría. En primer lugar, que la teoría es no paramétrica, por lo que la curva estímulo respuesta queda trazada con 0 grados de libertad. Por otro lado, también sucede que conociendo la curva estímulo-respuesta se puede conocer la distribución de probabilidad del estímulo conociendo las propiedades del ser vivo. De esta manera, investigando al ser vivo, se puede caracterizar el entorno, como si el ser vivo fuera un espejo de su entorno.

1.7.4. Teoría de la información aplicada al análisis de la multiplexación en levaduras.

Con posterioridad a los experimentos descritos en la Sec. 1.5, Hansen & O'Shea [47] usaron la teoría de la información para cuantificar cuánta información podía transducir el FT Msn2 en células de levadura al inducir los genes involucrados en la generación de la respuesta a estrés osmótico (tomando como ejemplo al gen SIP18) y a la privación de glucosa (tomando como ejemplo al

gen HXK1). De acuerdo a la Fig. 1.17 las vías en las que participan estos genes codifican la intensidad del estímulo externo en duración y frecuencia, respectivamente. En [47], sin embargo, los autores analizaron la capacidad de codificar por amplitud y por frecuencia para ambas vías. Más específicamente, analizaron la capacidad del canal, es decir, la información máxima posible de ser transmitida al variar la distribución de entradas. Para poder calcular la información a partir de datos experimentales debieron distinguir el ruido intrínseco (variabilidad dentro de una célula) del extrínseco (variabilidad entre células), ya que la información transmitida en cada célula depende del ruido intrínseco. Para tal fin usaron un método basado en el uso de dos reporteros transcripcionales fluorescentes [33]. Realizaron los experimentos para células silvestres y para células con el gen SIP18 (el que responde en la vía que codifica por amplitud/duración) mutado de dos formas posibles (mutA y mutB).

Los resultados obtenidos en células silvestres mostraron que, ante la modulación por amplitud, la información transmitida era de 1,2-1,3 bits para los dos genes analizados. Para la modulación por frecuencia, por otro lado, la información transmitida por el HXK1 (el que responde en la vía que codifica por frecuencia) era $\sim 1,1$ bits mientras que la transmitida por el SIP18 era despreciable. En el caso de los dos tipos de mutantes la información transmitida por amplitud (1,42 y 1,55 bits para mutA y mutB, respectivamente) era mayor que para las que transmiten las células silvestres. Esto llevó a los autores a concluir que el gen SIP18 silvestre no está optimizado para transmitir máxima información a través de la modulación por amplitud, a pesar de estar involucrado en la vía que “usa” esta codificación. Finalmente, se encontró también que en las cepas mutadas, la información modulada por frecuencia que transmite el gen SIP18 (0,88 y 1,39 bits) era mayor que en el caso silvestre. La Fig. 1.24 ilustra estos resultados.

Los resultados de la Fig. 1.24 llevaron a los autores de [47] a preguntarse por qué los genes de las células silvestres analizados funcionaban como “switches ruidosos” (es decir, transmitían solo un poco más de 1bit de información) cuando en realidad con pequeñas modificaciones podían mejorar dicha transmisión (como lo muestran los experimentos realizados con mutantes). ¿Por qué no estaban optimizados? El trabajo desarrollado en el Capítulo 3, publicado en [39], provee respuestas posibles a estas preguntas.

1.8. Organización de la Tesis y anticipo de los resultados principales.

En esta Introducción se intentó mostrar la diversidad y complejidad de los comportamientos mediante los cuales las células evalúan las condiciones del medio que las rodea para generar las respuestas necesarias. Mostramos ejemplos de situaciones en las que intensidades crecientes de un estímulo externo se traducen en concentraciones crecientes de intermediarios de la señalización (codificación por amplitud) incluyendo la de los FTs que modulan la expresión génica. Mos-

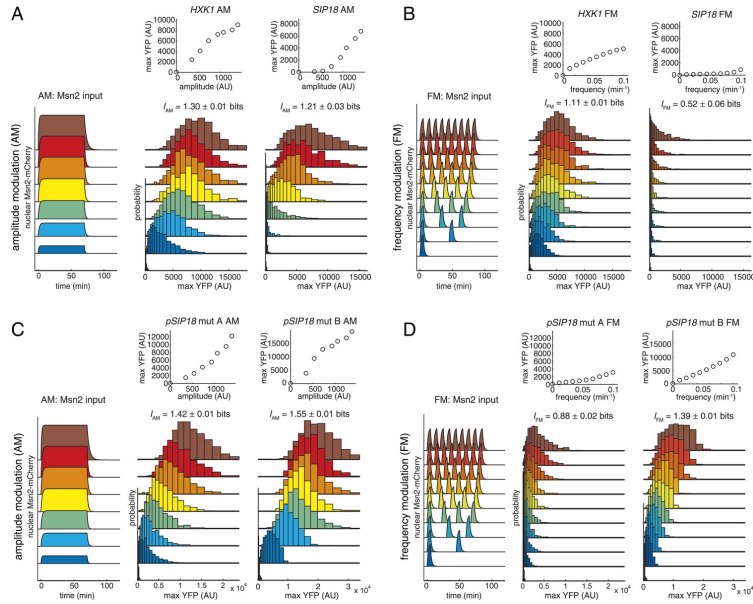


Figura 1.24: Figura tomada de [47] donde se ilustra la transducción de información mediada por un mismo FT (Msn2) sobre dos de los genes que puede inducir, HXK1 y SIP18, en células de *S. cerevisiae*. En células silvestres, HXK1 participa de la vía donde el estímulo externo se codifica en la frecuencia de los pulsos de FT nuclear mientras que SIP18 lo hace en la que codifica por amplitud. Los experimentos manipulan directamente la fracción nuclear del FT Msn2. (A) Se muestran los resultados obtenidos al exponer las células silvestres a ninguna activación o a un pulso de 70 minutos de siete amplitudes distintas. En particular, se muestra la distribución de la expresión génica para ambos genes como un histograma. La dosis-respuesta promediada sobre toda la población (arriba) se obtiene calculando la media del histograma para cada condición de entrada de Msn2. Se muestra también el valor calculado de la información mutua, I , entre la expresión génica y la amplitud del FT. (B) Se muestran resultados similares a los de A pero obtenidos al exponer las células silvestres a ninguna activación o a entre uno y nueve pulsos de localización nuclear del Msn2 de 5 minutos de duración cada uno. Como se ve en la figura, la frecuencia del Msn2 aumenta con el número de pulsos. Se muestra también el valor calculado de la información mutua, I , entre la expresión génica y la frecuencia del FT. En (C) y (D) se siguen los mismos protocolos que en (A) y en (B), respectivamente, pero usando las células mutantes, mutA y mutB.

tramos otros ejemplos en los que estímulos constantes inducen comportamientos pulsátiles en los intermediarios de la señalización con frecuencias crecientes con la intensidad del estímulo (codificación por frecuencia). Mostramos también ejemplos de células que “usan” las distintas dinámicas de traslocación nuclear que un mismo FT puede exhibir dependiendo del estímulo externo, generando, a su vez respuestas diferenciadas dependiendo de dicha dinámica (multiplexación). Finalmente, describimos ejemplos de estímulos que inducen respuestas que codifican en amplitud y, simultáneamente, otras que codifican en frecuen-

cia, en el mismo tipo celular. ¿Cuáles son las similitudes y diferencias entre estos distintos tipos de “evaluar” el entorno y generar respuestas? ¿Cuáles son las ventajas relativas de un modo de codificación sobre el otro? ¿Qué propiedades caracterizan a la codificación por frecuencia la que es usada en muchas situaciones a pesar de que requiere una mayor inversión energética? Estas son algunas de las preguntas motivadoras del trabajo contenido en esta Tesis, las que buscamos responder apelando a herramientas de la física, el modelado y la teoría de la información, varias de las cuales describimos en esta Introducción.

La Tesis está organizada como se describe a continuación. En el Capítulo 2 nos concentramos en la codificación de la intensidad de un estímulo externo en el tiempo medio entre pulsos de un intermediario pulsátil. Para ello combinamos las observaciones experimentales sobre pulsos de Ca^{2+} que llevan a las Ecs. (1.3) y (1.5) con el modelo introducido en [90] dado por la Ec. (1.8). Usando el modelo combinado, analizamos cómo la información mutua (MI) entre la concentración del ligando externo y alguna propiedad de la secuencia de pulsos varía al variar la relación entre la tasa de disparo, λ , y la de recuperación, ρ . Encontramos que esta última es fundamental para determinar la máxima información que puede obtenerse del ligando externo a partir de la secuencia de pulsos. Los resultados de este Capítulo fueron publicados en [42]. En los Capítulos 3 y 4 nos enfocamos en la transcripción para estudiar las similitudes y diferencias existentes entre las distintas formas de codificar la información en alguna propiedad de la fracción nuclear del FT involucrado. En ambos casos usamos el modelo sencillo dado por (1.13) del cual presentamos una derivación mecánica en el Apéndice B. Mientras que en el Capítulo 3 solo miramos el paso que va desde el FT hasta el ARNm producido, en el Capítulo 4 incorporamos la transducción desde el estímulo externo al FT suponiendo que sigue una ley del estilo de la Ec. (1.1) en el caso de codificación por amplitud y una basada en las Ecs. (1.3) y (1.5) en el caso de codificación por frecuencia. Más específicamente en el Capítulo 3 calculamos la MI entre la cantidad de ARNm producido durante un dado intervalo y la duración, amplitud o frecuencia media entre pulsos de la fracción nuclear del FT. Este análisis muestra que la principal diferencia entre los tres esquemas de codificación analizados parece residir en su diferente sensibilidad a cambios en los parámetros que caracterizan la transcripción. Debido a esta sensibilidad diferente, es posible encontrar promotores que son relativamente ciegos a las entradas moduladas en frecuencia pero que actúan como interruptores ruidosos para las moduladas en duración o amplitud, mientras que no es posible encontrar la situación opuesta. Los resultados de este Capítulo que establecen límites a la multiplexación de la transmisión de información y permiten explicar, entre otras cosas, las observaciones de la Fig. 1.24 fueron publicados en [39]. En el Capítulo 4 calculamos la MI entre la intensidad del estímulo externo y el ARNm producido cuando la primera se codifica en la amplitud o frecuencia medias de los pulsos del FT nuclear. Este estudio muestra que la forma diferente en que los dos modos de codificación mapean los estímulos externos da como resultado capacidades de transmisión de información cualitativamente diferentes: mientras que la codificación de amplitud es óptima para un rango limitado de intensidades alrededor de un valor determinado, la transmisión de información

para la codificación por frecuencia mejora o permanece relativamente constante a medida que aumenta la intensidad del estímulo. Estos resultados que se encuentran contenidos en el trabajo [40] podrían explicar la “utilidad” del uso de distintas estrategias para codificar un mismo estímulo discutido en la Sec. 1.6. La dependencia exponencial entre el tiempo medio entre pulsos y los estímulos externos, tomada de observaciones experimentales de pulsos intracelulares de Ca^{2+} (ver *e.g.* la Fig. 1.10), es fundamental para determinar el comportamiento cualitativamente distinto de las codificaciones por frecuencia y por amplitud. En el Capítulo 5 mostramos cómo esta dependencia puede ser obtenida a partir del modelo simplificado para la dinámica del Ca^{2+} intracelular dado por las Ecs. (1.6)–(1.7) para parámetros en los que esta dinámica es excitable. El abordaje seguido en este estudio que está contenido en [41] puede ser aplicado a otros sistemas excitables, en particular, a los que rigen la dinámica de traslocación nuclear de algunos FTs. Finalmente, el Capítulo 6 contiene un resumen con las conclusiones principales de la Tesis mientras que los Apéndices contienen algunos cálculos suplementarios.

Capítulo 2

Contenido de información en secuencias de pulsos estocásticos de mensajeros intracelulares.

En este Capítulo estudiamos algunas propiedades de la codificación de la intensidad de un estímulo externo en el tiempo medio entre pulsos de un intermediario pulsátil de la señalización. Para ello combinamos las observaciones experimentales sobre pulsos de Ca^{2+} que llevan a las Ecs. (1.3) y (1.5) con el modelo introducido en [90] dado por la Ec. (1.8). Usando este modelo combinado, analizamos cómo la información mutua entre la concentración del ligando externo y alguna propiedad de la secuencia de pulsos varía al variar la relación entre la tasa de disparo, λ , y la de recuperación, ρ . En particular, analizamos dos límites, aquél en que $\lambda/\rho \gg 1$ que corresponde a que los pulsos sigan una estadística de Poisson y el otro en el que $\lambda/\rho \ll 1$ en donde prepondera el *Feedback Negativo Global* (GNF). El trabajo contenido en este Capítulo, que fue publicado en [42], se enfoca entonces en determinar cuánta información en exceso se obtiene con la estadística del GNF respecto de la de Poisson.

2.1. Métodos

2.1.1. El modelo

Consideramos el modelo introducido en [90] en el que se supone que el tiempo entre pulsos sucesivos de Ca^{2+} es $t_{stoch} + T_{min}$ con T_{min} una componente determinista y $t \equiv t_{stoch}$ es una variable aleatoria (en esta sección se usará “t”). En particular, suponemos que, dado que un pulso ocurrió al tiempo 0, la densidad de probabilidad de que el primer pulso subsiguiente ocurra al tiempo,

$t + T_{min}$, está dada por la Ec. (1.8).

Estudiamos el modelo en el límite de *Feedback Negativo Global* (GNF) definido por $x \equiv \lambda/\rho \gg 1$ y en el de Poisson ($x \ll 1$). En este último límite la distribución de t es exponencial con media $1/\lambda$. Por otro lado, en base a los experimentos de [103] (ver Ec. (1.5)), suponemos que:

$$T \equiv \langle t \rangle|_{t|\lambda} = A \exp(-\gamma C), \quad (2.1)$$

con C la concentración del efector externo, $\langle \cdot \rangle|_{t|\lambda}$ la media sobre la distribución para un dado valor de λ , $p(t|\lambda)$ y A y B constantes. Como demostraremos más adelante, la Ec. (2.1) implica que λ y C están relacionados por $\lambda = \alpha \exp(\beta C)$ con $\alpha = \pi/(2\rho A^2)$ y $\beta = 2\gamma$ para el límite GNF y $\alpha = 1/A$ y $\beta = \gamma$ para el de Poisson. Las densidades de probabilidad de λ , T y C satisfacen entonces:

$$p_\lambda(\lambda) = \frac{1}{\lambda\beta} p_C \left(\frac{1}{\beta} \ln \left(\frac{\lambda}{\alpha} \right) \right), \quad p_T(T) = \frac{1}{BT} p_C \left(\frac{1}{B} \ln \left(\frac{A}{T} \right) \right). \quad (2.2)$$

En ambos límites, T , y la desviación estándar de t , σ_{stoch} , satisfacen [90], como se muestra en la Ec. (1.4), es $\sigma_{stoch} = kT$, con $k = 1$ en el caso de Poisson. Obtenemos $k = \sqrt{4/\pi - 1}$ en el límite GNF, como mostraremos después.

2.1.2. Cálculos

Después de derivar algunas propiedades del modelo en los límites de GNF y Poisson, calculamos, en primer lugar, la información mutua (ver Sec. 1.7.1) entre la concentración del efector externo, C , y (la parte estocástica de) el tiempo entre pulsos, t , que en este Capítulo llamamos $I(C, t)$. Esta información está dada por:

$$\begin{aligned} I(C, t) = I(\lambda, t) &= \iint p_\lambda(\lambda) p(t|\lambda) \log_2(p(t|\lambda)/p(t)) dt d\lambda \\ &= \int p_\lambda(\lambda) \int p(t|\lambda) \log_2(p(t|\lambda)) dt d\lambda - \int p_t(t) \log_2(p_t(t)) dt \\ &\equiv -H(t|\lambda) + H(t), \end{aligned} \quad (2.3)$$

donde

$$p_t(t) = \int p_{\{t,\lambda\}}(t, \lambda) d\lambda = \int p(t|\lambda) p_\lambda(\lambda) d\lambda, \quad (2.4)$$

con $p_{\{t,\lambda\}}$ la densidad de probabilidad conjunta de t y λ . Obtenemos una expresión general y luego suponemos que C se distribuye uniformemente en $[0, C_M]$:

$$p_C(C) = \begin{cases} \frac{1}{C_M}, & \text{si } 0 \leq C \leq C_M, \\ 0, & \text{sino.} \end{cases} \quad (2.5)$$

En segundo lugar procedemos a calcular la información mutua entre C y el número de pulsos, N , que ocurren durante un tiempo $t_{tot} \gg T$. Esta información satisface $I(C, N) = I(T, N)$, con $N \gg 1$. Para el cálculo eliminamos la parte

determinista del tiempo entre pulsos, T_{min} y consideramos que $p_C(C)$ está dada por la Ec. (2.5).

Más adelante presentamos los cálculos analíticos que se realizan bajo la suposición de que C , A , B y ρ son constantes. Estos cálculos son válidos siempre que las variaciones en C sean mucho más lentas que el tiempo medio, T .

2.1.3. Simulaciones Numéricas

Para verificar la validez de la $I(N, C)$ determinada analíticamente, realizamos simulaciones numéricas estocásticas para $A = 1$, $B = 1$ y diferentes valores de $t_{tot} \leq 1000$ usando el lenguaje de programación R. Se realizaron 500 realizaciones fijando $T = e^{-C}$ para cada uno de los 100 valores de C que se eligieron aleatoriamente con distribución uniforme en $[0, 1]$. Usamos $\rho = 0,01$ y $\lambda = 50\pi \exp(C)$ para el límite GNF y $\rho = 0$ y $\lambda = \exp(C)$ para el de Poisson. Así, $1/e \leq T \leq 1$ en ambos tipos de simulaciones. Para cada realización discretizamos el tiempo con paso de tiempo, $dt = 0,01$, y decidimos con probabilidad λdt que un pulso ocurra en cada paso de tiempo en el límite de Poisson y con probabilidad $\lambda(1 - \exp(-\rho(t - t_{prev})))dt$ con t_{prev} el momento en que se produjo el pulso anterior en el límite GNF. Contamos el número total de pulsos, N , para cada realización y agrupamos los resultados en bins de tamaño dependiente de t_{tot} . Utilizamos una biblioteca integrada de R para calcular la información mutua entre los valores de N obtenidos y los de C .

2.2. Resultados

Aquí comparamos los resultados sobre información mutua obtenidos para el modelo descrito en la Sec. 2.1.1 en los límites de Poisson ($x = \lambda/\rho \ll 1$) y GNF ($x \gg 1$).

2.2.1. Información contenida en el tiempo entre pulsos.

Como se muestra en A.2, la información mutua entre la concentración del efector externo, C , y el tiempo entre pulsos o, equivalentemente, su parte estocástica, t , es:

$$MI(C, t) = MI(\lambda, t) = \frac{1}{\ln(2)} \left(- \int_{-\infty}^0 \frac{\partial M_{\tilde{\lambda}}(\tilde{\tau})}{\partial \tilde{\tau}} \ln \left(\frac{\partial M_{\tilde{\lambda}}(\tilde{\tau})}{\partial \tilde{\tau}} \right) d\tilde{\tau} - 1 + \langle \ln(\tilde{\lambda}) \rangle |_{\lambda} \right), \quad (2.6)$$

con $\tilde{\lambda} = \lambda/\alpha$, $M_{\tilde{\lambda}}(\tilde{\tau})$ dada por:

$$M_{\tilde{\lambda}}(\tilde{\tau}) = \int e^{\tilde{\lambda}\tilde{\tau}} p_{\tilde{\lambda}}(\tilde{\lambda}) d\tilde{\lambda} = \int e^{\tilde{\tau} \exp(\beta C)} p_C(C) dC, \quad (2.7)$$

y

$$\langle \ln(\tilde{\lambda}) \rangle |_{\lambda} = \int \ln \left(\frac{\lambda}{\alpha} \right) p_{\lambda}(\lambda) d\lambda = \int \beta C p_C(C) dC = \beta \langle C \rangle. \quad (2.8)$$

Las Ecs. (2.6)–(2.8) implican que $I(C, t)$ es independiente de α . El hecho de que su forma sea similar independientemente del límite (Poisson o GNF) no significa que la información sea la misma en ambos casos: dados T y C , el valor, λ , es diferente dependiendo del límite (ver Ec. (A.2)). El hecho de que I dependa de β pero no de α y que $\beta = B$ en el límite de Poisson y $\beta = 2B$ en el GNF implica que simplemente multiplicando β por 2 en $I(C, t)$, pasamos de un límite al otro para el mismo valor de B . Suponiendo que p_C viene dado por la Ec. (2.5) obtenemos:

$$I(C, t) = \frac{1}{\ln(2)} \left(\frac{\sinh(\beta C_M)}{\beta C_M} + \ln \left(\beta \frac{C_M}{2} \right) - \ln \left(\frac{e^{\beta C_M} - 1}{2} \right) - 1 + \beta \frac{C_M}{2} \right) - \frac{1}{\beta C_M \ln(2)} \int_{-\infty}^0 \frac{e^\zeta}{\zeta} \ln \left(\frac{e^{\beta C_M} \sinh(\zeta e^{-\beta C_M} \frac{e^{\beta C_M} - 1}{2})}{\sinh(\zeta \frac{e^{\beta C_M} - 1}{2})} \right) d\zeta. \quad (2.9)$$

Mostramos en la Fig. 2.1 esta información mutua como función de βC_M para los dos límites del modelo. Allí observamos que siempre es mayor en el límite GNF y que la diferencia aumenta con βC_M acercándose a ~ 1 bit.

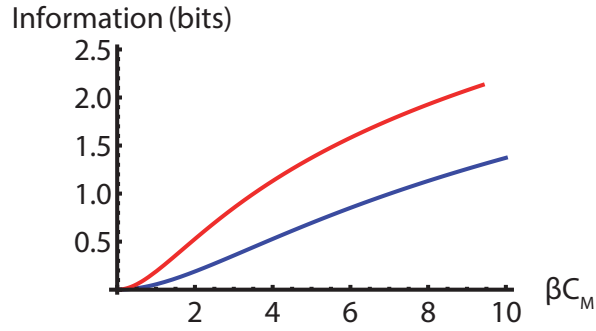


Figura 2.1: Información mutua, $I(C, t)$ en la Ec. (2.9), entre la concentración del efector, C , y la parte estocástica del tiempo entre pulsos, t , en función de C_M para los límites de Poisson (azul) y el GNF (rojo) del modelo con C distribuido uniformemente sobre $[0, C_M]$.

2.2.2. Información contenida en el número de pulsos, N , que ocurren durante un intervalo.

Los cálculos mostrados en el Apéndice A, para la información entre C y N para un tiempo de observación dado, t_{tot} , $N \gg 1$ y p_C dada por la Ec. (2.5) resultan en:

$$I(C, N) = I(N, T) = \frac{3}{2} H(T) - 2 \langle \log_2(T) \rangle_T - \log_2(2\pi e C_M B) + \frac{1}{2} \log_2 \left(\frac{t_{tot}}{k^2} \right) + \mathcal{O}(t_{tot}^{-1/3}) \quad (2.10)$$

con H la entropía como en la Ec. (2.3) y $k = 1$ ($k = \sqrt{4/\pi - 1}$) para el límite de Poisson (GNF). En este caso la única diferencia entre los límites de Poisson y GNF radica en la constante de proporcionalidad, k , entre el desvío estándar, σ y la media, T , de la (parte estocástica del) tiempo entre pulsos. Verificamos el cálculo analítico vía simulaciones numéricas realizadas como se describe en los Métodos. Mostramos los resultados en la Fig. 2.2. Allí observamos que, también en este caso, I es mayor para el GNF que para el límite de Poisson y que la diferencia aumenta muy lentamente con t_{tot} acercándose al valor asintótico prescrito por la teoría, $\Delta I = -\log_2(4/\pi - 1)/2 \approx 0,936$. Las curvas de ajuste confirman esto y la dependencia de I con t_{tot} , en particular, el término $\mathcal{O}(t^{-1/3})$ que atribuimos a la asimetría de la densidad de probabilidad de N que no se incluye al usar la aproximación normal.

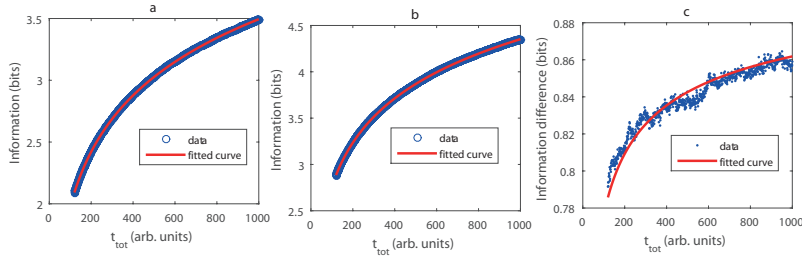


Figura 2.2: Información mutua, $I(N, C|t_{tot})$, en función de t_{tot} obtenida numéricamente con p_C dada por la Ec. (2.5), $A = 1$, $B = 1$ y $C_M = 1$ (ver Métodos). (a) Límite de Poisson. Los círculos son el resultado de la simulación. La curva de ajuste (sólida) de la forma $I = \log_2(t_{tot})/2 + a + bt_{tot}^{-1/3}$ dio $a = (-1,6111 \pm 0,0011)$ y $b = (1,2275 \pm 0,0086)$. (b) Similar a (a) pero para el límite GNF con $\rho = 0,01$. El ajuste dio $a = (-0,69121 \pm 0,00076)$ y $b = (0,6061 \pm 0,0058)$. (c) Diferencia de información entre ambos límites. En este caso la curva de ajuste es $\Delta I = -\log_2(\sqrt{4/\pi - 1}) + bt_{tot}^{-1/3}$ con $b = (-0,741 \pm 0,022)$.

2.3. Discusión y conclusiones del Capítulo.

La comparación de la información que proporciona el tiempo entre pulsos (o el número de pulsos) generados acerca de la concentración del efector externo, C , en los límites de GNF y Poisson mostró que, para $0 \leq C \leq C_M$ distribuida uniformemente, la información mutua, $I(t, C)$ y $I(N, C)$, fue mayor en el primero que en el segundo límite, que la diferencia fue como máximo $\sim 1bit$, resultando menor a 0,5 para los valores de los parámetros explorados experimentalmente en [103] ($\beta C_M \sim 1,6$). Esta comparación difiere del cálculo de la entropía de Kullback presentado en [90], que representa la cantidad de información que se pierde cuando se usa una probabilidad (por ejemplo, Poisson) para describir una observación en lugar de usar la distribución verdadera (por ejemplo, GNF). Tener un efecto inhibitorio global que actúa al ocurrir el pulso

(que se recupera a una tasa, ρ) es fundamental para la robustez funcional de la codificación de señalización en secuencias de pulsos estocásticos [102]. Es responsable del proceso descrito por la Ec. 1.8 y Ec. 1.4 con el mismo k para células individuales del mismo tipo. Aunque ρ determina la máxima información que se extrae del entorno, el incremento de información de a lo sumo $\sim 1bit$ que hemos encontrado respecto del proceso puramente de Poisson significa que la precisión con la que se infiere la concentración efectora puede, como máximo, duplicarse para procesos en los que la recuperación es suficientemente lenta ($\rho \ll \lambda$). La Ec. 2.6 también implica que, si fijamos λ y variamos ρ , no hay ganancia en la información mutua entre t y C . De esta manera, regular únicamente la tasa de recuperación de las células podría cambiar la frecuencia media entre pulsos sin cambiar la información que se puede obtener del ambiente.

Capítulo 3

La modulación de la dinámica de los Factores de Transcripción permite una transmisión versátil de la información.

En éste y en el próximo Capítulo nos enfocamos en la transcripción para estudiar las similitudes y diferencias existentes entre las distintas formas comúnmente utilizadas por las células para codificar la información en alguna propiedad de la fracción nuclear de los FTs. En ambos casos usamos el modelo transcripcional sencillo dado por (1.13) parte del cual modelamos mediante un proceso de Markov. En este Capítulo solo miramos el paso que va desde el FT hasta el ARNm producido. Más específicamente, calculamos la MI entre la cantidad de ARNm producido durante un dado intervalo y la duración, amplitud o frecuencia media entre pulsos de la fracción nuclear del FT. Está implícita en el abordaje la suposición de que el estímulo externo está codificado en la duración, amplitud o frecuencia media del FT nuclear de un modo que no exploramos en este Capítulo. Buscamos luego los parámetros del modelo de transcripción que maximizan MI para cada tipo de codificación o *modulación*. Para determinar la posibilidad de multiplexar la información usando un mismo FT y dos formas de codificación (ver Sec. 1.7.4) exploramos en cuánto puede diferir la información transmitida de uno u otro modo a través de dos promotores distintos. Los resultados de este Capítulo, que establecen límites a la multiplexación y permiten explicar, entre otras cosas, las observaciones de la Fig. 1.24, fueron publicados en [39].

3.1. Métodos

3.1.1. Modelo, entrada y salida

Para analizar cómo la información mutua, MI, entre el ARNm producido y la fracción nuclear del FT depende de si la intensidad del estímulo se codifica en la amplitud, frecuencia o duración de un estado de activación del FT, realizamos simulaciones del modelo ilustrado en la Fig. 3.1 donde el paso transcripcional corresponde al modelo sencillo dado por (1.13) del que presentamos una derivación mecanística en el Apéndice B. Acá la entrada corresponde a alguna propiedad de la fracción nuclear del FT y la salida al ARNm acumulado durante un cierto tiempo. Más específicamente, suponemos que un aumento en la intensidad del estímulo externo puede traducirse de tres formas posibles: en mayores concentraciones de la fracción nuclear del FT (modulación por amplitud); en tiempos de permanencia más largos de una concentración de FT nuclear relativamente alta (modulación por duración), o en una mayor frecuencia de pulsos de FT nuclear (modulación por frecuencia). Para la salida consideramos tres opciones posibles para el tiempo durante el cual se calcula el ARNm acumulado como describimos más adelante y se ilustra también en la Figura.

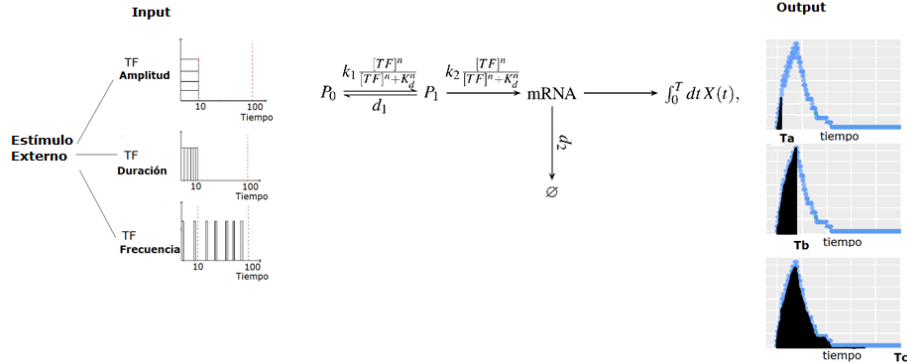


Figura 3.1: El modelo. Un estímulo externo se codifica en la duración o amplitud del pulso o la frecuencia de pulsos de la concentración nuclear de FT, $[FT]$. El modelo no describe este primer paso del proceso de codificación, sino que comienza directamente con el curso temporal del FT nuclear. La información mutua, MI, entre la entrada (duración de pulso, amplitud de pulso o frecuencia media entre pulsos, según el tipo de codificación del estímulo; 3 subfiguras a la izquierda) y la salida (la integral temporal del número de moléculas de ARNm producidas, X) se calcula eligiendo tres posibles cortes de tiempo para su cálculo (3 subfiguras a la derecha que corresponden a un ejemplo con un solo pulso de FT de amplitud, 100, y 10 minutos de duración). Entre la salida y la entrada está el modelo de transcripción [46] descrito con más detalle en el texto principal.

En el esquema de la Fig. 3.1, FT es la *fracción nuclear* del FT que se

supone varía en el tiempo con una dinámica conocida (constante de a tramos en el tiempo). Como se explica más adelante con más detalle, las unidades de concentración arbitrarias que se utilizan a lo largo del Capítulo son tales que 100 corresponde al valor máximo posible de $[FT]$ (a menos del ruido aditivo usado en las simulaciones). $P_0(t)$ y $P_1(t)$ representan el promotor en su estado inactivo o activo, respectivamente ($P_0 + P_1 = 1$). El promotor es activado por el FT en un modo cooperativo, como se refleja en el término, $[FT]^n / ([FT]^n + K_d^n)$, donde n indica la cooperatividad y K_d representa una constante efectiva de disociación, medida con las mismas unidades arbitrarias que $[FT]$, de la reacción de unión/desunión que se supone que ocurre en una escala de tiempo más rápida que el resto de los procesos y, por tanto, está en equilibrio. En el límite de $[FT]$ alta el promotor se vuelve activo con tasa k_1 . La flecha de P_1 a ARNm representa la transcripción que ocurre a una velocidad, $k_2 [FT]^n / ([FT]^n + K_d^n)$, lo que significa que el FT debe estar ligado al promotor para que se lleve a cabo la transcripción. Más detalles sobre el modelo se pueden encontrar en el Apéndice B. Dado que $0 \leq [FT]^n / ([FT]^n + K_d^n) < 1$, k_1 y k_2 son, respectivamente, las tasas máximas a las que se produce la activación del promotor y la transcripción. El modelo incluye la degradación del ARNm a la velocidad, d_2 , pero no la traducción a la proteína que se supone que se degrada con suficiente lentitud como para que la cantidad de ARNm acumulada se pueda utilizar como indicador del resultado del proceso.

3.1.2. Simulaciones numéricas

En las simulaciones, la transcripción se modela estocásticamente utilizando la ecuación maestra, mientras que el resto de los pasos se modelan de forma determinista, utilizando el método de Euler para resolver las ecuaciones diferenciales ordinarias. La concentración nuclear del FT, $[FT](t)$, se modela mediante un solo pulso o una secuencia de pulsos cuadrados. La amplitud máxima de $[FT]$ considerada es 100 (en unidades adimensionales) si no se tiene en cuenta el ruido aditivo usado en las simulaciones y la duración máxima del pulso es de 10 min. El tiempo total de las simulaciones es de 100min. El valor, $P_1(t)$, derivado de la integración numérica de los dos primeros pasos y $[FT](t)$ se introducen en el paso transcripcional que constituye el núcleo de la aleatoriedad y al que modelamos como un proceso de Markov. Finalmente, el ARNm se integra como función del tiempo para obtener la salida. Dado que $P_0 + P_1 = 1$, las simulaciones implican resolver las siguientes ecuaciones:

$$\begin{aligned} \dot{P}_1 &= \frac{k_1 [FT](t)^n}{K_d^n + [FT](t)^n} - \left(\frac{k_1 [FT](t)^n}{K_d^n + [FT](t)^n} + d_1 \right) P_1, \\ X &= N \xrightarrow{\zeta} X = N + 1, \quad X = N \xrightarrow{d_2 X} X = N - 1 \end{aligned} \quad (3.1)$$

donde $X(t)$ es el número (estocástico) de moléculas de ARNm presentes al tiempo, t , $N = \{0, 1, 2, \dots\}$ y la tasa de producción de ARNm es $\zeta = \frac{k_2 [FT](t)^n}{K_d^n + [FT](t)^n} P_1(t)$ si $\frac{k_2 [FT](t)^n}{K_d^n + [FT](t)^n} P_1(t) < 5/min$ mientras que una nueva molécula de ARNm se pro-

duce con probabilidad 1 cada $0.2min$ en caso contrario. Esta elección corresponde a limitar el número máximo de moléculas de ARNm que se pueden producir por unidad de tiempo. Como argumentamos en la Sec. 3.3, ésta es una forma sencilla de lograr que las fluctuaciones de la salida de nuestro modelo (el ARNm acumulado) sean similares a las de la salida en los experimentos (la proteína producida). En el Apéndice B mostramos los resultados obtenidos si no se impone esta condición pero se limita k_2 de modo que $k_2 \leq 10/min$. Dicho Apéndice contiene también una descripción más detallada del modelo como así también su derivación a partir de una descripción mecanística de la transcripción.

En este Capítulo se estudian tres tipos de modulaciones del FT : duración (identificada con la etiqueta “0”), amplitud (identificada con la etiqueta “1”) y frecuencia (identificada con la etiqueta “2”). En los casos de modulación por duración o amplitud, $[FT](t)$ viene dada, respectivamente, por un solo pulso de amplitud, 100, más ruido y duración variable o de amplitud variable y $10min$ de duración, que comienza en $t = 0$. Despreciando el ruido que se suma a la amplitud de la concentración para todas las modulaciones, $[FT] = 100$ es el valor máximo que puede tomar esta concentración en todos los casos probados (a menos del ruido aditivo usado en las simulaciones). En el caso de la modulación por frecuencia, el $[FT]$ está dado por una secuencia de pulsos cuadrados de $1min$ de duración y amplitud media 100, separados por un tiempo entre pulsos estocástico distribuido exponencialmente con valor medio elegido como se describe más adelante. En todos los casos, $[FT](t)$ es la norma de la suma de la amplitud del pulso determinista al tiempo t y una variable aleatoria con distribución normal de desvío estándar, 10, que es elegida y sumada a la amplitud determinista cada $0,2min$ para imitar el relativamente largo tiempo de correlación entre fluctuaciones. La salida es:

$$Out(T) = \int_0^T dt X(t), \quad (3.2)$$

donde, en la mayoría de los casos acá ilustrados, T toma uno de tres valores finitos, $T = T_a, T_b$ o T_c . Estos valores son $T_a = 10min$, $T_b = 50min$ y $T_c = 100min$ y $T_a = 1min$, $T_b = 10min$ y $T_c = 100min$ para la modulación por frecuencia y amplitud, respectivamente. Para este último, T_b corresponde al momento en que finaliza el pulso del FT independientemente de la amplitud. Para la modulación de duración, usamos $T_a = t_{end}$, $T_b = 10min$ y $T_c = 100min$, con t_{end} la duración del pulso de $[FT]$. Nótese que el corte de tiempo T_a de este caso es cualitativamente diferente a los otros 8 cortes, ya que no corresponde a un tiempo fijo, sino que está definido por un evento. Esto hace que la sensibilidad a los parámetros del modelo sea cualitativamente diferente en este caso. Con fines ilustrativos, también calculamos, $Out(T)$, y la información mutua correspondiente, como funciones continuas del tiempo, T .

3.1.3. Cómputo de la Información Mutua

Discretizando el conjunto de valores que pueden tomar la entrada, I , y la salida, O , ($\{I_i\}_{i=1}^{N_I}$ y $\{O_i\}_{i=1}^{N_O}$, respectivamente), su información mutua se puede

escribir como (ver Ec. (1.18)):

$$\begin{aligned} \text{MI}(I, O) = & - \sum_{i=1}^{N_I} p_I(I_i) \log_2(p_I(I_i)) - \sum_{i=1}^{N_O} p_O(O_i) \log_2(p_O(O_i)) \\ & + \sum_{i=1}^{N_I} \sum_{j=1}^{N_O} p_{IO}(I_i, O_j) \log_2(p_{IO}(I_i, O_j)), \end{aligned} \quad (3.3)$$

donde $p_{I,O}$ es la distribución de probabilidad conjunta de I y O y p_I y p_O son las distribuciones marginales correspondientes. Todas las simulaciones se realizan suponiendo una distribución uniforme de los valores de entrada: sobre el intervalo $[1, 10]min$ para la modulación por duración; sobre el intervalo $[0, 100]$ para la modulación por amplitud; sobre el intervalo $(0, 0,1]/min$ para la frecuencia media en el caso de modulación por frecuencia. La frecuencia más alta considerada da, en promedio, una integral en el tiempo de $[FT](t)$, sobre el intervalo de tiempo de la simulación ($100min$), equivalente a la de los casos de modulación por amplitud o duración ($100 \times 10min$). La distribución uniforme implica que $p_I(I_i) = 1/N_I \forall i$, de modo que el primer término en la parte derecha de la Ec. (3.3) es igual a $\log_2(N_I)$. Todas las simulaciones se realizan discretizando la entrada correspondiente en $N_I = 200$ valores. Dados los parámetros de entrada que se mantienen fijos dependiendo del tipo de modulación (amplitud y/o duración del pulso) y un conjunto de parámetros del modelo (k_1, k_2, n, K_d, d_1 y d_2), se realizaron 15.000 simulaciones para cada uno de los 200 valores de entrada que corresponden al tipo de modulación analizado (amplitud, duración o frecuencia media entre pulsos). Los resultados de estas $3 \cdot 10^6$ simulaciones se utilizan luego para calcular los 3 tipos de cortes de salida descritos anteriormente que solo difieren en el intervalo de tiempo en el que se integra el número de moléculas de ARNm, ver Ec. (3.2). La información mutua entre cada uno de los tres tipos de entrada y cada uno de los tres cortes de salida (9 combinaciones) se calcula, para cada conjunto de valores de parámetros cinéticos, mediante el método Jackknife, que corrige el submuestreo [20, 97]. Para algunos estudios también calculamos MI en función del tiempo, en cuyo caso lo llamamos $\text{MI}(tiempo)$. Calculamos $\text{MI}(tiempo)$ para cada tipo de entrada integrando X en la Ec. (3.2) entre 0 y $tiempo$.

3.1.4. Muestreo del espacio de parámetros, optimización y comparaciones.

Uno de los objetivos del presente estudio es determinar los parámetros del modelo que maximizan MI para cada modulación de entrada y corte de tiempo de salida. Para garantizar un muestreo homogéneo del espacio de parámetros, se utilizó el método Latin Hypersquare Sampling (LHS) [54], dividiendo el rango logarítmico de valores de cada parámetro (k_1, k_2, K_d, n, d_1) en intervalos equiprobables, no superpuestos [89]. Los intervalos muestreados para cada parámetro se muestran en la Tabla (3.1). En total, se probaron 17500 conjuntos de parámetros.

Parámetro	Rango
k_1 (1/min)	(0.01,1)
k_2 (1/min)	(1,100)
n	(1,10)
K_d (a.u.)	(10,100)
d_1 (1/min)	(0.01,1)
d_2 (1/min)	0.12

Tabla 3.1: Intervalos sampleados para cada uno de los parámetros del problema.

Para obtener los resultados del apartado 3.2.4 buscamos el conjunto de parámetros que da el menor valor, MI_m , para un tipo de modulación y corte de tiempo de salida, i ($i = 0, 1, 2$ como se describió anteriormente) restringiendo la búsqueda al conjunto de parámetros que dan valores, MI , que están dentro del 90% del máximo, MI_M , para otro tipo de modulación de entrada ($j \neq i$) y el mismo tiempo de corte. Luego definimos la región *Minmax* de parámetros para el par correspondiente de tipos de entrada, ij , como los conjuntos que dan valores de MI que difieren en menos del 10% de MI_M y MI_m para los tipos de modulación de entrada i y j , respectivamente. También decimos que los conjuntos de parámetros en *Minmax* cumplen la condición *Minmax*.

3.2. Resultados

3.2.1. Información mutua máxima y parámetros del modelo de transcripción.

Aquí mostramos los resultados de la optimización de MI para cada combinación de tipo de modulación de entrada y corte de tiempo de salida cuando los parámetros, d_1 , k_1 , k_2 , K_d y n , son variados mientras que $d_2 = 0,12/min$ permanece fijo (ver Sec. 3.1). La Tabla 3.2 muestra los conjuntos de parámetros que maximizan MI en cada caso y el rango de valores sobre los cuales se realizó la optimización. Observamos que MI_M es $\sim 1 - 2$ bits para todas las combinaciones (con un valor $\sim \log_2(3)$, *i.e.*, tres valores de entrada distinguibles), con la excepción del corte temporal, T_a , y la modulación por duración que da 2.6. Hay que señalar que en este caso la salida se integra solo mientras $[FT](t)$ es diferente de 0 (es decir, mientras el pulso está activo), que es cualitativamente diferente al resto. Para los demás casos, el sistema puede comportarse un poco mejor que un interruptor *binario (ruidoso)* (1 bit). Los parámetros que maximizan MI tienen varias características en común para casi todas las combinaciones de modulación de entrada y corte temporal de salida. Si dejamos de lado el corte, T_a , para la modulación de duración, observamos que $k_2 \gg k_1, d_1, d_2$, lo que implica que la escala de tiempo de la transcripción debe ser tan rápida como sea posible para garantizar una buena transmisión de la información. Segundo, $n \geq 3$ en la mayoría de los casos y la constante de disociación de la reacción de

unión/desunión del FT es $K_d \gtrsim 50$, del mismo orden de magnitud pero menor que la concentración máxima de FT (100). El límite inferior, 50, también es 5 veces mayor que la amplitud del ruido que se considera en el modelo (10). MI alcanza su máximo para $k_1 \gg d_1$, una condición que garantiza que, en estado estacionario, $P_1 \approx [FT]^n / ([FT]^n + K_d^n)$. Por lo tanto, siempre que $k_1 \gg d_1$, la capacidad de distinguir diferentes amplitudes que oscilan entre 0 y 100 dependerá de cómo la función $[FT]^n / ([FT]^n + K_d^n)$ mapee el intervalo $[0, 100]$.

Modulación de Entrada	Corte Salida	k_1 [1/min] (0.01,1)	k_2 [1/min] (1,100)	n [a.u.] (1,10)	K_d [a.u.] (10,100)	d_1 [1/min] (0.01,1)	MI_M [bits]
Duración	T_a	0.32	91	1	10	0.056	2.61
Duración	T_b	0.6	84	7.3	79	0.017	1.61
Duración	T_c	0.13	63	8.3	94	0.016	1.52
Amplitud	T_a	0.62	92	9.5	49	0.024	1.65
Amplitud	T_b	0.93	31	3.3	96	0.13	1.91
Amplitud	T_c	0.72	45	4.2	81	0.84	1.66
Frecuencia	T_a	0.77	26	9.9	100	0.031	0.82
Frecuencia	T_b	0.08	45	6.5	55	0.012	1.52
Frecuencia	T_c	0.13	35	7.3	54	0.017	1.87

Tabla 3.2: Conjuntos de parámetros que maximizan la información mutua, MI, para cada combinación de modulación de entrada y tiempo de integración de la salida, y el valor correspondiente, MI_M , obtenido en cada caso. La tabla muestra el rango en el que cada valor del parámetro fue variado al buscar este máximo. El parámetro, d_2 , siempre se fijó en $d_2 = 0,12/min$. Recordamos que las unidades de concentración arbitrarias son tales que 100 es la concentración de FT nuclear (activa) que se utilizó para calcular MI en el caso de modulación de frecuencia o duración y la concentración máxima en el caso de modulación de amplitud.

Para analizar con más detalle el efecto de cada parámetro sobre MI, se realizó un análisis de sensibilidad variando cada parámetro manteniendo los demás fijos en los valores que dieron MI_M . Mostramos los resultados en la Fig. 3.2. Esta figura confirma las conclusiones iniciales que extrajimos de la Tabla 3.2. Es decir, para todos los casos salvo uno (el de modulación por duración y corte temporal, T_a), se alcanza el valor MI_M cuando d_1 es bajo y k_1 y k_2 son altos, en comparación con la escala de tiempo fija del modelo, $d_2 = 0,12/min$. Si nos fijamos en los parámetros relacionados con la relación entre el promotor y el FT, observamos que K_d tiende a ser medio-alta y el Coeficiente de Hill, n , tiende a ser alto para que MI alcance su valor máximo. Por otro lado, observamos que el rango sobre el cual se pueden variar los parámetros sin cambiar significativamente a MI es diferente dependiendo del parámetro y la modulación de la entrada. En particular, observamos que la transmisión por modulación de amplitud es la más sensible a cambios separados en K_d y n , la modulación por frecuencia es más sensible a cambios en d_1 , mientras que MI apenas cambia con las variaciones en d_1 durante la modulación por duración. En la mayoría de los casos, una vez que los parámetros del modelo superan un umbral, MI per-

manece aproximadamente constante y, en promedio, las mesetas resultan más extendidas para la modulación por duración.

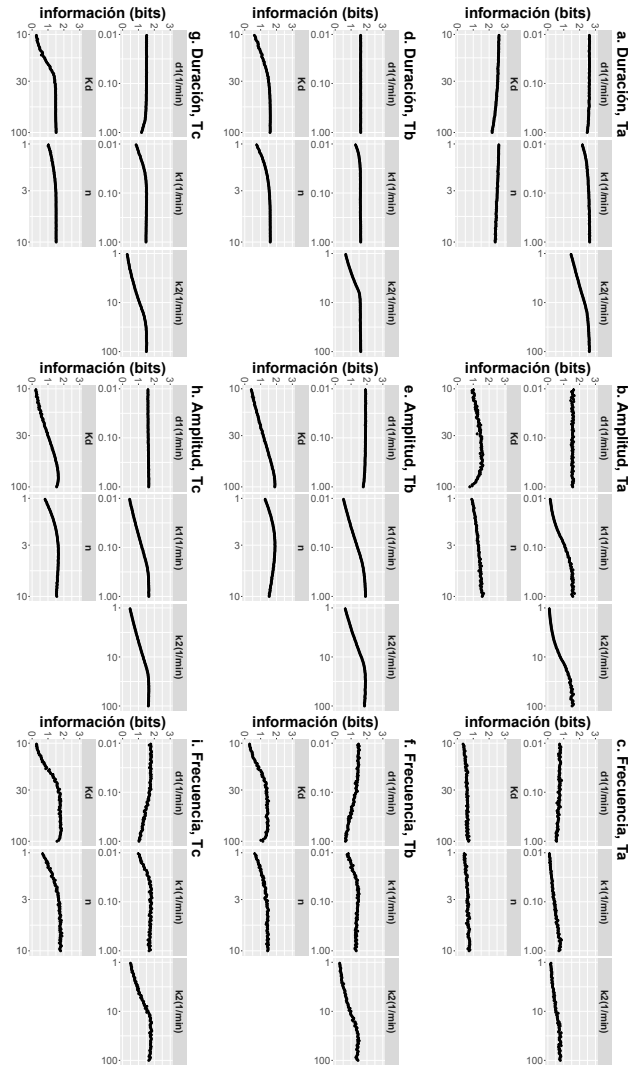


Figura 3.2: (Figura rotada). Comportamiento de la información mutua, MI, alrededor de su máximo valor, MI_M , cuando se varía uno de los parámetros cinéticos del modelo transcripcional. En la figura mostramos MI como función del parámetro correspondiente para cada tipo de modulación de entrada y corte temporal de salida (**a-c**, **d-f** y **g-i** corresponden, respectivamente, al tiempo de corte T_a , T_b y T_c ; **a-d-g**, **b-e-h** y **c-f-i** corresponden, respectivamente, al de tipo de modulación de entrada por duración, amplitud y frecuencia). Excepto para la modulación por duración y el tiempo de corte, T_a , los otros comportamientos son muy similares entre sí. En particular, la mayor información es obtenida si d_1 es baja, k_1 , k_2 y n son altas y K_d está en el rango medio-alto, pero no muy alto. Para el corte temporal T_a y el tipo de modulación por duración, K_d y n necesitan ser bajas para maximizar MI.

3.2.2. Dinámica del modelo y parámetros que maximizan MI.

Para entender por qué la máxima transmisión de información se obtiene, en la mayoría de los casos, para parámetros que satisfacen $k_2 \gg k_1, d_1, d_2$; $n \geq 3$ y $K_d \gtrsim 50$, observamos cómo cambia el curso temporal de algunas variables clave del modelo ($[FT]$, P_1 , ARNm y $Out(time)$) dependiendo de si los parámetros son tales que dan un valor relativamente alto o bajo de MI para alguna combinación de tipo de entrada y corte temporal de salida.

Mostramos en las Figs. 3.3a-c los resultados obtenidos con simulaciones en las que se consideró un único pulso de FT de $10min$ de duración y distintas amplitudes medias. Cuando los parámetros satisfacen $k_2 \gg k_1, d_1, d_2$, $n \geq 3$ y $K_d \gtrsim 50$ observamos que las tres amplitudes se traducen en tres salidas distinguibles (Fig. 3.3a). En los otros ejemplos, la dinámica de activación del promotor es demasiado lenta y nunca se activa (Fig. 3.3b, donde $k_1 \ll d_1$) o el promotor se activa al máximo para todas las entradas (Fig. 3.3c, donde el umbral determinado por K_d y n es demasiado bajo y la escala temporal no es lenta). En el último caso, las tres amplitudes de entrada se traducen en salidas casi indistinguibles. Mostramos en la Fig. 3.3 d, con símbolos, el curso de tiempo de MI calculado para modulación por amplitud y el corte temporal, T_c , usando los parámetros cinéticos del ejemplo en a. Observamos que MI alcanza un máximo aproximadamente a $t = 10min$, *i.e.*, cuando terminan los pulsos de FT , y luego decae hacia un valor ligeramente inferior a una tasa $\lambda \sim 0,09/min$ del orden de la tasa de degradación del ARNm, $d_2 = 0,12/min$, como se ilustra en la curva de ajuste representada con una línea continua en la figura. Obtenemos un comportamiento temporal no monótono similar de MI para la modulación por duración mientras que para la modulación por frecuencia MI aumenta monótonamente a lo largo del tiempo, *i.e.* cuanto más tiempo se le permite al sistema procesar el *input*, más información se puede extraer, como se ilustra en la Fig. 3.4). Estos resultados pueden verse en más detalle en el Apéndice B, más específicamente, en la Fig. B.1.

3.2.3. Pares de parámetros e información mutua

Hasta ahora, hemos analizado cómo varía MI cuando se varía un solo parámetro. En esta Sección analizamos si los parámetros son interdependientes para determinar MI_M . Para ello variamos los valores de todas las posibles parejas de parámetros manteniendo los demás fijos en los valores que daban MI_M y calculamos MI (Fig. 3.5). A partir del análisis de los resultados, determinamos cómo debían variarse los pares para mantener constante MI. Descubrimos que, para todos los tipos de entrada y salida, d_1 tendía a estar positivamente correlacionado con k_1 y k_2 . Es decir, cambiar d_1 en una dirección determinada modifica MI de una forma que se compensa modificando k_1 o k_2 en la misma dirección. Luego encontramos que k_1 y k_2 estaban correlacionados negativamente entre sí, *i.e.*, si k_1 aumenta, k_2 tiene que disminuirse para mantener MI fijo, y viceversa. De manera similar, los parámetros relacionados con la relación entre el promo-

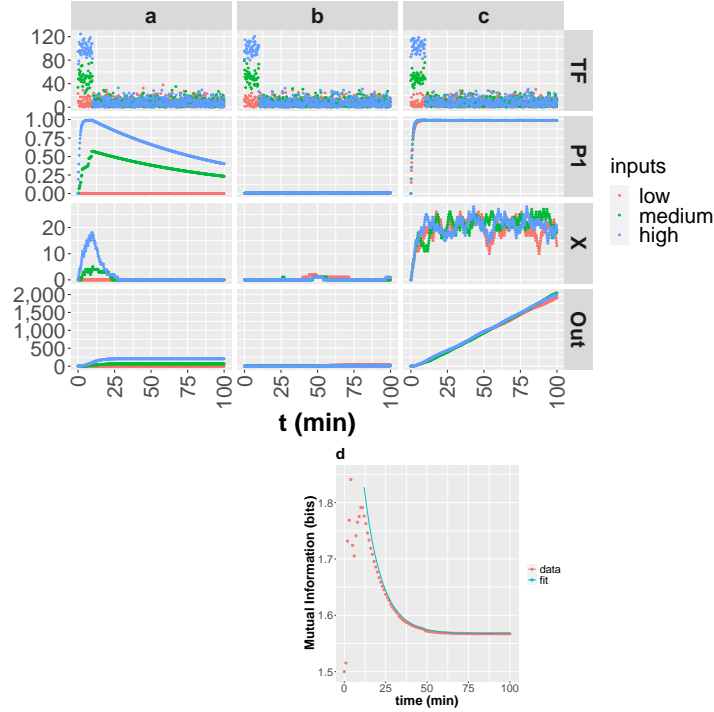


Figura 3.3: Dinámica temporal de $[FT]$, P_1 , X y $Out(time)$ para 3 conjuntos de valores de parámetros y 3 elecciones de curso temporal de $F(t)$ (a-c) y curso temporal de MI para la modulación por amplitud, el tiempo de corte T_c y los parámetros cinéticos del ejemplo en (a). (a-c): En todas las simulaciones, se consideró un único pulso de FT de 10min de duración. Las amplitudes medias fueron 0 (rojo), 10 (verde) y 100 (azul) con un ruido aditivo superpuesto de distribución Gaussian y desvío estándar 10. Los parámetros usados en las simulaciones de la primera columna son $k_1 = 1/min$, $k_2 = 100/min$, $d_1 = 0,01/min$, $Kd = 70$, $n = 10$ para los que MI, computada para la modulación por amplitud y el corte T_c , es $MI = 1,56$ (i.e. 94% de MI_M). Los de la segunda columna son $k_1 = 0,01/min$, $k_2 = 1/min$, $d_1 = 1/min$, $Kd = 10$, $n = 1$ para los que $MI = 0,03$, y los de la tercera: $k_1 = 1/min$, $k_2 = 100/min$, $d_1 = 0,01/min$, $Kd = 10$, $n = 1$ para los que $MI = 0,02$. (d) MI como función del tiempo (símbolos), computada para modulación por amplitud y el corte temporal T_c usando los parámetros cinéticos del ejemplo en (a). La curva en línea llena es de la forma $A \exp(-\lambda(t - 10min)) + MI_a$ la cual ajusta el eventual decaimiento al valor asintótico $\sim MI_a = 1,57$ bits con $\lambda = 0,09/min$.

tor y el FT, K_d y n , están negativamente correlacionados en la mayoría de los casos, con la excepción del corte de tiempo de salida T_a para la modulación de duración, para el cual n y K_d deben ser bajos para que MI alcance su valor máximo.

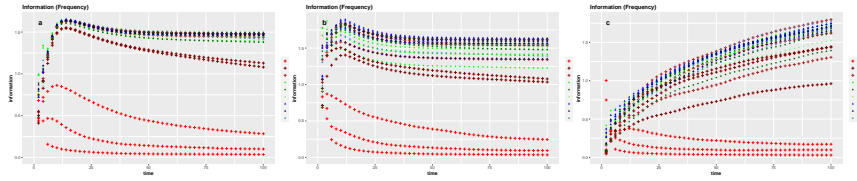


Figura 3.4: Se muestra MI como función del tiempo para las modulaciones por (a) duración, (b) amplitud y (c) frecuencia para distintos cortes temporales del *output*. Los parámetros fueron elegidos de forma que MI difiriera en no más de un 10 % respecto del valor máximo para la combinación de modulación de entrada y corte de tiempo indicada en las leyendas. En todos los casos se muestran 3 ejemplos de parámetros por cada combinación de tipo de modulación de entrada y corte de salida. En cada subfigura, las curvas muestran una dinámica similar entre sí con excepción de aquellas obtenidas para los parámetros que maximizan MI para la modulación por duración y el corte, T_a . (a), (b) En estos dos casos el comportamiento genérico es tal que MI crece hasta la duración máxima del pulso ($10min$), y luego decae. (c) En este caso el comportamiento genérico es tal que MI crece monótonamente con el tiempo. Se observa que MI tiende a ser mayor para los parámetros que optimizan la transmisión por frecuencia.

3.2.4. Un factor de transcripción, dos genes

Uno de los objetivos del presente trabajo es determinar si pueden existir dos promotores, modulados por un mismo FT, capaces de discriminar intensidades de estímulo codificadas en una propiedad diferente de la concentración nuclear de FT pero que no pueden discriminar dichas intensidades si están codificadas en la propiedad para la que el otro promotor discrimina correctamente. Por ejemplo, un promotor que permite discriminar amplitudes del *FT* pero que es “ciego” a los cambios de frecuencia entre pulsos sucesivos del *FT* y un segundo promotor para el que la situación inversa es válida. Para explorar esta posibilidad, buscamos los conjuntos de parámetros cinéticos que dan MI dentro del 90 % del máximo, MI_M , para cierto tipo de modulación de entrada y, al mismo tiempo, dan un valor relativamente bajo para otro tipo de modulación de entrada (y el mismo corte de tiempo de salida), *i.e.*, los conjuntos de parámetros que cumplen la condición *Minmax* (ver Métodos).

La Fig. 3.6 muestra la proyección de las seis regiones *Minmax*, para el corte de tiempo T_c , sobre cada uno de los cinco ejes de parámetros del modelo de transcripción. Aunque tal proyección debe analizarse con cuidado, la Fig. 3.6 ilustra cuánto se puede variar cada parámetro mientras se satisface la condición *Minmax*. El “0”, “1” y “2” representan la modulación por duración, amplitud y frecuencia respectivamente. En otras palabras, la región “01” es aquella que está en una región óptima para la modulación por duración, pero que intenta minimizar la información en la modulación por amplitud, y así el resto. Podemos observar que la sensibilidad de MI a las variaciones de los parámetros es diferente según el parámetro y el par de tipos de modulación de la entrada. Por ejemplo,

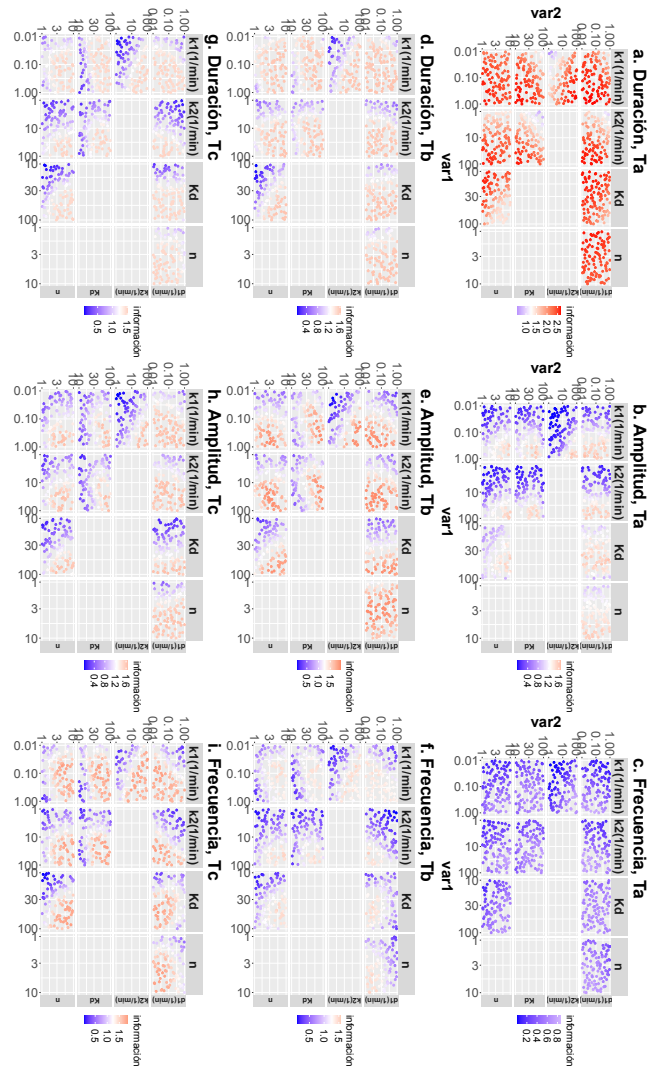


Figura 3.5: (Figura rotada). Información mutua, MI (en bits), como función de dos parámetros cuando los otros están fijos a valores que maximizan MI para los tipos de entrada y salida analizados en cada caso (**a-c**, **d-f** y **g-i** corresponden, respectivamente, al corte temporal de la salida T_a , T_b y T_c ; **a-d-g**, **b-e-h** y **c-f-i** corresponden, respectivamente, a los tipos de modulación de la entrada duración, amplitud y frecuencia). Comportamientos comunes son observados a través de la mayoría de las modulaciones y cortes temporales.

k_1 y d_1 pueden variar en más de un orden de magnitud y, sin embargo, MI difiere en menos del 10% de MI_M para la modulación por duración y del mínimo condicional, MI_m , para la modulación por amplitud, y *viceversa*. El parámetro,

d_1 , por otro lado, apenas puede variarse para mantener MI dentro del 90% de MI_M para la modulación por amplitud o duración y diferir en menos del 10% de MI_m para la modulación por frecuencia. En general, vemos que mantener MI lo más pequeño posible para la modulación por frecuencia en cada región *Minmax* *i2* requiere ajustar los valores de los parámetros relacionados con el tiempo (k_1 , k_2 y d_1) y, en el caso de la región 12, la constante de disociación, K_d , también. Como mostramos a continuación, este ajuste fino permite encontrar promotores (caracterizados por conjuntos de valores de parámetros) que son “ciegos” a los entradas moduladas por frecuencia pero que son buenos para transmitir información para otros tipos de modulación, mientras que lo contrario parece ser imposible.

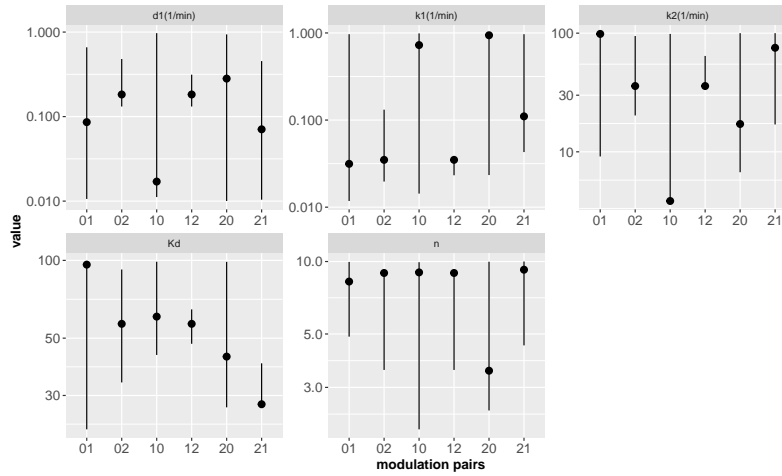


Figura 3.6: Proyección sobre los ejes correspondientes a cada parámetro de las regiones *Minmax* obtenidas para cada par de tipos de modulación y el corte temporal T_c . Los pares de dígitos en los ejes horizontales identifican los modos para los cuales la transmisión de información fue maximizada (primer dígito) y subsecuentemente minimizada (segundo dígito) para determinar las regiones *Minmax*, con 0: duración, 1: amplitud y 2: frecuencia. Los símbolos corresponden a los valores que minimizan MI para la modulación del segundo dígito que MI permanecer dentro del 90% del MI_M para la modulación del primer dígito.

Ahora analizamos la diferencia en la transmisión de información que se logra al elegir algunos de los parámetros del modelo determinados con el condicionamiento *Minmax*. Esto se ilustra en la Fig. 3.7 donde mostramos los valores de MI obtenidos para los modos de entrada i y j usando un conjunto de parámetros en la región ij *Minmax*, para los seis pares ij . Los parámetros son aquellos que minimizan MI para un tipo de modulación de entrada mientras aún están dentro del 90% del máximo del otro (trazado con círculos en la Fig. 3.6). Observamos que MI, para la modulación de duración, tiene aproximadamente el mismo valor ($\sim 1,3 - 1,4$) para todas las combinaciones, incluidas aquellas para las que se minimiza MI por duración. Dado que cada conjunto de parámetros

corresponde a un promotor diferente, esto implica que las duraciones siempre se discriminan por igual independientemente del promotor. En contraste, la modulación por frecuencia es el modo más sensible, con variaciones en MI de casi un bit dependiendo de los parámetros que caracterizan al promotor ($MI \sim 0,8$ para los conjuntos representados en rojo o azul y $MI \sim 1,7$ para los representados en naranja o turquesa). La situación para la modulación de amplitud es intermedia con variaciones de $\sim 0,5$ bits ($MI \sim 1,1$ para los parámetros representados en marrón o turquesa mientras que $MI \sim 1,5$ para los de verde o azul). Estos resultados indican que debería ser posible encontrar promotores que, al estar regulados por un FT, sean buenos para decodificar entradas codificadas en duración o amplitud y, al mismo tiempo, sean “ciegos” a las moduladas en frecuencia, pero que la separación de comportamientos no sería tan claro en la situación contraria, especialmente en el caso de la combinación frecuencia-duración. Ahora analizamos cómo la capacidad de discriminación se refleja en la cantidad de ARNm que se produce utilizando algunos ejemplos (Fig. 3.7).

Mostramos en la Fig. 3.7c el ARNm que se acumula en función del tiempo para varios ejemplos que se obtuvieron utilizando los conjuntos de parámetros de la Fig. 3.7a. Probamos la dinámica del modelo utilizando estos conjuntos de parámetros (cada uno de los cuales puede asociarse con un promotor diferente que responde al mismo FT) y conjuntos de entradas, $[FT](t)$ -A, D, F, caracterizados, respectivamente, por la disminución de la amplitud, duración, y frecuencia del pulso o pulsos de FT , como se muestra en la Fig. 3.7b, relacionados con los modos de transmisión que se han maximizado y minimizado condicionalmente para determinar los conjuntos de parámetros.

De acuerdo con la discusión anterior (Fig. 3.7a), es la codificación por frecuencia (Entradas F) la que produce mayores diferencias dependiendo del conjunto de parámetros. No solo la cantidad de ARNm producido para las Entradas F es mucho mayor para los dos conjuntos en los que se maximizó la transmisión por modulación de frecuencia (regiones *Minmax* 20 y 21) que en aquellos en los que se la minimizó condicionalmente (02 y 12), sino también que los tres cursos temporales de FT entre los de las Entradas F son más fácilmente distinguibles en el primero que en los dos últimos casos.

En el caso de la codificación por duración (Entradas D), las cantidades de ARNm acumulado son ligeramente mayores para el conjunto que pertenece a la región *Minmax* 01 que para el que pertenece a la 10 y ligeramente mayores para este último en comparación con los de las regiones *Minmax* 02 o 20. La capacidad de distinguir las 3 duraciones no parece variar mucho con el conjunto de parámetros.

En el caso de la codificación por amplitud (Entradas A), la cantidad de ARNm acumulado es mayor para el conjunto que se determinó minimizando condicionalmente la transmisión de información para la modulación por amplitud. Para este tipo de entradas, la falta de transmisión de información suficiente para el parámetro establecido en la región *Minmax* 01 en comparación con el establecido en la región 10 (comparar la cruz verde con la cruz marrón en la Fig. 3.7a) parece estar relacionado con la producción de cantidades similares de ARNm independientemente de la amplitud de FT en el primero (comparar

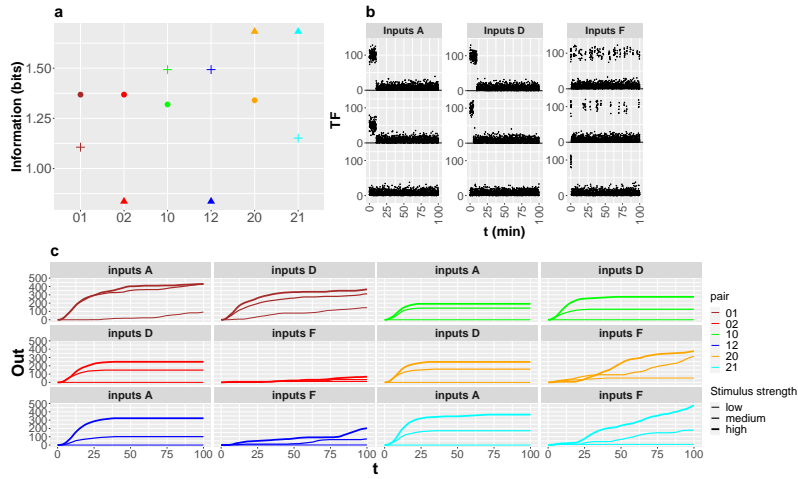


Figura 3.7: Información mutua y comportamientos dinámicos obtenidos usando el conjunto de parámetros de cada región *Minmax* que minimiza MI para el “segundo” tipo de modulación. **a:** MI computada usando el conjunto seleccionado en cada región *Minmax* para cada una de las dos modulaciones usadas para determinar dicha región (círculos: modulación por duración, cruces: modulación por amplitud, triángulos: modulación por frecuencia) y el corte temporal T_c . Las etiquetas en el eje horizontal identifican las regiones *Minmax* con el orden de los dígitos como en la Fig. 3.6. Cada conjunto de parámetros está identificado por un color diferente. Los valores de los parámetros en cada conjunto están mostrados con símbolos en la Fig 3.6. **b-c:** ARNm acumulado ($Out[a.u.]$) como función del tiempo (en min) (**c**) obtenido mediante simulaciones del modelo usando las dinámicas de $[FT]$ mostradas en **b** (las curvas en **c** son mostradas con grosor creciente en relación a la intensidad del estímulo) y el mismo conjunto de parámetros (identificado con el mismo color) que en **a**.

las curvas marrón y verde para las Entradas A). Algo similar ocurre cuando se compara el comportamiento obtenido para las Entradas A y los conjuntos en las regiones *Minmax* 12 y 21 (curvas azul y turquesa de las Entradas A, respectivamente).

En la Fig. 3.8 podemos ver la curva de información como función del tiempo cuando optimizamos los promotores para transmitir información selectivamente a través de un tipo de modulación y no del otro. Se puede observar que la Modulación más sensible es la de Frecuencia, en comparación con la de Duración y la de Amplitud. Esto significa que cuando un promotor está optimizado para transmitir la máxima información posible a través de la Modulación por Frecuencia y, dada esta condición, la mínima posible por Duración, el promotor transmite mucha más información a través de la Modulación por Frecuencia que si estuviera optimizado para Modulación por Duración. Sin embargo, no encontramos que ocurriera lo contrario. La sensibilidad en relación a la Amplitud, es levemente mayor que la de Duración.

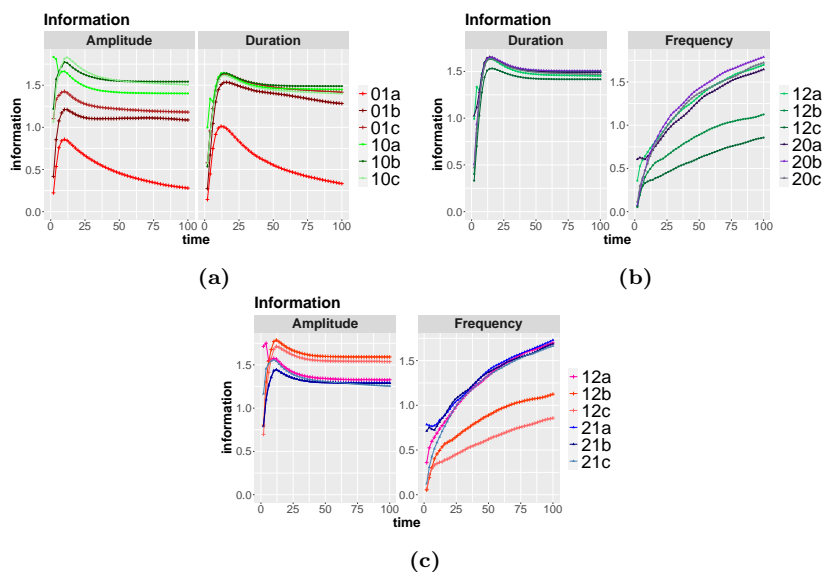


Figura 3.8: Se muestra MI (en bits) como función del tiempo (en minutos). Los parámetros se eligen dentro de las regiones Minmax resultantes de realizar el mismo proceso en otro corte temporal. En particular (con el código de “0”: duración, “1”: amplitud y “2”: frecuencia), el primer código es la modulación para cuya información se intenta mantener cerca del máximo, mientras que el segundo código se corresponde a aquella que se minimiza. Se usan los cortes de tiempo *a*, *b* y *c* mencionados previamente. (3.8a) La Modulación de Duración es mucho menos sensible que la Modulación de Amplitud, con parámetros optimizados para transmisión de información por Duración brindando menos información que los optimizados para transmisión por Amplitud en el caso de la modulación por Amplitud, pero no al revés. (3.8b) Comparación de Modulación por Duración y Frecuencia. La Modulación de Duración es mucho menos sensible que la Modulación de Frecuencia, con parámetros optimizados para transmisión de información por Duración que brindan menos información que los optimizados para transmisión por Frecuencia en el caso de la modulación por Frecuencia, pero no al revés. (3.8c) Modulación por Amplitud y Frecuencia comparadas. La Modulación en Amplitud es menos sensible que la Modulación en Frecuencia, con parámetros optimizados para transmisión de información por Amplitud que proporcionan menos información que los optimizados para transmisión por Frecuencia en el caso de la modulación por Frecuencia, mientras que el efecto en el otro sentido es menor.

3.3. Discusión y conclusiones del Capítulo.

En este Capítulo comparamos la transmisión de información de la que es capaz el paso de transcripción cuando los estímulos extracelulares están codificados en la amplitud, duración o frecuencia de la concentración nuclear de un FT. Esta comparación mostró que, para el modelo analizado en el que se limita la producción de ARNm, el valor máximo que puede tomar MI está entre uno y dos bits (Tabla 3.2). La única excepción corresponde a la modulación por du-

ración en el corte de tiempo de la salida, T_a , que es cualitativamente diferente al resto. Los valores máximos obtenidos para los otros casos fueron ligeramente mayores a los encontrados para la expresión de genes modulada por el FT Msn2 en células salvajes de levadura [47] ilustrados en la Fig. 1.24. Estos experimentos mostraron que la transmisión de información funciona como un interruptor ruidoso (transmitiendo $\sim 1,1 - 1,3$ bits) para dos genes regulados por Msn2, *HXK1* y *SIP18*, cuando son modulados, respectivamente, por la frecuencia y la amplitud del FT. Como se ve en esa figura, cuando los autores mutaron los promotores con el objetivo de mejorar la transmisión de información, la MI para la modulación por amplitud aumentó a $\sim 1,5$ bits [47], muy similar a los valores máximos que encontramos con nuestra exploración de parámetros. Usando el modelo sin imponer un límite al número máximo de moléculas de ARNm producidas por unidad de tiempo, los valores de MI obtenidos (ver Apéndice B) resultan mayores que los observados experimentalmente. Los cálculos de MI reportados en [47] e ilustrados en la Fig. 1.24 fueron obtenidos usando como salida la proteína producida, a la que estimamos como la cantidad de ARNm acumulada a lo largo del tiempo. Si bien ésta es indicativa de la cantidad de proteína producida en la medida que la degradación de esta última ocurra en una escala temporal mayor que la duración de los experimentos, la producción de proteínas está limitada por otros factores que afectan las fluctuaciones en el número de moléculas de proteína producidas (haciendo a las fluctuaciones relativamente mayores respecto de la media que si dichos limitantes no son tenidos en cuenta). Limitar la cantidad de moléculas de ARNm producidas es un modo sencillo de permitir que las fluctuaciones en el reportero de la salida usado en nuestro modelo sean más similares a las de las proteínas producidas (la “salida” en los experimentos). Así obtenemos valores de MI del orden de los obtenidos en los experimentos. Por todas estas razones en este Capítulo nos concentramos en los resultados obtenidos usando el modelo donde se limita el número de moléculas de ARNm producidas por unidad de tiempo.

El modelo transcripcional utilizado (ver Fig. 3.1) se caracteriza por cuatro parámetros directamente asociados a escalas de tiempo (k_1 , k_2 , d_1 , d_2) y por otros dos (K_d y n) relacionados con la relación FT-promotor. La búsqueda de parámetros que maximizan MI mostró (Tabla 3.2) que aquellos que maximizaron MI para todas las combinaciones de modulaciones de entrada y cortes de tiempo de salida correspondieron a una escala de tiempo de transcripción rápida ($k_2 \gg k_1, d_1, d_2$ con $k_1 \gg d_1$ en la mayoría de los casos). Esto se debe a que una tasa de transcripción alta (dada una tasa de degradación de ARNm constante, d_2) genera variaciones de ARNm amplias y sensibles a la entrada (Fig. 3.3). Este resultado concuerda con los de Hansen y O’Shea [46] de acuerdo a los cuales una escala de tiempo rápida permite responder mejor a los estímulos. Con respecto a la relación FT-promotor, nuevamente encontramos un resultado distintivo para la entrada de modulación por duración y corte de tiempo T_a . En particular, encontramos que en este caso se favorecía un rango dinámico más amplio (K_d y n más bajos) en comparación con los demás. Restringiendo la comparación a los parámetros que maximizaron MI para el corte de tiempo de salida T_c , obtuvimos la relación FT-unido al promotor en función de FT

total más “empinada” ($n \sim 7 - 8$) para las entradas moduladas por duración y frecuencia y valores de K_d varias veces mayores que la amplitud del ruido, 10, y que variaban entre el 50 % (para modulación por frecuencia) y el 94 % (para modulación por duración) de la concentración máxima de FT, 100. Suponemos que estos valores relativamente grandes de K_d (que corresponden a una afinidad de unión baja entre el FT y el promotor) son importantes para evitar la aparición de eventos de transcripción gatillados por ruido (un caso límite de los cuales se ilustra en la Fig. 3.3c). Por lo tanto, una afinidad relativamente baja es una condición para garantizar una distinción clara entre los estados “ON” y “OFF” del promotor y está de acuerdo con el hecho de que los sitios de unión funcionales de baja afinidad son comunes en eucariotas, tal vez para distinguir mejor entre FTs similares [58]. Los valores de n que maximizan MI también son relativamente altos y pueden parecer poco realistas si solo se considera la unión FT-promotor. Sin embargo, en nuestro modelo simplificado, tanto el primero como el segundo paso sintetizan un agregado de reacciones/procesos elementales. Una serie de reacciones cooperativas, cada una de las cuales tiene un modesto $n \gtrsim 1$, pueden combinarse dando lugar a un n relativamente alto. Hay múltiples ejemplos de este efecto en ciclos de modificación covalente concatenados, como los que ocurren en las cascadas de proteínas quinasas [35, 36]. De todos modos, los valores de los parámetros que dan lugar al máximo absoluto, MI_M , deben considerarse con cuidado ya que algunos de ellos pueden ser variados sin producir muchos cambios en MI (Figs. 3.2 y 3.5). Por ejemplo, cuando n varía entre 3 y 10, MI varía en menos del 10 % para todos los tipos de modulación de entrada y el corte de tiempo T_c (Figs. 3.2g-i). En cualquier caso, nuestros estudios sobre los parámetros de maximización de MI deberían ser útiles a la hora de diseñar promotores sintéticos. Algunas de las propiedades requeridas para maximizar MI (Tabla 3.2) concuerdan y otras parecen diferir de los resultados experimentales de Hansen y O’Shea [46] descritos en la Sec. 1.5. Estos autores clasificaron a los promotores activados por Msn2 que estudiaron en sus trabajos (ver Fig. 1.18) como de umbral alto (H) o bajo (L) (*i.e.*, que requieren una concentración alta o baja de Msn2 para la inducción, respectivamente, lo que corresponde a K_d alto o bajo) y como lentos (S) o rápidos (F) (*i.e.*, que requieren que el Msn2 permanezca unido más o menos tiempo, respectivamente, para inducir la transcripción). Sus experimentos mostraron que la propiedad de tener un umbral de activación alto iba acompañada de escalas de tiempo lentas (*promotores HS*) y que estos genes eran los modulados por duración (estrés osmótico en el caso analizado por estos autores, Fig. 1.17). Descubrieron, a su vez, que los de umbral bajo estaban caracterizados por escalas de tiempo rápidas (*promotores LF*) y que estos genes eran los modulados por frecuencia (inducidos por la falta de glucosa, ver Fig. 1.17). En base al estudio descrito en este Capítulo, encontramos, en concordancia con estos resultados experimentales, que para el corte de tiempo de salida T_c el máximo MI se obtenía para un K_d mayor (es decir, un umbral mayor) para la modulación por duración que para la de frecuencia. Pero para la escala de tiempo de transcripción (determinada por la relación $\frac{k_1 k_2}{k_1 + d_1}$) encontramos, a diferencia de las observaciones experimentales, un valor dos veces mayor para la modulación por duración en comparación con

la modulación por frecuencia. ¿Cuál es la causa de esta discrepancia?

Podemos explicar las diferencias entre los valores óptimos derivados de nuestro estudio y los calculados con datos experimentales obtenidos en células de levadura por Hansen y O’Shea [47] en términos de la diferente sensibilidad que muestra la MI a variaciones en los parámetros para las distintas modulaciones de entrada. En general, una vez que los parámetros del modelo superan un umbral, MI permanece aproximadamente constante (Fig. 3.2). Por otro lado, en la mayoría de los casos el umbral para un parámetro depende de los valores de otros parámetros, por ejemplo, el umbral de k_2 disminuye al aumentar k_1 (Fig. 3.5). Por lo tanto, los parámetros “óptimos” no son tan significativos *per se*, en el sentido de que podrían variarse (especialmente, de a pares) sin muchos cambios en MI. Nuestro estudio demostró que el rango en el que se podían variar los parámetros sin cambiar mucho la MI era diferente para los distintos parámetros y tipos de modulación de entrada. Por ejemplo, para el corte de tiempo T_c obtuvimos que la transmisión para modulación por amplitud fue la más sensible a cambios en K_d y n , que para la modulación por frecuencia MI varió en más de medio bit si d_1 o k_1 no se mantenían dentro del mismo orden de magnitud que los valores que dan máxima transmisión para este modo y que la modulación por duración fue la menos sensible a variaciones simultáneas en d_1 y k_1 (Figs.3.2,3.5). Esta diferente sensibilidad se reflejó claramente en los resultados de la búsqueda de parámetros que pudieran dar una buena transmisión para una modulación de entrada y una relativamente baja para otra (Figs. 3.6–3.7a). El estudio ilustrado en esta figura mostró que la modulación por frecuencia es tan sensible a los cambios de parámetros que se pueden encontrar promotores que producen una transmisión de información relativamente grande para la modulación por duración o amplitud y una mucho menor para las entradas moduladas por frecuencia: uno que da $MI \sim 1,4$ bits para modulación por duración y $MI \sim 0,8$ bits para modulación por frecuencia (Fig. 3.7a, símbolos rojos) y otro que da $MI \sim 1,5$ bits para modulación por amplitud y $MI \sim 0,8$ bits para modulación por frecuencia (Fig. 3.7a, símbolos azules). Observamos en la Fig. 3.6 que los parámetros que caracterizan a estos ejemplos difieren de los que arrojan mayores valores de MI para la modulación por frecuencia, entre otras cosas, en los valores de la tasa de transición de la conformación inactiva a la activa del promotor, k_1 , y de la constante de disociación, K_d : mientras que $K_d \sim 60$ y $k_1 \sim 0,02/min$ para los casos que dan $MI \sim 0,8$ bits (triángulos rojo y azul en Fig. 3.7a), es $k_1 > 0,1/min$ y $K_d \sim 30 - 40$, en los casos en los que $MI \sim 1,7$ bits (triángulos naranja y turquesa en la Fig. 3.7a). La modulación por frecuencia requiere una mayor afinidad de unión y una transición más rápida al estado activo que los otros modos porque la transcripción debe comenzar durante la duración relativamente corta de los pulsos individuales. Esto concuerda con observaciones del FT de células de mamífero, p53, el que, como mencionamos en la Introducción, cumple un rol fundamental en el desarrollo del cáncer y presenta comportamiento pulsátil dependiendo del estímulo [5, 7, 64, 108] (ver Sec. 1.5). Distintos experimentos mostraron que la afinidad de unión de p53 a los genes proapoptóticos es menor que a la de los que inducen el arresto celular [108] y que, por otro lado, los pulsos de p53 llevan a las células a recuperarse del daño en

el ADN, mientras que una elevación sostenida de p53 conduce habitualmente a la senescencia [83]. Por lo tanto, hasta cierto punto nuestros resultados permiten concluir que la codificación por frecuencia es selectiva no sólo porque requiere un ajuste relativamente fino de los parámetros del promotor sino también porque necesita mayores afinidades de unión.

Tener un promotor tal que $MI \sim 1,5$ bits para la modulación por amplitud y $MI \sim 0,8$ bits para la modulación por frecuencia (símbolos azules en la Fig. 3.7a) es similar a la situación encontrada por Hansen y O'Shea [47] utilizando una versión mutada del promotor, SIP18, que se induce tras una acumulación nuclear prolongada de Msn2 (ver Fig. 1.24). Ampliando nuestra búsqueda más allá de los límites de las regiones *Minmax* ya definidas, encontramos un conjunto de parámetros (o sea, un promotor) que dio $MI \sim 1,1$ bits para la modulación por duración, $MI \sim 1,2$ bits para la modulación por amplitud y $MI \sim 0,7$ bits para la modulación por frecuencia. Los dos últimos valores están más cerca de los obtenidos por Hansen y O'Shea para SIP18 en células salvajes de levadura. Creemos que la transmisión de información para la modulación por frecuencia podría haberse reducido aún más si hubiéramos utilizado un modelo más sofisticado (y realista) que el muy simple de la Fig. 3.1. Hansen y O'Shea [47] se preguntaron por qué la célula no debería "ajustar el nivel de expresión de los genes de estrés a la intensidad del estrés". Nuestros estudios parecen indicar que, al menos en el caso de SIP18, los parámetros podrían haberse ajustado para hacer al promotor lo más "ciego" posible para la modulación por frecuencia sin perder la capacidad de distinguir ON de OFF sin error en el caso de una acumulación nuclear prolongada del FT, Msn2. La menor sensibilidad a la variación de parámetros de las codificaciones por duración y amplitud puede ser la causa de que el promotor, HXK1, que es inducido fisiológicamente por pulsos de Msn2, pueda distinguir intensidades de estímulos codificados en estas dos propiedades. En cualquier caso, según nuestros estudios, incluso bajo parámetros óptimos del promotor, la transmisión de información solo podría ser ligeramente mejor que 1 bit, es decir, cada promotor no puede actuar como un reóstato. Estímulos externos codificados en la amplitud o duración de la fracción nuclear de un FT, que eventualmente pueden activar más de un promotor (siempre que se superen los umbrales de detección), podrían aumentar la capacidad de transmisión de información con una estrategia combinatoria. Estos resultados son aplicables a otras vías de señalización. En particular, procesos similares podrían subyacer a la multiplexación dinámica en la expresión génica dentro de la vía reguladora de p53 en humanos [83], algo que podría investigarse introduciendo mutaciones que podrían cambiar el umbral de activación o las escalas de tiempo características de los promotores involucrados.

Capítulo 4

Las codificaciones por amplitud y frecuencia dan a las células una lente diferente para detectar el entorno.

En este Capítulo extendemos el modelo explorado en el Capítulo anterior para incluir el paso que va desde el estímulo externo hasta la fracción nuclear del FT. Más específicamente, calculamos la MI entre la intensidad del estímulo externo, I_{ext} , y el ARNm producido cuando la primera se codifica en la amplitud o frecuencia medias de los pulsos del FT nuclear. Para tal fin usamos formas funcionales distintas entre I_{ext} y la amplitud o la frecuencia medias dependiendo de la codificación. Mientras que para la codificación por amplitud usamos una función de Hill (ver Ec. (1.1) para la codificación por frecuencia suponemos una dependencia como en la Ec. (1.3) donde el valor medio de la parte estocástica depende exponencialmente del estímulo externo como en la Ec. (1.5). Si bien esta dependencia exponencial fue originalmente tomada de observaciones experimentales de pulsos intracelulares de Ca^{2+} (ver *e.g.* la Fig. 1.10), mostramos en el Capítulo siguiente cómo dicha dependencia puede derivarse a partir de un modelo excitable sencillo. Esto da mayor sustento a la hipótesis subyacente al trabajo del presente Capítulo de que la codificación por frecuencia involucra genéricamente una dependencia exponencial con la intensidad del estímulo. Los resultados acá descritos están contenidos en el trabajo [40].

4.1. Métodos

Como en el Capítulo anterior, para la modulación por amplitud, $[FT]$ se modela como un pulso único de 10 min de duración y amplitud (adimensional), $A = A_{FT} + \zeta$, con ζ un término de ruido y A_{FT} relacionado con la intensidad del estímulo externo (la concentración adimensional de un ligando externo), I_{ext} , por una función cooperativa de Hill 1.1:

$$A_{FT}(I_{ext}) = 100 \frac{I_{ext}^h}{I_{ext}^h + EC_{50}^h}, \quad (4.1)$$

que combina en un solo paso los diversos procesos que van desde el estímulo externo hasta la fracción nuclear del FT. Para la modulación por frecuencia, $[FT]$ se modela como una secuencia de pulsos cuadrados de 1 min de duración cada uno y amplitud (adimensional) 100 más ruido, separados por intervalos de tiempo, $\tau = T_{min} + \eta$, con T_{min} fijo [103]1.3 (en este capítulo llamamos τ a t_{stoch}) y η una variable aleatoria distribuida exponencialmente con un parámetro que depende exponencialmente de I_{ext} [42]1.5 de modo que $T_{IS} \equiv \langle \tau \rangle$ es:

$$T_{IS}(I_{ext}) = T_{min} (1 + \kappa \exp(-bI_{ext})), \quad (4.2)$$

con $\kappa, b > 0$. Usando una distribución para I_{ext} de soporte compacto en el intervalo $[I_m, I_M]$, calculamos numéricamente la información mutua, MI, entre I_{ext} y la integral en el tiempo del ARNm durante los 100 minutos de duración de la simulación. En las simulaciones usamos $T_{min} = 5, 10 \text{ min}$, varios valores de κ y, para la distribución de I_{ext} :

$$I_{ext} = x, \quad (4.3)$$

$$I_{ext} = \exp(4(x - 0,5)) \quad (4.4)$$

con $0 \leq x \leq 1$ distribuida mediante la distribución Beta:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (4.5)$$

para diferentes opciones de $\alpha, \beta > 1$ tales que $\alpha + \beta \geq 3$, de modo de obtener diferentes valores para la mediana y la varianza de I_{ext} . Eligiendo la Ec. (4.3) o la Ec. (4.4), el rango correspondiente de valores de I_{ext} se abarca “lineal” o “logarítmicamente”, respectivamente, con un desvío estándar aproximadamente uniforme para todas las medianas. Dentro de esos rangos, valores crecientes de $\alpha + \beta \geq 3$ producen medianas más diversas y desvíos más similares entre sí, como se ilustra en la Fig. 4.1. Esta CDF se calcula tomando la PDF de la distribución Beta correspondiente, e integrándola numéricamente.

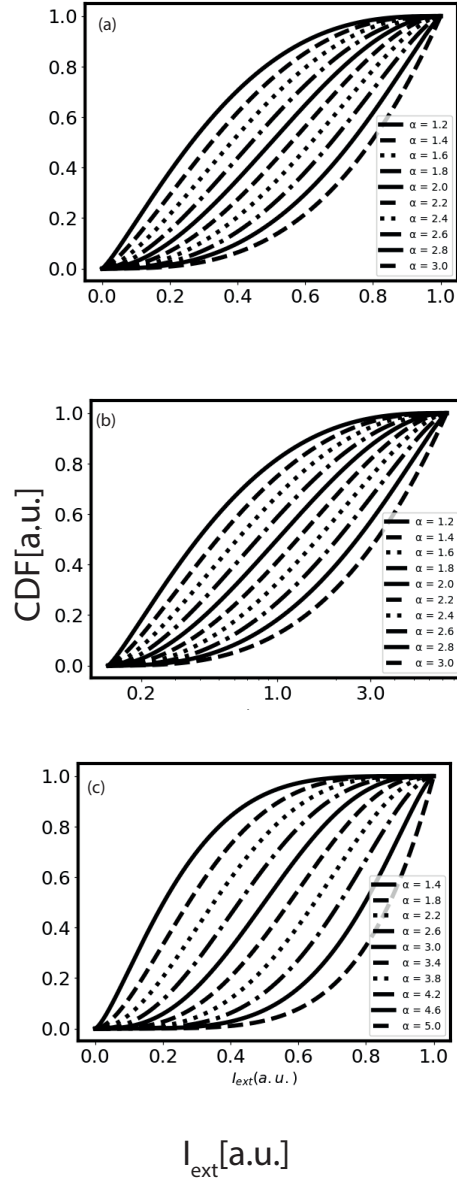


Figura 4.1: Función de distribución acumulada (CDF) de la intensidad del estímulo, I_{ext} , obtenida combinando la Ec. (4.3) ((a), (c)) o la Ec. (4.4) (b) con la Ec. (4.5) para 10 pares (α, β) tales que $\alpha + \beta = 4$ ((a), (b)) o $\alpha + \beta = 6$ ((c)). Variar α y β para valores crecientes de $\alpha + \beta > 2$ produce rangos más amplios (más estrechos) de la mediana I_{ext} (desvío estándar). Para evitar problemas debido a condiciones extremas de los parámetros (como escasez de pulsos o competencia con el T_{min}), se debe usar $4 \leq \alpha + \beta \leq 6$.

4.2. Resultados

Mostramos en las Figs. 4.2 (a),(b) los valores de MI obtenidos para la codificación por amplitud utilizando las 10 distribuciones de I_{ext} de la Fig. 4.1 (a). En ambos casos mostramos MI en función de las medianas correspondientes para diferentes opciones de h y EC_{50} en la Ec. (4.1). Resultados cualitativamente similares fueron obtenidos para las distribuciones de la Fig. 4.1 (b) y para otros valores de parámetros que produjeron máximos locales más pronunciados en la función entre MI y la mediana al aumentar h o $\alpha + \beta$. Observamos en las Figs. 4.2 (a)-(b) que MI es máxima para una mediana de I_{ext} menor o igual a EC_{50} que se aproxima a EC_{50} al aumentar h . Por otro lado, MI permanece dentro de un pequeño porcentaje de su máximo para un rango limitado de medianas de I_{ext} . Esta situación es cualitativamente diferente de la obtenida para la codificación por frecuencia siempre que κ y b en la Ec. (4.2) sean tales que la mayor parte de la distribución de T_{IS} permita la discriminación de frecuencias medias cercanas entre sí (ver más adelante) y, al mismo tiempo, abarque valores que no son demasiado grandes para que la probabilidad de que ocurra un pulso durante el tiempo finito de la simulación ($100min$) no sea despreciable. Estas dos condiciones pueden satisfacerse simultáneamente, como se ilustra en la Fig. 4.2 (c) donde observamos que MI puede permanecer cerca de su valor máximo para un rango más amplio de intensidades de estímulo externo que en el caso de codificación por amplitud. La capacidad de la codificación por frecuencia para distinguir un rango relativamente grande de intensidades de estímulo externo se ha observado en pulsos de Ca^{2+} (ver Fig. 1.10) y de localización nuclear del FT (ver Fig. 1.14). Debemos recordar que el tiempo finito de las simulaciones impone un límite a la frecuencia mínima que probablemente conducirá a la producción de ARNm. Esta limitación también es relevante en situaciones fisiológicas, debido al tiempo de vida finito de las proteínas y a la necesidad de generar respuestas dentro de un marco de tiempo determinado. De hecho, MI disminuye con la mediana de I_{ext} si la probabilidad de provocar al menos un pulso durante el tiempo finito de observación se vuelve demasiado pequeña. Este es el caso de los ejemplos representados con símbolos sólidos en la figura para los cuales esta probabilidad es $\sim 0,54$ y $\sim 0,14$ para $I_{ext} = 0,2$ en los casos con $b = 4$ y $b = 6$, respectivamente. Estas probabilidades son $\sim 0,99$ y $\sim 0,69$ para los ejemplos representados con símbolos abiertos y $\sim 0,9$ para el que tiene asteriscos. Como se explica más adelante, la disminución de MI al aumentar la mediana de I_{ext} se puede atribuir a algunas de las escalas de tiempo más lentas del modelo de transcripción que impiden la discriminación entre tiempos de interpulso muy cortos.

La diferente dependencia de MI con la mediana de I_{ext} para la codificación por amplitud y frecuencia se debe a la forma diferente en la que el conjunto de estímulos externos se “mapea” en los pasos posteriores. Mientras que un mapeo en la forma de la Ec. (4.1) solamente permite discernir un conjunto relativamente angosto de estímulos externos alrededor o por debajo del valor EC_{50} , la Ec. (4.2) puede asignar todo el conjunto de estímulos externos de entrada a un conjunto de valores medios de tiempo entre pulsos discernibles entre sí. Dados dos valores

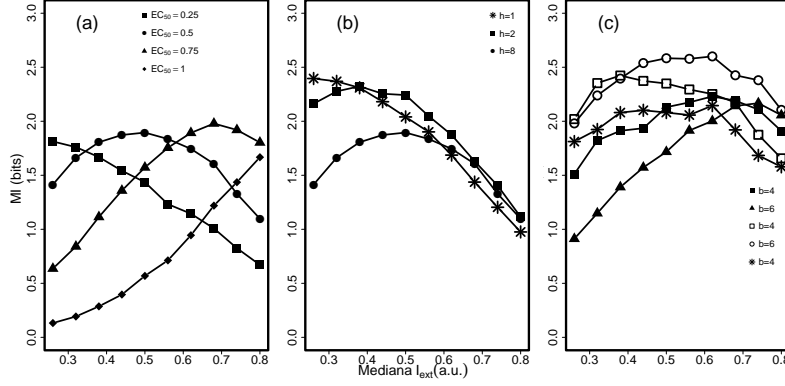


Figura 4.2: Información mutua entre O (Ec. (3.2)) e I_{ext} , para las 10 distribuciones de I_{ext} de la Fig. 4.1 (a) (símbolos), en función de las medianas de I_{ext} , para la codificación por amplitud (en (a), (b)) y por frecuencia (en (c)). Se usó la Ec. (4.1) con $h = 8$ y $EC_{50} = 0,25, 0,5, 0,75, 1$ en (a) y con $EC_{50} = 0,5$ y $h = 1, 2, 8$ en (b). Se usó la Ec. (4.2) en (c) con $T_{min} = 5min$, $b = 4, 6$ y $\kappa = \exp(b)$ (símbolos sólidos); con $T_{min} = 5min$, y las combinaciones $b = 4$, $\kappa = 9$ y $b = 6$, $\kappa = e^4$ (símbolos abiertos) y con $T_{min} = 10min$, $b = 4$ y $\kappa = 9$ (asteriscos).

cercanos, I_{ext} y $I_{ext} + \Delta_I$, el primer paso del mapeo para la codificación por amplitud permitirá que sean distinguibles siempre que la relación, Δ_A , entre la diferencia de las amplitudes medias de FT correspondientes, A_{TF} (Ec. (4.1)), y la suma de los desvíos estándar, $2\sigma_\zeta$, satisfaga:

$$\Delta_A \equiv \frac{\Delta_I}{EC_{50}} \frac{100h}{2\sigma_\zeta} \frac{I_{ext}^{h-1}/EC_{50}^{h-1}}{(I_{ext}^h/EC_{50}^h + 1)^2} > 1, \quad (4.6)$$

donde $\sigma_\zeta = 10$ es el desvío estándar de la variable aleatoria que se suma a A_{TF} en las simulaciones. En el caso de la codificación por frecuencia, los tiempos entre los pulsos resultantes se podrán distinguir siempre que los cuantiles, p y $1-p$, de las distribuciones entre pulsos para I_{ext} e $I_{ext} + \Delta_I$ estén suficientemente separados. Esto se cumple si

$$\Delta_F \equiv \exp(b\Delta_I) \frac{\log(p)}{\log(1-p)} > 1, \quad (4.7)$$

para algún $p > 1/2$ (por ejemplo, $p = 3/4$ garantiza que la superposición de las dos distribuciones no supere $1/4$ de la probabilidad total). Las Ecs. (4.6) y (4.7) son cualitativamente diferentes: Δ_A es una función no monótona de I_{ext} , mientras que Δ_F no depende de I_{ext} . Para $h = 8$, el término que multiplica a Δ_I/EC_{50} en la Ec. (4.6) alcanza su valor máximo en $x \equiv I_{ext}/EC_{50} \approx 1$ y decae en un 50% para $x \approx 0,77$ y $x \approx 1,20$. Así, este primer paso podría distinguir valores, I_{ext} , que difieren entre sí en $\sim 0,2EC_{50}$ si $0,77 < I_{ext}/EC_{50} < 1,20$. Un discernimiento similar se logra sobre los rangos $I_{ext}/EC_{50} \leq 0,4$ para $h = 1$

y $0,17 < I_{ext}/EC_{50} < 1,4$ para $h = 2$. La Ec. (4.7) muestra que para valores de b suficientemente grandes, este paso permitirá discernir intensidades de estímulo cercanas entre sí con la misma resolución sobre todo el rango de valores de I_{ext} . Este comportamiento diferente es evidente en la Fig. 4.3 donde mostramos las CDF de la amplitud del TF, A , (a) y de la frecuencia entre pulsos, $1/\tau$, (b) para cada tipo de codificación con un conjunto de parámetros dado para el mapeo I_{ext} -TF y 5 distribuciones de I_{ext} cuyas medianas se indican en la leyenda. Analizando la Fig. 4.3 (a) en términos de la MI que eventualmente se transmite (curva con círculos sólidos en la Fig. 4.2 (a)) podemos concluir que el valor máximo de MI (aquél para el que la mediana de I_{ext} es 0,5) corresponde a la CDF de amplitud (triángulos sólidos en la Fig. 4.3 (a)) más cercana a una distribución uniforme. En el caso de codificación por frecuencia, casi todas las CDF representadas en la Fig. 4.3 (b) son similarmente cercanas a las de una distribución uniforme, lo que podría explicar la dependencia suave de MI con la mediana de I_{ext} para la curva representada con cuadrados abiertos en la Fig. 4.2 (b). Esta uniformidad se asemeja a la relación entrada/salida óptima obtenida para casos con ruido pequeño e independiente de la media [14, 59]. Ahora bien, éste es solo el primer paso en la generación de la respuesta y otras incertezas son añadidas posteriormente que degradan aún más la información. En el caso de la codificación por frecuencia, intensidades de estímulos de entrada demasiado grandes pueden volverse indistinguibles si la diferencia entre los valores medios correspondientes del tiempo entre pulsos, T_{IP} (Ec. (4.2)), es tan pequeña que es filtrada por algunos de los procesos más lentos del modelo de transcripción. Considerando que $\Delta T_{IP} \approx \kappa T_{min} \exp(-bI_{ext})b\Delta_I$ para dos entradas externas, I_{ext} , que difieren en Δ_I y que, para los ejemplos con símbolos abiertos de la Fig. 4.2 (c) las condiciones que producen MI máximo permiten su discriminación si difieren en al menos $\Delta_I \sim 1/2^{2,5} \approx 0,18$, estimamos que el mínimo discernible de ΔT_{IP} para estos dos ejemplos es $\sim 7 - 8min$ que es aproximadamente el tiempo de degradación característico del ARNm de las simulaciones. Observamos previamente que esta escala de tiempo es clave para limitar la información transmitida (ver Fig. 3.3 d). La observación de tiempos medios entre pulsos indistinguibles en experimentos (ver por ejemplo, Fig. 1.10 o Fig. 1.19) pueden servir para determinar el rango de entradas para el que puede funcionar la codificación por frecuencia.

4.3. Discusión y conclusiones del Capítulo.

Los resultados de este Capítulo muestran cómo la codificación de estímulos externos en la amplitud o en la frecuencia de un intermediario de la señalización (en particular, la fracción nuclear de un FT) otorgan a las células modos cualitativamente distintos de discernir la intensidad de dichos estímulos. Esta diferencia cualitativa que se basa en el distinto modo en que se mapean las intensidades del estímulo para ambas codificaciones, se refleja en que el rango de intensidades para el que hay una buena transmisión de información es limitado para la codificación por amplitud y puede ser mucho más amplio para la codi-

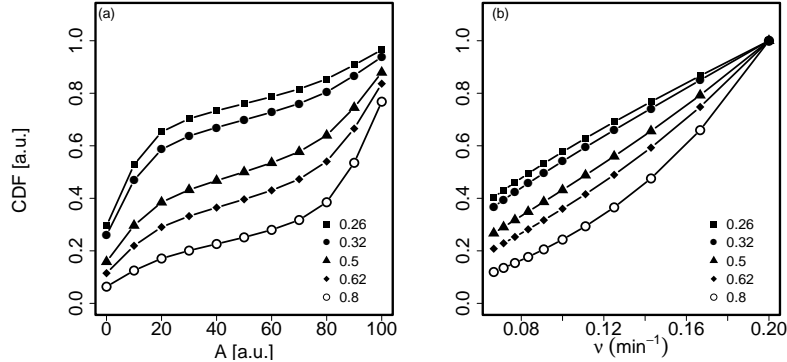


Figura 4.3: CDF de la amplitud del FT, A , (a) y de la frecuencia entre pulsos, $1/\tau$, (b) obtenidas para 5 de las distribuciones de I_{ext} representadas en la Fig. 4.1 (a) usando $EC_{50} = 0,5$ y $h = 8$ en (a) (los mismos parámetros que la curva con círculos sólidos en la Fig. 4.2 (a)) y $T_{min} = 5min$, $b = 4$ y $\kappa = 9$ en (b) (los mismos parámetros que la curva con cuadrados abiertos en la Fig. 4.2 (c)). Las leyendas indican los valores de las medianas de I_{ext} correspondientes.

ficación por frecuencia. Estos dos modos podrían combinarse en un mismo tipo celular para ampliar el rango de intensidades de estímulo distinguibles. Esto podría suceder en el sistema de respuesta a feromona sexual (SRF) en células de levadura [16, 19, 110]. La respuesta canónica en células haploides de tipo **a** de *S. cerevisiae* a la feromona (α -factor) secretada por su potencial compañera, utiliza la codificación por amplitud como en la Ec. (4.1) con $h \sim 1$ y constante dimensional $EC_{50} \approx 3 - 5nM$ [16, 21, 110]. Suponiendo que la curva con $h = 1$ en la Fig. 4.2 (b) representa esta situación ($[\alpha\text{-factor}] = 6-10nM I_{ext}$ para hacer la equivalencia), podemos concluir que MI disminuye en ~ 1 bit cuando la mediana de $[\alpha\text{-factor}]$ aumenta de 1,6-2,6 a 4-6,8 nM . En estas células se observan pulsos de Ca^{2+} con muy baja probabilidad para $[\alpha\text{-factor}] = 0$. La frecuencia de estos pulsos aumenta con $[\alpha\text{-factor}]$ hasta alcanzar un valor de saturación para $[\alpha\text{-factor}] > 10nM$ [19]. Aunque el papel de estos pulsos en la vía de respuesta a feromona aún no es claro, es concebible que éstos resulten en un comportamiento pulsátil de la fracción nuclear de un FT. En particular, como mostramos en la Sec. 1.1, la fracción nuclear del FT, Crz1, presenta comportamiento pulsátil en levaduras en respuesta al estrés por Ca^{2+} , situación en la que también se observaron pulsos intracelulares de Ca^{2+} [17]. Una estimación aproximada en la forma de la Ec. (4.2) obtenida a partir de la Fig. 1.19 da $T_{min} \approx (10 - 12)min$, $\kappa \approx 8 - 9$ y un valor dimensional de la constante $b \sim (0, 4 - 0, 5)/nM$. Suponiendo que la curva trazada con asteriscos en la Fig. 4.2 (c) ($T_{min} = 10min$, $\kappa = 9$, $b = 4$) corresponde a esta situación ($[\alpha\text{-factor}] = (8-10)nM I_{ext}$ para hacer la equivalencia) concluimos que MI difiere de su máximo en menos del 20% para todo el soporte de la distribución de $[\alpha\text{-factor}]$, $[0, 8 - 10nM]$. Bajo condiciones fisiológicas, la posibilidad de distinguir diferencias relativamente sutiles en

[α -factor] es importante para que la célula crezca hacia las regiones de mayor [α -factor] y pueda encontrar a su potencial pareja. Supongamos que las células del par están inicialmente suficientemente separadas entre sí. Es de esperar entonces que la codificación por amplitud sea la utilizada en esas primeras etapas de la detección, por un lado, porque, como ilustra la curva con $h = 1$ de la Fig. 4.2 (b), funciona correctamente para concentraciones relativamente pequeñas (la transición entre las conductas que son características de la ausencia y presencia de la feromona ocurre en [α -factor] $\sim 1nM$ [76]). Por otro lado, porque si [α -factor] es demasiado bajo, llevará un tiempo relativamente largo para que una célula individual recopile suficiente estadística y “responda” correctamente usando la codificación por frecuencia (nuestra estimación de la media del tiempo entre pulsos derivado de [19] da $\sim 70min$ en [α -factor] $=1nM$). Además, existe un desfase temporal entre la exposición de las células al factor α y la aparición de los pulsos de Ca^{2+} que, para niveles saturantes de [α -factor] $=100nM$, es de 30min en promedio [19]. A medida que las células cambian su forma, acercándose a sus parejas, la concentración de feromona alrededor de la creciente proyección de apareamiento aumenta. La codificación por amplitud podría entonces dejar de funcionar para discriminar [α -factor], pero la codificación por frecuencia todavía podría hacer su trabajo correctamente. Por lo tanto, el diferente rango sobre el que cada estrategia de transmisión de información funciona correctamente podría ser la base para que las células de levadura encontraran sus parejas cuando están cerca y cuando están más separadas también.

Capítulo 5

Procesamiento de información celular mediante codificación de frecuencia y excitabilidad.

En el Capítulo anterior vimos cómo la dependencia exponencial del tiempo medio entre pulsos con la intensidad del estímulo externo hacía que la transmisión de información en células usando la codificación por frecuencia fuera cualitativamente distinta de cuando se usa la codificación por amplitud. Allí basamos la elección de dicha dependencia en las propiedades de los pulsos intracelulares de Ca^{2+} mediados por IP_3Rs observados en numerosos tipos celulares distintos (ver Sec. 1.4.1). En este Capítulo mostramos cómo esta dependencia puede ser obtenida a partir del modelo simplificado para la dinámica del Ca^{2+} intracelular mediada por estos canales definido por las Ecs. (1.6)–(1.7) para parámetros en los que esta dinámica es excitable. Más específicamente, incluyendo fluctuaciones a través de un término de ruido aditivo, derivamos la tasa de escape media del estado estacionario y, por lo tanto, el tiempo medio entre pulsos, como función de la fracción, β , de IP_3Rs pasibles de abrirse al ligar Ca^{2+} . Usando un modelo cinético para los IP_3Rs , resultados derivados del análisis de imágenes espacio-temporales de señales de Ca^{2+} y estimaciones previas del IP_3 producido frente a distintos estímulos, cuantificamos las fluctuaciones y relacionamos β con $[\text{IP}_3]$ y la intensidad del estímulo. De esta manera determinamos que el tiempo medio entre pulsos se puede aproximar por una función exponencial de este último para rangos tales que los intervalos de tiempo medio cubiertos son similares o mayores que los observados experimentalmente.

El abordaje seguido en este estudio que está contenido en [41] puede ser aplicado a otros sistemas excitables, en particular, a los que rigen la dinámica de traslocación nuclear de algunos FTs.

5.1. Métodos

Para facilitar los cálculos, como se hizo en [85], introducimos la variable $Z \equiv C + Y$ y reescribimos las Ecs. (1.6)–(1.7) como:

$$\begin{aligned} \dot{C} = & v_0 + v_1\beta - v_{M2} \frac{C^n}{K_2^n + C^n} + \beta v_{M3} \frac{C^p}{K_A^p + C^p} \frac{(Z - C)^m}{K_R^m + (Z - C)^m} \\ & + k_f(Z - C) - kC, \end{aligned} \quad (5.1)$$

$$\dot{Z} = v_0 + v_1\beta - kC. \quad (5.2)$$

En esta sección utilizamos los valores de los parámetros introducidos en [27] como se enumeran en la Tabla 1.1. El único parámetro que cambiamos con respecto a los de [27] es β , ya que el valor $\beta = 0,4$ usado en ese artículo produce oscilaciones mientras que acá buscamos un comportamiento excitable.

5.2. Resultados

5.2.1. Análisis del modelo: estímulos externos y el parámetro β .

Las Ecs. (5.1)–(5.2) modelan la dinámica del Ca^{2+} intracelular cuando está involucrada la liberación de Ca^{2+} a través de IP_3Rs . Como ya mencionamos, los IP_3Rs necesitan unir Ca^{2+} (citósólico) e IP_3 para abrirse y, normalmente, el IP_3 se produce tras la estimulación con efectores externos tales como hormonas o neurotransmisores. El modelo no incluye una descripción explícita de la producción de IP_3 . El efecto del estímulo externo está incluido implícitamente *via* el parámetro β que representa la fracción de IP_3Rs con IP_3 ligado que están listos para abrirse al ligar Ca^{2+} . Por lo tanto, los incrementos en la intensidad del estímulo están representados por incrementos en β . Dupont y Goldbeter analizaron el diagrama de bifurcación del modelo para los parámetros en la Tabla 1.1 y $\beta \in (0, 1)$ [27]. Así determinaron que existe un punto fijo estable para $\beta \lesssim 0,24$ y para $\beta \gtrsim 0,8$ y un ciclo límite estable para valores de β entre ambas regiones. A continuación analizamos los valores de β para los cuales el punto fijo estable es excitable.

5.2.2. Análisis del modelo: soluciones de equilibrio y nulclinas.

Mostramos en la Fig. 5.1 las nulclinas del modelo para tres valores de β . Observamos que para todo $0 \leq \beta \leq 1$ solo hay un punto fijo, (C_{eq}, Z_{eq}) , el cual satisface:

$$C_{eq} = \frac{v_0 + v_1\beta}{k}. \quad (5.3)$$

Es decir, C_{eq} crece con β . Por otro lado, para todos los casos es $\dot{Z} > 0$ ($\dot{Z} < 0$) a la izquierda (derecha) de la nulclina representada con líneas discontinuas mientras que $\dot{C} > 0$ ($\dot{C} < 0$) en la región arriba (abajo) de la nulclina representada con

líneas continuas. Esto es equivalente a la estructura de nulclinas de un ejemplo prototípico de excitabilidad, el modelo de FitzHugh-Nagumo (FHN) dado por las Ecs. 1.10, si identificamos, respectivamente, a $-Z$ y C de las Ecs. (5.1)–(5.2) con las variables lenta, w , y rápida, v , del modelo FHN. Por lo tanto, cuando las nulclinas se cruzan a la izquierda del valor, C , en el cual la nulclina $\dot{C} = 0$ tiene su máximo local, el punto fijo es excitable como se ilustra a continuación. En efecto, para el caso de la Fig. 5.1 (a), que corresponde a $\beta = 0,09$, el único punto fijo es excitable: condiciones iniciales que comienzan a uno u otro lado del “rama media” de la nulclina $\dot{C} = 0$ dan lugar, respectivamente, a una excursión relativamente larga en el espacio de fases o decaen monótonamente hasta el punto fijo. Esto se ilustra en la Fig. 5.2 (a) donde mostramos la evolución temporal de C para $\beta = 0,09$ y dos de esas condiciones iniciales. A medida que β aumenta, C_{eq} se acerca al valor en el que la nulclina $\dot{C} = 0$ tiene su máximo local. La solución de equilibrio continúa siendo excitable siempre que C_{eq} se mantenga lejos de este máximo local, por ejemplo, para $\beta = 0,15$ para el que mostramos las nulclinas y dos trayectorias en las Figs. 5.1 (b) y 5.2 (b), respectivamente. Se puede observar en la Fig. 5.2 (b) que la excursión en el espacio de fases es mayor que en la Fig. 5.2 (a), consistente con la estructura de las nulclinas. En $\beta \approx 0,183$ el punto fijo se convierte en un foco estable y luego pierde estabilidad en $\beta \approx 0,237$ a través de una bifurcación de Hopf. Para $0,237 \lesssim \beta \lesssim 0,775$ la solución (asintóticamente) estable es un ciclo límite, como se ilustra en la Fig. 5.2 (c) que corresponde a $\beta = 0,3$ para el cual las nulclinas se muestran en la Fig. 5.1 (c). A medida que β se acerca desde abajo a $\beta = 0,775$, las oscilaciones se vuelven más sinusoidales y de menor amplitud hasta que una nueva bifurcación de Hopf, en $\beta \approx 0,775$, vuelve a estabilizar el punto fijo.

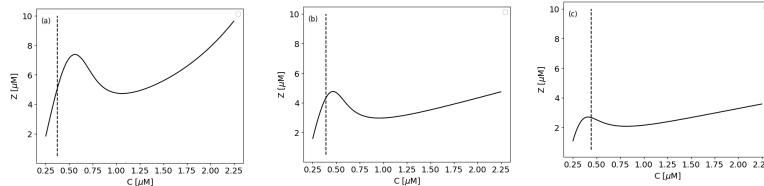


Figura 5.1: Nulclinas del modelo dadas por las ecuaciones (5.1)–(5.2) para los valores de los parámetros de la Tabla 1.1 y $\beta = 0,09$ (a), $\beta = 0,15$ (b) y $\beta = 0,3$ (c). La línea lina muestra la nulclina $\dot{C} = 0$ y la línea de puntos muestra la nulclina $\dot{Z} = 0$.

El modelo simple proporciona valores razonables de C_{eq} y Y_{eq} para el rango de parámetros para los cuales la única solución estable es excitable. A saber, en la mayoría de los tipos celulares, la concentración de Ca^{2+} citosólico libre es muy pequeña en el estado basal ($\sim 50 - 100 \text{ nM}$ [81]), consistente con los valores representados en la Fig. 5.3 (que satisfacen $C_{eq} \sim 0,4 \mu\text{M}$ en toda la figura). En los ovocitos de *X. laevis*, la concentración de Ca^{2+} libre en el lumen del retículo endoplasmático (RE) se estimó en $\sim 200 \mu\text{M}$, el cociente entre el volumen luminal y el citosólico en 0,7 y el de la fracción de Ca^{2+} libre en el

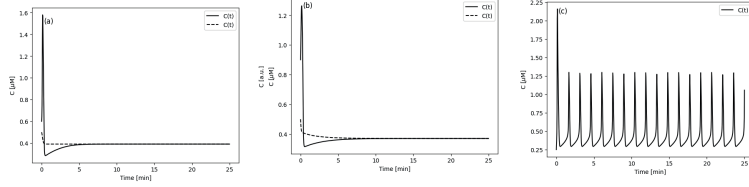


Figura 5.2: Curso temporal de C obtenido resolviendo las ecuaciones (5.1)–(5.2) para los mismos valores de parámetros que en la Fig. 5.1. En (a) y (b) mostramos la evolución temporal de dos condiciones iniciales para ilustrar que el único punto fijo del sistema es estable pero excitable. En (c) hay un ciclo límite estable.

citosol y en el RE como 0.00125 [66]. En una interpretación mecanística de las Ecs. (1.6)–(1.7), Y podría identificarse con el producto de estos dos cocientes ($\sim 0,009$) por la concentración de Ca^{2+} luminal libre. Así, los valores de equilibrio, $Y_{eq} \sim 3,5 - 4 \mu\text{M}$, como los que se pueden derivar de la Fig. 5.3, resultan consistentes con las estimaciones de [66].

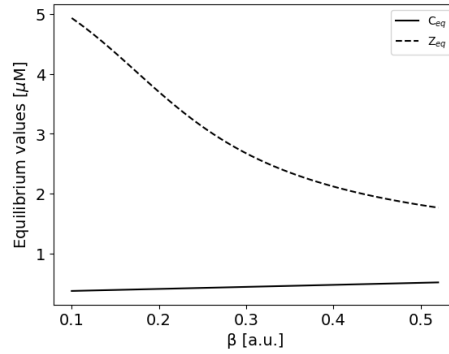


Figura 5.3: Concentraciones de equilibrio, C_{eq} y Z_{eq} , para valores de β para los cuales la única solución de equilibrio es excitable. La concentración de Ca^{2+} correspondiente en el lumen del RE se puede derivar de $Y_{eq} = Z_{eq} - C_{eq}$.

5.2.3. Descripción del modelo en términos de variables lentas y rápidas

Los valores de equilibrio, C_{eq} y Z_{eq} , que se muestran en la Fig. 5.3 difieren en aproximadamente un orden de magnitud. Por otro lado, durante los pulsos (relativamente cortos) de C como el que se ilustra en la Fig. 5.2 (b), C aumenta en un factor menor que 4. Como es estándar para el análisis de reacciones con enzimas en la aproximación cuasiestacionaria [57](pp. 9-12), introducimos concentraciones, $\bar{C} = C/C_{eq}$ y $\bar{Z} = Z/Z_{eq}$ y tiempo adimensional, $\tau = tv_0/C_{max}$,

para reescribir las ecuaciones (5.1)–(5.2) como:

$$\begin{aligned} \frac{d\bar{C}}{d\tau} = & 1 + \frac{v_1}{v_0}\beta - \frac{v_{M2}}{v_0} \frac{\bar{C}^n}{\bar{K}_2^n + \bar{C}^n} + \beta \frac{v_{M3}}{v_0} \frac{\bar{C}^p}{\bar{K}_A^p + \bar{C}^p} \frac{(\bar{Z}/\epsilon - \bar{C})^m}{\bar{K}_R^m + (\bar{Z}/\epsilon - \bar{C})^m} \\ & + \frac{k_f}{v_0}(\bar{Z}/\epsilon - \bar{C}) - \frac{kC_{max}\bar{C}}{v_0}, \end{aligned} \quad (5.4)$$

$$\frac{d\bar{Z}}{d\tau} = \epsilon \left(1 + \frac{v_1}{v_0}\beta - \frac{kC_{max}\bar{C}}{v_0} \right), \quad (5.5)$$

donde $\epsilon \equiv C_{eq}/Z_{eq} \ll 1$. De las Ecs. (5.4-5.5) se desprende claramente que \bar{Z} (y, por lo tanto, Z) es la variable lenta. Esta separación de escalas temporales se ve reforzada aún más por la elección de los valores de los parámetros de la Tabla 1.1 que son tales que v_{M3} , el cual solo afecta la dinámica de C , corresponde a una velocidad mucho más rápida que la de todos los otros procesos.

5.2.4. Probabilidad por unidad de tiempo y período promedio entre pulsos sucesivos.

Estudiamos la probabilidad por unidad de tiempo de que ocurra un pulso en C restringiendo el análisis a la dinámica en la variedad rápida. Fijamos $0,1 \leq \beta \leq 0,22$, $Z = Z_{eq}(\beta)$ y trabajamos con la ecuación:

$$\begin{aligned} \dot{C} = & v_0 + v_1\beta - v_{M2} \frac{C^n}{K_2^n + C^n} + \beta v_{M3} \frac{C^p}{K_A^p + C^p} \frac{(Z_{eq}(\beta) - C)^m}{K_R^m + (Z_{eq}(\beta) - C)^m} \\ & + k_f(Z_{eq}(\beta) - C) - kC \equiv -\frac{dV_\beta}{dC}, \end{aligned} \quad (5.6)$$

donde hemos introducido la definición del potencial, $V_\beta(C)$. La tasa de aparición de pulsos obtenida a partir del análisis de la Ec. (5.6) dará la componente estocástica del tiempo entre pulsos. El tiempo total entre pulsos será entonces la suma de esta componente estocástica, la duración del pulso y el tiempo de recuperación (que no se tiene en cuenta en la Ec. (5.6)).

En la Ec. (5.6) v_0 y $k_f Y = k_f(Z_{eq} - C)$ representan un flujo de Ca^{2+} hacia el citosol desde el medio extracelular y el RE, respectivamente, que es independiente de Ca^{2+} y del IP_3 citosólicos. Hay un término adicional que depende de IP_3 , $v_1\beta$, que se incluyó para dar cuenta de la observación de que la $[\text{Ca}^{2+}]$ citosólica media aumenta con el nivel de estimulación [27]. Todos estos procesos se ven afectados por el ruido, ya que son el resultado de eventos localizados de ingreso de Ca^{2+} . El IP_3 producido como consecuencia del estímulo externo también puede fluctuar, de modo que β también está sujeto a fluctuaciones aleatorias. Suponiendo que, para cada nivel de estimulación externa, I_{ext} , las fluctuaciones en β , $\delta\beta(I_{ext})$, satisfacen $\delta\beta(I_{ext}) \ll \langle\beta\rangle(I_{ext})$, podemos abarcar toda la parte ruidosa de estos procesos en un solo término aditivo y reescribir la ecuación (5.6) como:

$$\dot{C} = -\frac{dV_\beta}{dC} + \zeta, \quad (5.7)$$

con $\beta \cong \langle \beta \rangle (I_{ext})$ y ζ un término de ruido blanco gaussiano tal que $\langle \zeta(t)\zeta(t') \rangle = 2D\delta(t-t')$ donde D puede, en principio, variar con β . Discutimos más adelante hasta qué punto es razonable suponer que $\delta\beta(I_{ext}) \ll \langle \beta \rangle (I_{ext})$ y las implicancias de la dependencia de β con el coeficiente, D .

Mostramos en las Figs. 5.4 (a) y (b) una gráfica de $V_\beta(C)$ para los mismos valores de β que en las Figs. 5.1 (a) y (b). Observamos que, en ambos casos, hay dos mínimos locales, C_1 y C_2 , que corresponden a puntos fijos estables de la Ec. (5.7) y entre ellos un punto silla, C_S , que determina la separatriz entre las cuencas de atracción (en la variedad rápida) de los puntos fijos estables. Uno de estos puntos fijos es $C_1 = C_{eq}(\beta)$, es decir, corresponde a la solución estacionaria (excitable) del modelo bidimensional completo dado por las Ecs. (5.1)-(5.2). La diferencia, $\Delta V_\beta \equiv V_\beta(C_S) - V_\beta(C_1)$, es la altura de la barrera que debe superarse para que el sistema (restringido a la variedad rápida) pase de C_1 a C_2 . En el modelo real, dado por las Ecs. (5.1)-(5.2), si $\beta \lesssim 0,22$, superar esta barrera implica que el sistema sufre una larga excursión en el espacio de fases y, por lo tanto, se produce un pulso. La altura de la barrera disminuye con β como se muestra en la Fig. 5.4 (c) donde graficamos ΔV_β .

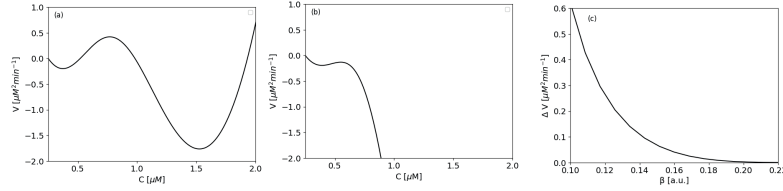


Figura 5.4: El potencial, V_β , que determina la dinámica en la variedad rápida (Ec. (5.7)) para $\beta = 0,09$ (a) y $\beta = 0,15$ (b) y la altura de la barrera, $\Delta V_\beta \equiv V_\beta(C_S) - V_\beta(C_1)$, en función de β , donde C_S corresponde al máximo local de V_β y C_1 al mínimo local a la izquierda de C_S (que es el valor en la solución de equilibrio de las Ecs. (5.1)-(5.2)).

Para un V_β biestable (como en las Figs. 5.4 (a)-(b)) la Ec. (5.7) corresponde a la ecuación paradigmática analizada por Kramers [45] quien obtuvo que el tiempo que le toma al sistema salir de la vecindad del mínimo local, C_1 (que corresponde al punto fijo excitable de las Ecs. (5.1)-(5.2)) está distribuido exponencialmente con una tasa de escape, r_β , que, en el límite de bajo ruido ($D \ll \Delta V_\beta$), puede aproximarse por:

$$r_\beta = \frac{\left(V_\beta''(C_1) |V_\beta''(C_S)| \right)^{1/2}}{2\pi} \exp(-\Delta V_\beta/D), \quad (5.8)$$

con $V_\beta'' = d^2 V_\beta / dC^2$. En base a esto, el tiempo medio entre pulsos, $\langle T_{IS} \rangle$, sería:

$$\langle T_{IS} \rangle = T_{min} + \frac{2\pi}{\left(V_\beta''(C_1) |V_\beta''(C_S)| \right)^{1/2}} \exp(\Delta V_\beta/D), \quad (5.9)$$

con T_{min} la suma de la duración del pulso y el tiempo de recuperación. De acuerdo a la Ec. (5.9), $\langle T_{IS} \rangle$ depende exponencialmente de ΔV_β que es una función decreciente de β (ver Fig. 5.4 (c)) y, por lo tanto, de la intensidad del estímulo externo, I_{ext} . En la siguiente sección combinamos una variedad de resultados publicados previamente para derivar la dependencia de $\langle T_{IS} \rangle$ con I_{ext} dentro del marco del modelo simple considerado en este Capítulo.

5.2.5. Tiempo medio entre pulsos e intensidad del estímulo externo

Para determinar la dependencia de ΔV_β con I_{ext} necesitamos estimar el nivel de ruido (que entra a través de D en la Ec. (5.9)) y la relación entre β e I_{ext} . Para esto último necesitamos relacionar β con la concentración de IP_3 . Recordamos aquí que β es la fracción de IP_3 Rs con IP_3 ligado que están listos para abrirse una vez que pegan Ca^{2+} . Para estimar la relación entre β e IP_3 utilizamos el modelo de DeYoung-Keizer [24] que considera la estructura tetramérica del canal y supone que tres de los monómeros deben tener IP_3 unido y el sitio inhibidor de Ca^{2+} vacío para que el canal eventualmente pueda abrirse (al unirse Ca^{2+} al sitio activador en los tres monómeros con IP_3 unido). La constante de disociación de la reacción de pegado/despegado del IP_3 a sus sitios de ligadura depende de si el Ca^{2+} está unido o no al sitio inhibidor. Suponiendo que inmediatamente antes de que se produzca el pulso, la $[Ca^{2+}]$ es aproximadamente igual a su (bajo) valor basal, de modo que el sitio inhibidor del receptor probablemente esté libre de Ca^{2+} , utilizamos la constante de disociación que corresponde a la situación sin Ca^{2+} pegado al sitio inhibidor, $K_{IP} = 0,13\mu M$, y relacionamos β y $[IP_3]$ como:

$$\beta = \left(\frac{[IP_3]}{[IP_3] + K_{IP}} \right)^3. \quad (5.10)$$

Mostramos en la Fig. 5.5 (a) un gráfico de $[IP_3]$ vs β derivado de la Ec. (5.10) con $K_{IP} = 0,13\mu M$. Observamos que la relación entre $[IP_3]$ y β es aproximadamente lineal en el rango de valores de β para los cuales la dinámica del modelo simple (Ecs. (5.1)– (5.2)) es excitable.

Para estimar el nivel de ruido, utilizamos los experimentos de [69] realizados en ovocitos de *Xenopus Laevis* en presencia de diferentes valores constantes de $[IP_3]$ intracelular (ver Fig. 1.7). Más específicamente, usamos el resultado del análisis presentado en [69] de la frecuencia y amplitud de los *puffs* de Ca^{2+} que ocurren antes de una onda de Ca^{2+} global (un pulso). Para nuestro cálculo suponemos que los *puffs* de Ca^{2+} son los principales eventos de liberación aleatoria que entran en el término de ruido, ζ . En los experimentos [69], las señales de Ca^{2+} fueron observadas usando el fluoróforo de Ca^{2+} Fluo4 y microscopía confocal escaneando una región de $65\mu m \times 65\mu m$. La fluorescencia asociada a cada *puff* fue recolectada de una región de $8\mu m \times 8\mu m$. Para el análisis, los autores separaron las observaciones según el tiempo medio típico entre pulsos (ondas) que, como se esperaba, disminuyó a medida que aumentó

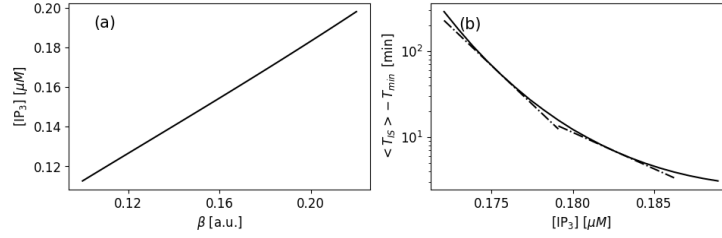


Figura 5.5: (a) $[IP_3]$ como función del nivel de saturación de IP_3Rs β , dado por la Ec. (5.10) con $K_{IP} = 0,13\mu M$. (b) Parte estocástica media del tiempo entre picos, $\langle T_{IS} \rangle - T_{min}$ en función de $[IP_3]$ obtenida combinando las ecuaciones (5.10) y (5.9), con $D = 0,0015\mu M^2/min$ y $K_{IP} = 0,13\mu M$. Dos rectas tangentes a la trazada. También se muestran la curva en $[IP_3] \approx 0,176\mu M$ y en $[IP_3] \approx 0,183\mu M$.

la concentración de IP_3 . Así analizaron los *puffs* que precedían a pulsos separados por tiempos medios, $\langle T_{IS} \rangle$, intermedios (de entre 25 y 50s) y altos (en su mayoría, entre 50 y 70s). En ambas situaciones observaron que la amplitud de los *puffs* se mantenía aproximadamente igual. Midieron la amplitud en términos del aumento relativo de la fluorescencia con respecto a su nivel basal, $\Delta F/F_0$, obteniendo $\Delta F/F_0 \sim 1$ en la mayoría de los casos. Esto implica que el aumento local en $[Ca^{2+}]$ debido a un solo *puff* fue del orden de la concentración basal de Ca^{2+} ($0,4\mu M$ en nuestro modelo simple, ver Fig. 5.3). En nuestro modelo (Ecs. (5.1)–(5.2)), C corresponde a la concentración promediada sobre toda la célula. Considerando la región de $8\mu m \times 8\mu m$ que se usó para monitorear los *puffs* en [69], suponemos que el aumento en la concentración promediada sobre toda la célula debido a un solo *puff* es $\sim (8/16)^2 0,4\mu M \approx 0,006\mu M$. En el caso de tiempos intermedios entre pulsos (valores medios de 25–50 s), los *puffs* comienzan a ocurrir aproximadamente después de que haya transcurrido el 25% del tiempo entre pulsos (es decir, durante 19–38 s) y el número total de *puffs* es ~ 1000 . Por lo tanto, en este caso, el número de *puffs* por unidad de tiempo es $\sim (26–53)/s$. Esto significa que, en promedio, la amplitud del término de ruido es $\zeta \sim (26–53)0,006\mu M/s = (0,16–0,32)\mu M/s$. Para el cálculo de Kramers suponemos que $\langle \zeta(t)\zeta(t') \rangle = 2D\delta(t-t')$. Los *puffs* no son instantáneos, por lo que aproximaremos $\delta(t-t')$ por la duración típica de uno de ellos. Aquí tenemos dos opciones, tomar en cuenta la duración de la liberación de Ca^{2+} o la de la elevación de Ca^{2+} local (estimada, respectivamente, en 20 y 150ms en [15]). Usando esta última, entonces estimamos $D \sim (0,16^2 - 0,32^2)\mu M^2/s^2 \times 0,15s/2 = (0,002 - 0,007)\mu M^2/s = (0,13 - 0,46)\mu M^2/min$. En los experimentos con un tiempo medio entre pulsos alto (la mayoría de ellos entre 50 y 70s), el número de *puffs* entre pulsos es aproximadamente 120 y aunque la mayoría ocurre después de que haya transcurrido el 25% del tiempo entre pulsos (es decir, durante 38–45s), hay una fracción no despreciable que ocurre antes. Por lo tanto, en este caso, el número de *puffs* por unidad de tiempo es $\sim (2–3)/s$, mucho menor que para valores intermedios de $\langle T_{IS} \rangle$, mientras que la amplitud permanece aproximadamente igual. Repitiendo el cálculo para esta frecuencia de *puffs*

más pequeña obtenemos $\zeta \sim (2-3)0,006\mu M/s = (0,012-0,018)\mu M/s$ que da $D \sim (1,1-2,4)10^{-5}\mu M^2/s = (0,0007-0,0015)\mu m M^2/min$. Observamos entonces que D aumenta significativamente con el nivel de estimulación, es decir, con β . Si restringimos el cálculo a los experimentos con nivel bajo o con nivel intermedio de estimulación, sin embargo, D varía como máximo en un factor 2. Por lo tanto, es razonable considerar un valor fijo de D en tal caso como hacemos acá. Ahora bien, para que el cálculo de Kramers que lleva a la Ec. (5.9) sea válido, es necesario que $D \ll \Delta V_\beta$. Esta condición no se cumple para casi todo el rango de valores β de la Fig. 5.4 (c) si usamos $D \sim (0,13-0,46)\mu M^2/min$, la estimación derivada de los experimentos con tiempos entre pulsos intermedios, mientras que sí se satisface para $\beta \lesssim 0,18$ si utilizamos la estimación obtenida para tiempos entre pulsos largos, $D \sim (0,0007-0,0015)\mu M^2/min$.

Mostramos en la Fig. 5.5 (b) la media de la parte estocástica del tiempo entre pulsos, $\langle T_{IS} \rangle - T_{min}$, dada por la Ec. (5.9) en función de $[IP_3]$ para $D = 0,0015\mu M^2/min$, donde hemos usado la Ec. (5.10) con $K_{IP} = 0,13\mu M$ para relacionar β e $[IP_3]$. Las dos líneas rectas dibujadas en la figura ilustran que $\langle T_{IS} \rangle - T_{min} \approx A \exp(-\lambda[IP_3])$ con $\lambda \approx 414/\mu M$ para el rango, $0,172\mu M \lesssim [IP_3] \lesssim 0,179\mu M$ (para el cual $15min \lesssim \langle T_{IS} \rangle - T_{min} \lesssim 290min$) y con $\lambda \approx 195/\mu M$ para el rango, $0,179\mu M \lesssim [IP_3] \lesssim 0,186\mu M$ (para el cual $4min \lesssim \langle T_{IS} \rangle - T_{min} \lesssim 15min$). Debemos recordar que la curva de la Fig. 5.5 (b) se obtuvo utilizando un valor fijo de D y el análisis del párrafo anterior mostró que D aumenta significativamente con el nivel de estimulación, es decir, con $[IP_3]$. El uso de dos valores diferentes de D dentro de cada uno de los rangos de $[IP_3]$ cubiertos por las líneas rectas en la Fig. 5.5 (b) no permite extender la dependencia exponencial a una región más grande. En cualquier caso, el cociente de los períodos medios entre pulsos entre los cuales la curva en la Fig. 5.5 (b) es aproximadamente una función exponencial de $[IP_3]$ es mayor que para la mayoría de los experimentos reportados en [103]. En particular, para el rango $0,179\mu M \lesssim [IP_3] \lesssim 0,186\mu M$, este cociente es 3,75, que es dos veces mayor que el cociente análogo obtenido para los experimentos realizados en hepatocitos primarios de rata en presencia de vasopresina y muy similar al obtenido para células de riñón embrionario humano en presencia de carbacol (ver Fig. 1.10 tomada de [103]).

5.3. Discusión y conclusiones del Capítulo.

Los cálculos acá descritos muestran que es razonable suponer que el tiempo medio entre pulsos de Ca^{2+} intracelular mediados por IP_3 Rs es una función exponencial de $[IP_3]$. Ahora bien, en los experimentos [103], la dependencia exponencial fue encontrada entre la media de la parte estocástica del tiempo entre pulsos y la concentración del ligando externo (vasopresina o carbacol en los dos ejemplos descritos en el párrafo anterior e ilustrados en la Fig. 1.10). Debemos preguntarnos entonces si la dependencia exponencial que encontramos entre $\langle T_{IS} \rangle - T_{min}$ e $[IP_3]$ también es válida para la intensidad del estímulo externo, I_{ext} . El IP_3 es producido mediante la hidrólisis de precursores que

se encuentran en la membrana plasmática producida por la enzima fosfolipasa C (PLC). La PLC se activa como resultado de la unión de efectores externos a ciertos receptores de la membrana plasmática. La intensidad del estímulo externo en nuestra descripción es, por lo tanto, proporcional a la concentración del efector externo. La tasa de producción de IP_3 es proporcional al número medio de receptores unidos al efector que, en equilibrio, suele ser una función de Hill de la concentración del efector. El IP_3 , a su vez, se degrada a una velocidad $\sim 0,11 - 0,14\text{s}^{-1}$ [107]. Por lo tanto, podemos suponer que $[\text{IP}_3]$ e I_{ext} están relacionados por una función de Hill (Ec. (1.1)), es decir mediante:

$$[\text{IP}_3] = [\text{IP}]_0 \frac{I_{ext}^h}{I_{ext}^h + K_D^h}, \quad (5.11)$$

donde $h = 1$ en ausencia de cooperatividad como se considera en [107]. Ahora, en nuestro modelo sencillo, el rango de $[\text{IP}_3]$ sobre el cual la relación entre $\langle T_{IS} \rangle - T_{min}$ e $[\text{IP}_3]$ es aproximadamente exponencial ($[0,172\mu\text{M}, 0,179\mu\text{M}]$ o $[0,179\mu\text{M}, 0,186\mu\text{M}]$ en la Fig. 5.5 (b), para los que el rango de tiempos medios entre pulsos es relativamente amplio) supone como máximo una variación del 4% en $[\text{IP}_3]$. Por lo tanto, una linealización de la Ec. (5.11) sería relativamente buena en cada uno de estos rangos garantizando la dependencia exponencial entre $\langle T_{IS} \rangle - T_{min}$ e I_{ext} también. Podríamos también preguntarnos si la función dada por la Ec. (5.11) no está saturada ($I_{ext} \gg K_D$). Es poco probable que la saturación en la producción de IP_3 se alcance para una situación con menos del 20% de ocupación de los IP_3Rs (representada por β en nuestro modelo). Por otro lado, se estimó que $[\text{IP}_3]$ en los ovocitos de *X. Laevis* varía entre 40nM en condiciones basales y $\sim 1,8\mu\text{M}$ tras una fuerte estimulación [68]. Es por lo tanto razonable suponer que la relación entre $[\text{IP}_3]$ e I_{ext} es aproximadamente lineal ($[\text{IP}_3] \approx [\text{IP}]_0 I_{ext}/K_D$ de acuerdo a la Ec. (5.11) con $h = 1$) para situaciones, como las analizadas en este Capítulo, en las que la dinámica del Ca^{2+} intracelular mediada por IP_3Rs es excitable. Por lo tanto, bajo estas condiciones, encontramos una explicación mecanicista para la dependencia exponencial de $\langle T_{IS} \rangle - T_{min}$ con I_{ext} , como se observa experimentalmente en [103] (ver Fig. 1.10).

Aunque hemos obtenido este resultado para el caso de los pulsos de Ca^{2+} , la generalidad y sencillez del modelo utilizado permite suponer que es posible aplicar un enfoque similar a otros casos, teniendo en cuenta, además, la presencia generalizada de dinámicas excitables en biología. Los candidatos que podrían describirse con este enfoque incluyen FTs que muestran pulsos de translocación nuclear, para las cuales existen varios signos de procesos inducidos por ruido. Ya mencionamos el ejemplo de este comportamiento en la respuesta al estrés en levaduras mediada por el FT, Msn2 [78], ilustrado en la Sec. 1.5. Otros ejemplos incluyen la diferenciación en bacterias que puede describirse correctamente con un sistema excitable activado por ruido [32] o el comportamiento del supresor de tumores en mamíferos, p53 [6], cuya dinámica requiere una estructura de red excitable subyacente [73]. Tener un marco teórico para dar cuenta de la dependencia exponencial entre la intensidad del estímulo externo y el tiempo

medio entre pulsos de un mediador pulsátil de la señalización proporciona la base para suponer que dicha dependencia puede ser muy generalizada y que, por lo tanto, las células cuentan con un sistema de codificación de su entorno cualitativamente distinto del que les provee la codificación por amplitud como discutimos en el Capítulo anterior.

Capítulo 6

Discusión y conclusiones generales de la Tesis.

Esta Tesis aborda un tema interdisciplinario en el que la física y las personas formadas en dicha disciplina han hecho contribuciones clave: la transmisión de información en sistemas biológicos. Podemos mencionar acá los primeros intentos de Schödinger de determinar los portadores de la herencia, el descubrimiento de la estructura del ADN o los numerosos esfuerzos que crecieron enormemente junto con el proyecto de secuenciación del genoma humano. Quienes trabajan en física tienden a simplificar las complejidades de las estructuras y funciones biológicas, lo que limita el tipo de preguntas de interés biológico que pueden contestarse. Las descripciones simples, sin embargo, pueden revelar conexiones profundas e inspirar futuros desarrollos y descubrimientos. Esta Tesis se inscribe en este tipo de abordajes que, usando modelos minimales, busca contestar preguntas sobre los rasgos genéricos comunes a varios sistemas específicos distintos. En particular, en este caso la motivación del estudio fue la de indagar las diferencias y similitudes entre las diversas formas en que las células codifican los cambios en el medioambiente para generar respuestas acordes. Más aún, se buscó exponer cómo el acceso a una dinámica temporal les da a las células mayor sensibilidad y flexibilidad ante su entorno y así mostrar algunas de las muchas ventajas que proveen los comportamientos dinámicos de no equilibrio.

En esta Tesis se usaron herramientas de la teoría de la información de Shannon para estudiar, mediante cálculos analíticos y simulaciones numéricas, algunos aspectos del modo en que las células “codifican” los cambios del ambiente para actuar sobre ellos. Inspirándose en los desarrollos previos de Bialek y colaboradores (ver 1.7.3), la Tesis buscó comparar los dos grandes modos de codificación del entorno: por amplitud y por frecuencia. Mientras que en el primer caso intensidades crecientes de un estímulo externo (la concentración de un ligando externo en los problemas abordados en la Tesis) se “mapean” en concentraciones o nivel de activación crecientes de los intermediarios de la señalización, en el segundo se induce, en algún intermediario, una dinámica pulsátil

de frecuencia creciente con la intensidad del estímulo. Como mostramos en la Introducción, existen distintos tipos de ejemplos en donde ambas estrategias se combinan. Por un lado, hay situaciones en las que un mismo intermediario, en particular, un Factor de Transcripción (FT), puede presentar distintas dinámicas dependiendo del estímulo externo y generar distintas respuestas dependiendo de dicha dinámica. Esta combinación permite la *multiplexación dinámica* del procesamiento de información (ver *e.g.*, 1.5). Por otro lado, hay ejemplos de células que pueden codificar un mismo estímulo en amplitud o en frecuencia, probablemente para activar distintas vías de generación de respuestas (1.5). El trabajo contenido en el Capítulo 3 que estudia el paso transcripcional que va desde la componente nuclear de los FTs hasta la producción del ARNm permitió establecer límites al uso de la multiplexación dinámica cuando hay un único FT involucrado en el procesamiento de la información (ver más adelante). El trabajo contenido en el Capítulo 4 que agrega a este paso transcripcional el que va desde el estímulo externo hasta la fracción nuclear del FT llevó a la conclusión de que, debido a la diferente forma (matemática) en la que se mapea el entorno cuando sus cambios están codificados en la amplitud o en la frecuencia de los intermediarios, la capacidad de transmisión de información por amplitud funciona correctamente en un rango limitado de intensidades del estímulo, mientras que la de frecuencia puede funcionar adecuadamente en rangos mucho más amplios. Esto permite suponer que la combinación del uso de ambas estrategias para responder a un mismo estímulo puede servir para ampliar el rango de detección de gradientes de intensidad. La Tesis se completa con los Capítulos 2 y 5 que se centran en análisis más detallados de la codificación por frecuencia. En el Capítulo 2 se analiza la información sobre un ligando externo que puede codificarse en una secuencia de pulsos durante un tiempo dado dependiendo de si los pulsos siguen una distribución de Poisson o si el tiempo entre pulsos incluye un período de recuperación de la inhibición. En el Capítulo 5 se deriva a partir de un modelo de excitabilidad sencillo la dependencia exponencial entre el tiempo medio entre pulsos y la intensidad del estímulo externo que está en la base de que la codificación por frecuencia sea cualitativamente distinta de la codificación por amplitud, como se muestra en el Capítulo 4.

Con más detalle, en el Capítulo 2 se estudió la información mutua, MI, entre la concentración de un ligando externo y el tiempo entre pulsos o el número de pulsos durante un tiempo dado de un intermediario pulsátil de la señalización. Para tal fin se usó, por un lado, un modelo sencillo introducido previamente en la literatura para describir la dinámica de pulsos de Ca^{2+} intracelular que está caracterizado por una tasa de disparo “desnuda” constante, λ , y otra de recuperación de la inhibición, ρ . Por el otro lado, en base a propiedades observadas de los pulsos de Ca^{2+} , se consideró que el tiempo medio entre pulsos dependía exponencialmente de la concentración del ligando externo. Se analizó en cuánto podía diferir MI dependiendo del cociente λ/ρ , el que determina la distribución del tiempo entre pulsos (exponencial si $\lambda \gg \rho$ y de Gompertz-Makeham en el límite opuesto). Se encontró que MI es mayor en el límite dominado por la tasa de recuperación, pudiendo diferir hasta en $\sim 1\text{bit}$ respecto del caso con $\lambda \gg \rho$. Se encontró también, sin embargo, que la sola variación de ρ no modifica MI.

Esto significa que regular únicamente la tasa de recuperación podría cambiar la frecuencia media entre los pulsos sin modificar la información que es posible extraer del ambiente.

En el Capítulo 3 se estudió la MI entre la duración, amplitud o frecuencia entre pulsos de la fracción nuclear de un FT y el ARNm transcripto acumulado. Para tal fin se usó un modelo para la transcripción sencillo propuesto previamente en la literatura en base a observaciones del FT Msn2 en células de levadura y del que presentamos una derivación a partir de una descripción mecanística. El FT Msn2 es tal que responde con distintas dinámicas de traslocación nuclear dependiendo del tipo de estímulos a los que la célula es sometida. La limitación de glucosa induce pulsos de acumulación nuclear de frecuencia creciente cuanto más baja sea la concentración de glucosa. El estrés osmótico induce un pulso de localización nuclear de duración creciente con la intensidad del shock. El estrés oxidativo induce un pulso de amplitud creciente con la magnitud del estrés. En el Capítulo 3 se compararon estos tres modos de codificación buscando determinar si cada uno de ellos era óptimo en regiones distintas de los parámetros del modelo de transcripción. Esta búsqueda permitió concluir que los tres modos de codificación maximizan MI en regiones similares de los parámetros: aquellas con tasas de transcripción alta y con relaciones FT-promotor “empinadas” y con umbral alto (n y K_d suficientemente altos en ec. 3.1). Encontramos que la mayor diferencia entre los tres modos radica en su distinta sensibilidad a cambios en los parámetros del modelo de transcripción. En este sentido, la codificación por frecuencia es la que muestra mayor sensibilidad y la codificación por duración, la menor. Cada conjunto de parámetros del modelo transcripcional puede asociarse a un promotor diferente. La distinta sensibilidad de los distintos modos es la que nos permitió encontrar promotores (*i.e.*, conjuntos de valores de parámetros) que son “ciegos” a los entradas moduladas por frecuencia pero que son buenos para transmitir información para otros tipos de modulación. Fue imposible encontrar, sin embargo, la situación opuesta. Esto permite explicar las observaciones de [47] ilustradas en la Fig. 1.24. En esa figura se compara la transmisión de información debida a la transcripción cuando el FT Msn2 induce la expresión del gen involucrado en la respuesta a estrés osmótico (SIP18) con la situación en la que se induce el gen involucrado en la privación de glucosa (HXK1). En condiciones fisiológicas, las vías en las que participan estos genes codifican la intensidad del estímulo externo en duración y frecuencia, respectivamente (ver Fig. 1.17). Dado que en los experimentos es posible manipular *ad libitum* la dinámica de la traslocación nuclear del FT, en [47] se analizó la capacidad de ambas vías de transmitir información codificada en la amplitud y en la frecuencia del FT nuclear. En coincidencia con nuestro estudio teórico, los resultados obtenidos en células silvestres mostraron que, ante la modulación por amplitud, la información transmitida fue similar para los dos genes (es decir, poco sensible a cambios en los parámetros del modelo transcripcional o, equivalentemente, a un cambio de promotor). Para la modulación por frecuencia, por otro lado, la información transmitida por el HXK1 (el involucrado en la vía que codifica por frecuencia) fue del mismo orden mientras que la transmitida por el SIP18 fue despreciable (lo que refleja la mayor sensibilidad a los parámetros de

la codificación por frecuencia). Por otro lado, en [47] se repitieron los experimentos usando versiones mutadas del gen SIP18 encontrando que la información transmitida por amplitud era mayor en este caso que en las células silvestres. En [47] los autores se preguntaron por qué el gen de las cepas salvajes no estaba optimizado para la transmisión de información por amplitud, a pesar de estar involucrado en la vía que “usa” esta codificación. Nuestros estudios parecen indicar que, al menos en este caso, el promotor podría haber sido seleccionado de forma de ser lo más “ciego” posible para la modulación por frecuencia sin perder la capacidad de distinguir ON de OFF en el caso de modulación por amplitud. Los resultados de este Capítulo muestran también cómo la distinta sensibilidad de los diferentes modos de codificación pone límites a la multiplexación dinámica para el uso de un mismo FT con distintos fines.

En el Capítulo 4 se calculó MI entre la intensidad del estímulo externo y el ARNm acumulado cuando la primera se codifica en la amplitud o frecuencia medias de los pulsos del FT nuclear. Para tal fin introdujimos una dependencia entre la intensidad del estímulo y la amplitud o la frecuencia acá mencionadas. Mientras que para la primera relación usamos una ecuación de Hill, en el segundo caso, basándonos en numerosas observaciones experimentales, supusimos una dependencia exponencial entre el tiempo medio entre pulsos y la intensidad del estímulo. Estas dos formas de mapear los estímulos externos dan como resultado capacidades de transmisión de información cualitativamente diferentes. La codificación por amplitud es óptima para un rango limitado de intensidades alrededor de un valor determinado. La transmisión de información para la codificación por frecuencia mejora o permanece relativamente constante a medida que aumenta la intensidad del estímulo. La combinación de ambas estrategias podría servir para expandir el rango de detección de gradientes de intensidad. En el Capítulo 4 se discutió el caso de la respuesta a feromona en células de levadura donde coexisten la codificación por amplitud y por frecuencia como un posible ejemplo donde esta combinación podría ser de utilidad.

Finalmente, en el Capítulo 5 derivamos la dependencia exponencial entre el tiempo entre pulsos y el estímulo externo para un modelo sencillo de la dinámica del Ca^{2+} intracelular mediada por IP_3Rs , en presencia de ruido, para parámetros en que dicha dinámica es excitable. Usando observaciones experimentales para cuantificar el ruido y relacionar parámetros del modelo con el estímulo externo mostramos que la dependencia exponencial obtenida era válida sobre un rango tal que el valor de los tiempos correspondientes variaban sobre un intervalo igual o mayor al de las observaciones experimentales. Si bien este resultado fue formalmente derivado para el caso del Ca^{2+} , la generalidad y sencillez del modelo excitable utilizado permite suponer que es posible aplicar un enfoque similar a otros casos, en particular, al comportamiento pulsátil de FTs, para el que existe abundante evidencia de que, en muchos casos, se rige por una dinámica excitable.

Apéndice A

Cálculos suplementarios del Capítulo 2

El modelo que abordamos en el Capítulo 2 corresponde el proceso, río arriba, de transducción del estímulo externo en un intermediario pulsátil de la vía de señalización. Más específicamente, nos concentramos en las observaciones de pulsos intracelulares de Ca^{2+} en células estimuladas con ligandos externos. Para ello usamos un modelo caracterizado por dos parámetros: la tasa de disparo “desnuda”, λ , y la tasa de recuperación de la inhibición, ρ . En el Capítulo comparamos el comportamiento del modelo en dos límites que presentan una estadística entre pulsos diferente: al que llamamos de Poisson (donde $x = \lambda/\rho \ll 1$) y al que llamamos de Feedback Global Negativo o GNF (donde $x \gg 1$). En este Apéndice presentamos varios cálculos suplementarios en los que se basan algunos de los resultados mostrados en el Capítulo.

A.1. Propiedades del modelo en los límites de Poisson y GNF

El modelo presentado en esta sección de la tesis se caracteriza, por un lado, por la siguiente densidad de probabilidad de que, dado que se produjo un pulso en el tiempo 0, el primer pulso posterior se produzca en el tiempo $t + T_{min}$, como está mencionado en 1.8.

En el modelo de [90] y aquí, el tiempo entre picos es la suma, $t + T_{min}$, donde T_{min} es un componente determinista (fijo) y t un uno estocástico. El otro componente de nuestro modelo es la relación entre $T \equiv \langle t \rangle_{t|\lambda}$ y la concentración de ligando externo, C . Con este fin, usamos la relación presentada en [103] para explicar varias observaciones experimentales:

$$T \equiv \langle t \rangle_{t|\lambda} = A \exp(-\gamma C), \quad (\text{A.1})$$

donde $\langle \cdot \rangle_{t|\lambda}$ representa la media sobre la distribución, $p(t|\lambda)$. Aquí derivamos

algunas propiedades del modelo en los dos límites que nos interesan: el de *Global Negative Feedback* (GNF) que está definido por $x \equiv \lambda/\rho \gg 1$ y el límite de Poisson que está definido por $x \ll 1$.

A.1.1. Relación entre la tasa de “disparo”, λ , y la concentración de ligando externo, C .

Mostramos aquí que la Ec. (A.1) implica que:

$$\lambda = \alpha \exp(\beta C), \quad (\text{A.2})$$

en los límites GNF y Poisson, con α y β parámetros fijos que pueden ser funciones de ρ .

Y es claro que Ec. (1.8), para un dado valor de λ y en el límite $x \ll 1$, el modelo se reduce a un proceso de Poisson para cuando la distribución, $p(t|\lambda)$, es exponencial con $T = 1/\lambda$.

Introduciendo esta última relación Ec. (A.1) obtenemos:

$$\lambda = \frac{1}{A} e^{BC}, \quad \text{for } x \ll 1. \quad (\text{A.3})$$

Para relacionar λ y C en el límite $x \gg 1$ (GNF) usamos el resultado de [90] el cual muestra que la media, T , para el modelo de Ec. (1.8) está dada por:

$$T = \frac{e^x x^{1-x}}{\lambda} (\Gamma(x) - \Gamma(x, x)), \quad (\text{A.4})$$

donde $\Gamma(x)$ es la función Gamma, $\Gamma(s, t)$ es la función Γ incompleta por arriba y $x = \lambda/\rho$. La diferencia en el r.h.s de la Ec. (A.4) puede ser reescrita como:

$$\Gamma(x) - \Gamma(x, x) = \gamma(x, x), \quad (\text{A.5})$$

con $\gamma(s, t)$ la función Γ incompleta por abajo. Ahora probamos que:

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^x} \gamma(x+1, x) - \sqrt{\frac{\pi}{2}} x = 0. \quad (\text{A.6})$$

De hecho, usando el método generalizado de Laplace, obtenemos:

$$\gamma(x+1, x) = \int_0^x z^x e^{-z} dz \sim x^x e^{-x} \int_0^\infty e^{-\frac{1}{2x}(z-x)^2} dz = \frac{1}{2} x^x e^{-x} \sqrt{2\pi x}, \quad (\text{A.7})$$

para $x = \lambda/\rho$ suficientemente grande para el cual Ec. (A.6) sigue inmediatamente.

Esto, por su lado, implica que:

$$\gamma(x, x) = x^{x-1} e^{-x} + x^{-1} \gamma(x+1, x) \approx x^{x-1} e^{-x} + x^{x-1} e^{-x} \sqrt{\frac{\pi x}{2}}, \quad x \gg 1. \quad (\text{A.8})$$

Combinado Ecs. (A.8), (A.4) y (A.5) obtenemos:

$$T = \frac{1}{\lambda} \left(1 + \sqrt{\frac{\pi x}{2}} \right) \approx \sqrt{\frac{\pi x}{2\lambda^2}} = \sqrt{\frac{\pi}{2\lambda\rho}}, \quad x \gg 1. \quad (\text{A.9})$$

Combinando Ecs. (A.1) y (A.9) obtenemos:

$$\lambda = \frac{\pi}{2\rho A^2} e^{2\gamma C}, \quad \text{for } x \gg 1. \quad (\text{A.10})$$

Eqs. (A.3) y (A.10) implican que, en ambos límites, se da:

$$\lambda = \alpha e^{\beta C}, \quad (\text{A.11})$$

con $\alpha = \pi/(2\rho A^2)$ y $\beta = 2\gamma$ para el límite GNF y $\alpha = 1/A$ y $\beta = B$ para el límite de Poisson.

Relación entre la media y la desviación estándar de la componente estocástica del tiempo entre pulsos en los límites de GNF y de Poisson

Como se muestra en [90], la media de la parte estocástica del tiempo entre pulsos, T , viene dada por la ecuación (A.4), mientras que su varianza viene dada por:

$$\sigma^2 = \langle t^2 \rangle - T^2 = 2e^x {}_2F_2 \left(\begin{matrix} x & x \\ 1+x & 1+x \end{matrix}; -x \right) - T^2, \quad (\text{A.12})$$

donde ${}_2F_2$ es la función hipergeométrica generalizada (2,2). Para ambos límites del modelo, la media de la parte estocástica, T , y la desviación estándar, σ , están relacionadas por $\sigma = kT$, como se menciona en 1.4:

Esta relación es trivial en el límite de Poisson ($x \ll 1$) para el cual $T = \sigma = 1/\lambda$ y $k = 1$. Fuera de este límite, la relación se obtuvo numéricamente en [90]. Como mostramos ahora, se puede derivar analíticamente para $x \gg 1$ y producir $k = \sqrt{\frac{4}{\pi} - 1}$.

Para determinar $k = \sigma/T$ para el límite $x \gg 1$ calculamos:

$$1 + \frac{\sigma^2}{T^2} = \lim_{x \rightarrow \infty} \frac{\langle t^2 \rangle}{T^2} = \lim_{x \rightarrow \infty} \frac{2e^x {}_2F_2 \left(\begin{matrix} x & x \\ 1+x & 1+x \end{matrix}; -x \right)}{(e^x x^{1-x} (\Gamma(x) - \Gamma(x, x)))^2}. \quad (\text{A.13})$$

Ya probamos que $T = e^x x^{1-x} (\Gamma(x) - \Gamma(x, x)) \approx \sqrt{\pi x/2} + 1$ para $x \gg 1$ (mirar Ec. (A.4)).

Ahora computamos el numerador Ec. (A.13). Para este objetivo usamos la transformación integral de Euler (ver *e.g.*[93]):

$${}_{A+1}F_{B+1} \left(\begin{matrix} a_1 & \dots & a_A & c \\ b_1 & \dots & b_B & d \end{matrix}; z \right) = \frac{\Gamma(d)}{\Gamma(c)\Gamma(d-c)} \int_0^1 t^{c-1} (1-t)^{d-c-1} {}_A F_B \left(\begin{matrix} a_1 & \dots & a_A \\ b_1 & \dots & b_B \end{matrix}; tz \right) dt. \quad (\text{A.14})$$

Tomando en cuenta que $\frac{\Gamma(x+1)}{\Gamma(x)\Gamma(1)} = x$, aplicando esta fórmula dos veces para ir de ${}_0F_0$ a ${}_2F_2$ obtenemos:

$${}_2F_2 \left(\begin{matrix} x & x \\ 1+x & 1+x \end{matrix}; -x \right) = x^2 \int_0^1 \int_0^1 (tt')^{x-1} {}_0F_0 (; -tt'x) dt dt'. \quad (\text{A.15})$$

Dado que ${}_0F_0 (; -tt'x) = e^{-tt'x}$, Ec. (A.15) deviene:

$${}_2F_2 \left(\begin{matrix} x & x \\ 1+x & 1+x \end{matrix}; -x \right) = x^2 \int_0^1 \int_0^1 (tt')^{x-1} e^{-tt'x} dt dt'. \quad (\text{A.16})$$

Esta ecuación puede ser resuelta via la sustitución:

$$u = \ln \left(\sqrt{\frac{t}{t'}} \right), \quad (\text{A.17a})$$

$$v = \sqrt{tt'}, \quad (\text{A.17b})$$

cuyo Jacobiano es $2|v|$.

El límite de integración, por otro lado, se transforma en $0 < v < 1$, $\ln(v) < u < -\ln(v)$. Por lo tanto, la integral, en el r.h.s de Ec. (A.16) puede ser reescrita como:

$$\begin{aligned} \int_0^1 \int_0^1 (tt')^{x-1} e^{-tt'x} dt dt' &= \int_0^1 \int_{\ln(v)}^{-\ln(v)} (v^2)^{x-1} e^{-v^2x} 2v dudv = \\ &\int_0^1 (v^2)^{x-1} e^{-v^2x} 2v (-2 \ln(v)) dv. \end{aligned} \quad (\text{A.18})$$

Reescribiendo $2 \ln(v) = \ln(v^2)$ e introduciendo la sustitución $w = v^2$, puede ser reescrito Ec. (A.18) como:

$$\int_0^1 \int_0^1 (tt')^{x-1} e^{-tt'x} dt dt' = \int_0^1 (w)^{x-1} e^{-wx} (-\ln(w)) dw \quad (\text{A.19})$$

Por lo tanto, el límite de Ec. (A.13) es Ecuivalente a:

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{2e^x x^2 \int_0^1 (w)^{x-1} e^{-wx} (-\ln(w)) dw}{\frac{\pi}{2}x + \sqrt{2\pi x} + 1} &= \\ \lim_{x \rightarrow \infty} \frac{4x \int_0^1 (w)^{x-1} e^{(1-w)x} (-\ln(w)) dw}{\pi}. \end{aligned} \quad (\text{A.20})$$

Por lo tanto, probando que

$$\lim_{x \rightarrow \infty} x \int_0^1 (w)^{x-1} e^{(1-w)x} (-\ln(w)) dw = 1, \quad (\text{A.21})$$

es suficiente para demostrar que $1 + \frac{\sigma^2}{T^2} \rightarrow \frac{4}{\pi}$ para $x \gg 1$ y, por tanto, que $k = \sigma T = \sqrt{\frac{4}{\pi} - 1} \sim 0,523$. Para probar la Ec. (A.21) usamos el método de

Laplace modificado. Con esto en mente, podemos observar que el integrando en esta ecuación tiende a 0 para cada valor de w fuera de una vecindad de $w = 1$. Por lo tanto, podemos usar la aproximación $-\frac{\ln(w)}{w} \sim 1 - w + \mathcal{O}((1 - w)^2)$. Insertando en Ec. (A.21) obtenemos:

$$x \int_0^1 w^{x-1} e^{(1-w)x} (-\ln(w)) dw \sim \int_0^1 w^x e^{(1-w)x} x(1-w) dw = \int_0^1 e^{x(\ln(w)+(1-w))} x(1-w) dw. \quad (\text{A.22})$$

Aproximando $\ln(w) + (1 - w) \sim -\frac{1}{2}(w - 1)^2$ llegamos a:

$$x \int_0^1 (w)^{x-1} e^{(1-w)x} (-\ln(w)) dw \sim \int_0^1 e^{-\frac{1}{2}x(w-1)^2} x(1-w) dw, \quad (\text{A.23})$$

que se puede resolver induciendo el cambio de variable $y = \frac{1}{2}x(w - 1)^2$. De esta manera obtenemos

$$x \int_0^1 (w)^{x-1} e^{(\ln(w)+(1-w))x} (-\ln(w)) dw \approx \int_0^{\frac{1}{2}x} e^{-y} dy \approx \int_0^{\infty} e^{-y} dy = 1, \quad (\text{A.24})$$

y así probamos Ec. (A.21).

En resumen hemos probado aquí que en los límites de la GNF y Poisson T y σ están relacionados por la Ec. (1.4) con $k = 1$ en el caso de Poisson y $k = \sqrt{\frac{4}{\pi}} - 1$ en el GNF.

A.2. Cuantificación de la información contenida en un pulso.

A.2.1. Detalles de Cálculos

Aquí mostramos el cálculo de la información contenida en el tiempo de separación entre dos pulsos sucesivos. Más específicamente, calculamos la información mutua entre t y C que es la misma que la información mutua entre t y λ 2.3 y 2.4.

Ahora, dado que λ es una función de la concentración del ligando, que normalmente cambia con el tiempo, λ también cambiará. Los cálculos que vamos a realizar son buenos para λ constante o si λ varía lo suficientemente lento con respecto a $\langle t \rangle|_{t|\lambda}$. De esta forma, usamos la Ec. (A.2) para relacionar la densidad de probabilidad λ , p_λ , y la de la concentración del ligando externo, p_C , a través de:

$$p_\lambda(\lambda) = \frac{1}{\alpha\beta} e^{-\beta C} p_C(C) = \frac{1}{\lambda\beta} p_C\left(\frac{1}{\beta} \ln\left(\frac{\lambda}{\alpha}\right)\right). \quad (\text{A.25})$$

Primero computamos $H(t)$. Asumiendo que λ es constante, cuando escribimos:

$$\int_0^t -\lambda(1 - \exp(-\rho t')) dt' = -\lambda t - \frac{\lambda}{\rho} (\exp(-\rho t) - 1), \quad (\text{A.26})$$

y

$$p(t|\lambda) = -\frac{\partial e^{\lambda(-t - \frac{1}{\rho}(e^{-\rho t} - 1))}}{\partial t}. \quad (\text{A.27})$$

Definiendo $\tau = (-t - \frac{1}{\rho}(e^{-\rho t} - 1))$ cuando reescribimos

$$p_t(t) = \int -\frac{\partial e^{\lambda\tau(t)}}{\partial t} p_\lambda(\lambda) d\lambda = -\frac{\partial M_\lambda(\tau(t))}{\partial t} \equiv -M'_\lambda, \quad (\text{A.28})$$

donde introducimos la definición:

$$M_\lambda(\tau) = \int e^{\lambda\tau} p_\lambda(\lambda) d\lambda, \quad (\text{A.29})$$

la cual es la función generadora de momentos.

Luego reescribimos H en términos de M_λ como:

$$H(t) = \frac{1}{\ln(2)} \left(-\int -M'_\lambda(\tau) \ln(-M'_\lambda(\tau)) dt \right), \quad (\text{A.30})$$

donde la prima indica la derivada con respecto a t (ver Ec. (A.28)). Dado que $-M'_\lambda(\tau) = -\frac{\partial M_\lambda(\tau)}{\partial \tau} \tau'(t) = \frac{\partial M_\lambda(\tau)}{\partial \tau} (-\tau'(t))$ reescribimos Ec. (A.30) como:

$$H(t) = -\frac{1}{\ln(2)} \left(\int -M'_\lambda(\tau) \ln \left(\frac{\partial M_\lambda(\tau)}{\partial \tau} \right) dt + \int -M'_\lambda(x) \ln(-\tau'(t)) dt \right). \quad (\text{A.31})$$

El último término en esta ecuación puede ser escrita como:

$$\begin{aligned} & \frac{1}{\ln(2)} \int M'_\lambda(\tau) \ln(-\tau'(t)) dt = \\ & -\frac{1}{\ln(2)} \int p_t(t) \ln(1 - e^{-\rho t}) dt \equiv -\frac{1}{\ln(2)} \langle \ln(1 - e^{-\rho t}) \rangle|_t. \end{aligned} \quad (\text{A.32})$$

Cambiando más aún la variable de integración de t a $\tau = -t - 1/\rho(\exp(-\rho t) - 1)$, obtenemos:

$$H(t) = \frac{1}{\ln(2)} \left(-\int_{-\infty}^0 \frac{\partial M_\lambda(\tau)}{\partial \tau} \ln \left(\frac{\partial M_\lambda(\tau)}{\partial \tau} \right) d\tau - \langle \ln(1 - e^{-\rho t}) \rangle|_t \right), \quad (\text{A.33})$$

donde usamos $\tau(t=0) = 0$ y $\tau(t \rightarrow \infty) \rightarrow -\infty$.

Ahora computamos $H(t|\lambda)$.

Tomando en cuenta Ecs. (1.8) y (A.26) escribimos:

$$\begin{aligned} H(t|\lambda) &= \int p(t|\lambda) \log_2(p(t|\lambda)) dt = \\ & \frac{1}{\ln(2)} \left(\langle \ln(\lambda(1 - e^{-\rho t})) \rangle|_{t|\lambda} - \lambda \langle t \rangle|_{t|\lambda} - \frac{\lambda}{\rho} (\langle e^{-\rho t} \rangle|_{t|\lambda} - 1) \right) = \\ & \frac{1}{\ln(2)} \left(\langle \ln((1 - e^{-\rho t})) \rangle|_{t|\lambda} + \ln(\lambda) - \lambda \langle t \rangle|_{t|\lambda} - \frac{\lambda}{\rho} (\langle e^{-\rho t} \rangle|_{t|\lambda} - 1) \right). \end{aligned} \quad (\text{A.34})$$

El primer término en esta ecuación se cancela con el similar que proviene de $H(t)$ al calcular I (ver Ecs. (2.3), (A.31) y (A.34)). Por otro lado,

$$-\lambda \langle t \rangle_{|t|\lambda} - \frac{\lambda}{\rho} (\langle e^{-\rho t} \rangle_{|t|\lambda} - 1) = - \int \frac{\partial e^{\lambda\tau(t)}}{\partial t} \lambda\tau(t) dt \quad (\text{A.35})$$

Mediante integración por partes, y luego sustitución por $\lambda\tau$:

$$\begin{aligned} & - \int_0^\infty \frac{\partial e^{\lambda\tau(t)}}{\partial t} \lambda\tau(t) dt = \\ & - e^{\lambda\tau} \lambda\tau|_0^{-\infty} + \int_0^\infty e^{\lambda\tau(t)} \lambda\tau'(t) dt = \int_0^{-\infty} e^u du = 1 \end{aligned} \quad (\text{A.36})$$

Finalmente, así:

$$I(\lambda, t) = \frac{1}{\ln(2)} \left(- \int_{-\infty}^0 \frac{\partial M_\lambda}{\partial \tilde{\tau}} \ln \left(\frac{\partial M_\lambda}{\partial \tilde{\tau}} \right) d\tilde{\tau} - 1 + \langle \ln(\tilde{\lambda}) \rangle \right). \quad (\text{A.37})$$

Dependencia de la información sobre los parámetros del problema

Si elegimos $\mu = \alpha$, para definir las variables sin dimensión, la ecuación (A.11) implica que $\tilde{\alpha} = e^{\beta C}$. Las ecuaciones (A.37) y (A.29) implican que, al menos en los dos límites que estamos estudiando, I no depende de α . Similarmente, si elegimos $\mu = e^{\beta \langle C \rangle}$, esto es $\tilde{\lambda} = e^{\beta(C - \langle C \rangle)}$, concluimos que I no depende de $\langle C \rangle$. La dependencia de ρ es más sutil y la discutimos más adelante. Para estudiar la dependencia de β , ahora hacemos una suposición para la distribución de C . Es decir, asumimos una distribución uniforme en el intervalo $[0, C_M]$, como se la menciona en 2.5, de tal manera que

$$p_\lambda(\lambda) = \begin{cases} \frac{1}{C_M \beta \lambda}, & \text{si } \alpha \leq \lambda \leq \alpha \exp(\beta C_M), \\ 0, & \text{de otra manera.} \end{cases} \quad (\text{A.38})$$

Y luego, $\partial M_\lambda / \partial \tilde{\tau}$ (Eq. (2.7) con $\tilde{\tau} = \mu\tau$, $\mu = \alpha$) es dado por:

$$\frac{\partial M_\lambda}{\partial \tilde{\tau}} = \int_1^{\exp(\beta C_M)} e^{\tilde{\lambda} \tilde{\tau}} \frac{1}{C_M \beta} d\tilde{\lambda} = \frac{1}{C_M \beta \tilde{\tau}} (\exp(\tilde{\tau} e^{\beta C_M}) - \exp(\tilde{\tau})) \quad (\text{A.39})$$

Por otro lado:

$$\langle \ln(\tilde{\lambda}) \rangle = \beta \langle C \rangle = \beta \frac{C_M}{2} \quad (\text{A.40})$$

Entonces, lo único que queda por calcular I en Ec. (A.37) es:

$$HM'_\lambda \equiv - \int_{-\infty}^0 \frac{\partial M_\lambda}{\partial \tilde{\tau}} \ln \left(\frac{\partial M_\lambda}{\partial \tilde{\tau}} \right) d\tilde{\tau}. \quad (\text{A.41})$$

Definimos

$$g(\tilde{\tau}) \equiv \beta C_M \frac{\partial M_\lambda}{\partial \tilde{\tau}} = \frac{e^{\tilde{\tau} e^{C_M \beta}} - e^{\tilde{\tau}}}{\tilde{\tau}}, \quad (\text{A.42})$$

que puede ser reescrito como:

$$g(\tilde{\tau}) = \frac{2}{\tilde{\tau}} e^{\tilde{\tau} \frac{e^{\beta C_M} + 1}{2}} \sinh\left(\tilde{\tau} \frac{e^{\beta C_M} - 1}{2}\right). \quad (\text{A.43})$$

Y por tanto:

$$\ln(g(\tilde{\tau})) = \ln\left(\frac{\partial M_\lambda}{\partial \tilde{\tau}}\right) + \ln(\beta C_M) = \ln(2) - \ln(\tilde{\tau}) + \tilde{\tau} \frac{e^{\beta C_M} + 1}{2} + \ln\left(\sinh\left(\tilde{\tau} \frac{e^{\beta C_M} - 1}{2}\right)\right). \quad (\text{A.44})$$

Dados Ecs. (A.28)-(2.7) y que $\tilde{\tau}(t=0) = 0$ y $\tilde{\tau}(t \rightarrow \infty) \rightarrow -\infty$, entonces

$$-\int_0^\infty f(t) \frac{\partial M_\lambda}{\partial t} dt = \int_{-\infty}^0 f(t(\tilde{\tau})) \frac{\partial M_\lambda}{\partial \tilde{\tau}} d\tilde{\tau} = \int_0^\infty f(t) p_t(t) dt = \langle f \rangle_t, \quad (\text{A.45})$$

para cualquier función $f(t)$ (o de $\tilde{\tau}$ a través de la relación $\tilde{\tau} = \mu(-t - \frac{1}{\rho}(e^{-\rho t} - 1))$).

Por lo tanto, I puede ser reescrito como:

$$I(\lambda, t) = \frac{1}{\ln(2)} \left(-\int_{-\infty}^0 \frac{\partial M_\lambda}{\partial \tilde{\tau}} \ln(g(\tilde{\tau})) d\tilde{\tau} + \ln(\beta C_M) - 1 + \beta \frac{C_M}{2} \right) = \frac{1}{\ln(2)} \left(-\ln(2) + \langle \ln(\tilde{\tau}) \rangle_t - \langle \tilde{\tau} \rangle_t \frac{e^{\beta C_M} + 1}{2} - \langle \ln\left(\sinh\left(\tilde{\tau} \frac{e^{\beta C_M} - 1}{2}\right)\right) \rangle_t + \ln(\beta C_M) - 1 + \beta \frac{C_M}{2} \right). \quad (\text{A.46})$$

Calculando cada término de esta ecuación, teniendo especial cuidado con el límite de integración $\tilde{\tau} = 0$, llegamos a la siguiente expresión

$$I(\lambda, t) = \frac{1}{\ln(2)} \left(\frac{\sinh(\beta C_M)}{\beta C_M} + \ln\left(\beta \frac{C_M}{2}\right) - \ln\left(\frac{e^{\beta C_M} - 1}{2}\right) - 1 + \beta \frac{C_M}{2} \right) - \frac{1}{\beta C_M \ln(2)} \int_{-\infty}^0 \frac{e^\zeta}{\zeta} \ln\left(\frac{e^{\beta C_M} \sinh(\zeta e^{-\beta C_M} \frac{e^{\beta C_M} - 1}{2})}{\sinh(\zeta \frac{e^{\beta C_M} - 1}{2})}\right) d\zeta, \quad (\text{A.47})$$

cuyo último término calculamos numéricamente usando Mathematica. El resultado de este cálculo numérico se muestra en el cuerpo principal del artículo donde se puede observar que es una función creciente de βC_M .

Resultado final de la información contenida en el tiempo entre pulsos.

Los cálculos mencionados previamente a cuyos resultados queremos llegar son 2.6 y 2.8.

Eqs. (2.6)-(2.8) implican que $I(C, t)$ es independiente de α . El hecho de que Ecs. (2.6)-(2.8) se lean de manera similar independientemente del límite no significa que la información sea la misma en ambos casos: dados T y C , el valor, λ , es diferente dependiendo del límite (ver Ec. (A.2)). El hecho de que I depende

de β pero no de α y que $\beta = B$ en el límite de Poisson y $\beta = 2B$ en el GNF implica que simplemente multiplicando β por 2 en $I(C, t)$, pasamos de un límite al otro por el mismo valor de B . Suponiendo que p_C viene dado por Eq. (2.5) obtenemos lo mencionado en 2.9

Mostramos en la Fig. 2.1 esta información mutua como una función de βC_M para los dos límites del modelo. Allí observamos que siempre es mayor en el límite GNF y que la diferencia aumenta con βC_M acercándose a ~ 1 bit.

A.2.2. Información contenida en $N \gg 1$ pulsos.

Ahora miramos el problema cuando hay muchos (N) pulsos subsiguientes. Para realizar cálculos analíticos consideraremos el límite $N \gg 1$. Básicamente, asumimos que observamos el sistema durante un tiempo fijo, t_{tot} , tal que $t_{tot} \gg (T_{min} + T)$. Luego queremos calcular la información mutua entre N y la concentración del ligando externo, C , asumiendo valores fijos de ρ , α y β (o de A y B) como hemos hecho antes. Para ello, necesitamos conocer la probabilidad condicional de que se produzcan N pulsos durante un tiempo de observación, t_{tot} , para un valor dado de C , *i.e.*, $p(N|C, t_{tot})$. El problema en cuestión es un ejemplo de un *proceso de renovación* [84]. Según el Teorema Central del Límite para este tipo de procesos, para N suficientemente grande, es:

$$p(N|C, t_{tot}) \approx \left(2\pi \frac{t_{tot}}{T_{min} + T} \frac{\sigma^2}{(T_{min} + T)^2} \right)^{-1/2} \exp \left(-\frac{(N - \frac{t_{tot}}{T_{min} + T})^2}{2 \frac{t_{tot}}{T_{min} + T} \frac{\sigma^2}{(T_{min} + T)^2}} \right), \quad (\text{A.48})$$

donde tanto la media, $T_{min} + T$, como la desviación estándar, σ , del tiempo entre pulsos, t , dependen de la concentración del ligando externo, C . Para simplificar los cálculos, de ahora en adelante supondremos que $T_{min} \ll T$. Todos los cálculos se pueden realizar de la misma manera que se describe aquí si no se tiene en cuenta esta suposición. Considerando la simplificación y dada la Ec. (1.4) (que se cumple para los límites de Poisson y GNF), la Ec. (A.48) se puede reescribir como:

$$p(N|C, t_{tot}) = p(N|T, t_{tot}) \approx \left(\frac{2\pi k^2 t_{tot}}{T} \right)^{-1/2} \exp \left(-\frac{(N - \frac{t_{tot}}{T})^2}{2k^2 t_{tot}/T} \right). \quad (\text{A.49})$$

Para valores fijos de ρ , α y β , la media, T , está determinada únicamente por C . En tal caso, la información mutua entre N y C es la misma que la información entre N y T . Luego calcularemos este último tratando N como una variable continua, a saber:

$$I(C, N) = I(N, T) = - \int p(N|t_{tot}) \log_2(p(N|t_{tot})) dN + \iint p(N|T, t_{tot}) p_T(T) \log_2(p(N|T, t_{tot})) dT dN, \quad (\text{A.50})$$

donde $p_T(T)$ es la densidad de probabilidad de la media, T , que se derivará de la concentración del ligando, $p_C(C)$, y

$$p(N|t_{tot}) = \int p(N|T, t_{tot})p_T(T)dT. \quad (\text{A.51})$$

Es conveniente introducir el cambio de variable $\bar{N} = \frac{N}{t_{tot}}$, tal que:

$$I(N, T) = - \int p(\bar{N}|t_{tot}) \log_2 (p(\bar{N}|t_{tot})) d\bar{N} + \iint p(\bar{N}|T, t_{tot})p_T(T) \log_2 (p(\bar{N}|T, t_{tot})) dT d\bar{N}, \quad (\text{A.52a})$$

$$p(\bar{N}|T, t_{tot}) = \sqrt{\frac{t_{tot}T}{2\pi k^2}} \exp\left(-\frac{(\bar{N} - \frac{1}{T})^2}{\frac{2k^2}{t_{tot}T}}\right). \quad (\text{A.52b})$$

Para calcular el primer término en Ec. (A.52a) asumimos que t_{tot} es lo suficientemente grande para que podamos aproximar:

$$\begin{aligned} p(\bar{N}|t_{tot}) &= \int \frac{1}{\sqrt{\frac{2\pi k^2}{t_{tot}T}}} e^{-\frac{(\bar{N} - \frac{1}{T})^2}{\frac{2k^2}{t_{tot}T}}} p_T(T) dT \\ &\approx \int \delta(\bar{N} - \frac{1}{T}) p_T(T) dT + O\left(\frac{1}{t_{tot}}\right). \\ &= \frac{1}{N^2} p_T\left(\frac{1}{N}\right) + O\left(\frac{1}{t_{tot}}\right) \end{aligned} \quad (\text{A.53})$$

La última igualdad se demuestra en A.2.2. De esta forma, el primer término en Ec. (A.52a) se convierte en:

$$\begin{aligned} & - \int p(\bar{N}|t_{tot}) \log_2 (p(\bar{N}|t_{tot})) d\bar{N} \\ & \approx - \int \left(\frac{1}{N^2} p_T\left(\frac{1}{N}\right) + O\left(\frac{1}{t_{tot}}\right) \right) \log_2 \left(\frac{1}{N^2} p_T\left(\frac{1}{N}\right) + O\left(\frac{1}{t_{tot}}\right) \right) d\bar{N} \\ & = - \int p_T(T) (\log_2 (p_T(T)) + 2 \log_2(T)) dT + O\left(\frac{1}{t_{tot}}\right) \\ & = H(T) - 2\langle \log_2(T) \rangle_T, \end{aligned} \quad (\text{A.54})$$

con la entropía, H , como se define en Ec. (2.3) y $\langle \cdot \rangle_T$ la media calculada usando la densidad de probabilidad, p_T . Luego, recordando que la entropía de una distribución normal de desviación estándar σ es $\frac{1}{2} \log_2(2\pi e\sigma^2)$, el segundo término en la Ec. (A.52a) se puede escribir como:

$$\begin{aligned}
& \iint p(\bar{N}|T, t_{tot}) p_T(T) \log_2(p(\bar{N}|T, t_{tot})) dT d\bar{N} = \\
& \int p_T(T) \int \frac{1}{\sqrt{\frac{2\pi k^2}{t_{tot} T}}} e^{-\frac{(\bar{N}-\frac{1}{T})^2}{\frac{2k^2}{t_{tot} T}}} \log_2 \left(\frac{1}{\sqrt{\frac{2\pi k^2}{t_{tot} T}}} e^{-\frac{(\bar{N}-\frac{1}{T})^2}{\frac{2k^2}{t_{tot} T}}} \right) d\bar{N} dT = \\
& -\frac{1}{2} \int p_T(T) \log_2 \left(\frac{2\pi e}{t_{tot} T} k^2 \right) dT = \\
& -\frac{1}{2} \log_2 \left(\frac{2\pi e k^2}{t_{tot}} \right) - \frac{1}{2} \int p_T(T) \log_2 \left(\frac{1}{T} \right) dT.
\end{aligned} \tag{A.55}$$

Asumiendo que C está uniformemente distribuido como antes (Eq. (2.5)) y que T y C están relacionados por: $A \exp(-\gamma C)$ (ver Ec. (2) en el cuerpo principal del artículo) obtenemos:

$$p_T(T) = \begin{cases} \frac{1}{C_M \gamma T}, & \text{si } A \exp(-\gamma C_M) \leq T \leq A, \\ 0, & \text{de otra manera,} \end{cases} \tag{A.56}$$

de modo que $1/T = C_M \gamma p_T(T)$ para los valores de T para los cuales p_T no es cero. Insertando la Ec. (A.56) en la Ec. (A.55) obtenemos:

$$\begin{aligned}
& \iint p(\bar{N}|T, t_{tot}) p_T(T) \log_2(p(\bar{N}|T, t_{tot})) dT d\bar{N} \\
& = \frac{1}{2} \log_2 \left(\frac{t_{tot}}{2\pi e k^2 C_M \gamma} \right) - \frac{1}{2} \int p_T(T) \log_2(p_T) dT \\
& = \frac{1}{2} H(T) + \frac{1}{2} \log_2 \left(\frac{t_{tot}}{2\pi e k^2 C_M \gamma} \right),
\end{aligned} \tag{A.57}$$

con la entropía, H , como se define en Ec. (2.3). Poniendo todo junto, obtenemos lo señalado en 2.10

Por lo tanto, si asumimos que la relación entre T y C es fija y determinada por los experimentos (lo que también implica que los valores A y B se determinan experimentalmente), entonces la densidad de probabilidad, p_T , sólo depende de esta relación y de p_C . Esto significa que la única diferencia entre los dos límites del modelo depende de k que es 1 en el límite de Poisson y $\sqrt{4/\pi - 1} \approx 0,523$ en el GNF. En particular, la diferencia entre la información para el GNF y los límites de Poisson es de 0,936. Hay un término adicional de $\mathcal{O}(\frac{1}{\sqrt[3]{t_{tot}}})$ relacionado con la asimetría residual de la distribución de probabilidad $p(\bar{N}|T, t_{tot})$, que se explora en la sección A.3. Ec. (A.56) implica que la suma de los tres primeros términos de la r.h.s. de Ec. (2.10) es una función, f , de A y BC_M . Teniendo esto en cuenta e incluyendo el término $\mathcal{O}(\frac{1}{\sqrt[3]{t_{tot}}})$ el resultado final se puede escribir como:

$$I(N, T) = f(A, \gamma C_M) + \frac{1}{2} \log_2 \left(\frac{t_{tot}}{k^2} \right) + \mathcal{O} \left(\frac{1}{\sqrt[3]{t_{tot}}} \right). \tag{A.58}$$

Cálculo aproximado de $p(\bar{N}|t_{tot})$

Queremos calcular aproximadamente la función, $p(\bar{N}|T, t_{tot})$, dada por la Ec. (A.52b) para t_{tot} suficientemente grande. Realizando una transformada de Fourier, \mathcal{F} , (con respecto a \bar{N}) sobre ella obtenemos:

$$\begin{aligned} \mathcal{F}\left(\frac{1}{\sqrt{2\pi\frac{1}{t_{tot}T}k^2}}e^{-\frac{(\bar{N}-\frac{1}{T})^2}{\frac{1}{t_{tot}T}k^2}}\right) &= \frac{1}{\sqrt{2\pi}}e^{i[\frac{1}{t_{tot}T}k^2]\omega}e^{-\frac{1}{2t_{tot}T}k^2\omega^2} \sim \\ \frac{1}{\sqrt{4\pi}}e^{i[\frac{1}{T}]\omega}e^{-\frac{1}{2t_{tot}T}k^2\omega^2} &\sim \frac{1}{\sqrt{4\pi}}e^{i(\frac{1}{T})\omega}\left(1 - \frac{k^2\omega^2}{2t_{tot}T} + \mathcal{O}\left(\frac{1}{t_{tot}^2}\right)\right). \end{aligned} \quad (\text{A.59})$$

Antitransformando la última expresión que obtenemos $\delta(\bar{N} - \frac{1}{T}) + \mathcal{O}(\frac{1}{t_{tot}})$.

A.2.3. Resultado final de la información contenida en el número de pulsos, N , que ocurren durante un intervalo.

Los cálculos previantes detallados, para la información mutua entre C y N para un tiempo de observación dado, t_{tot} , $N \gg 1$ y p_C dados por la Ec. (2.5) resultan en:

$$\begin{aligned} I(C, N) &= I(N, T) = \\ \frac{3}{2}H(T) - 2\langle \log_2(T) \rangle_T - \log_2(2\pi e C_M \gamma) + \frac{1}{2} \log_2\left(\frac{t_{tot}}{k^2}\right) + \mathcal{O}(t_{tot}^{-1/3}), \end{aligned} \quad (\text{A.60})$$

con H la entropía como en la Ec. (2.3) y $k = 1$ ($k = \sqrt{4/\pi - 1}$) para el límite de Poisson (GNF). En este caso la única diferencia entre los límites de Poisson y GNF radica en la constante de proporcionalidad, k , entre la desviación estándar, σ y la media, T , de la (parte estocástica del) tiempo entre pulsos. Verificamos el cálculo analítico vía simulaciones numéricas. Allí observamos que, también en este caso, I es mayor para el GNF que para el límite de Poisson y que la diferencia aumenta muy lentamente con t_{tot} acercándose al valor asintótico prescrito por la teoría, $\Delta I = -\log_2(4/\pi - 1)/2 \approx 0,936$. Las curvas de ajuste confirman esto y la dependencia de I de t_{tot} , en particular, el $\mathcal{O}(t^{-1/3})$ término que atribuimos a la asimetría del N densidad de probabilidad que no se incluye al usar la aproximación normal.

A.3. Más allá del Teorema Central del Límite.

En Ec. (A.49) (o, Ecuivalentemente, Ec. (A.52b)) asumimos que la función de densidad de probabilidad del número de pulsos, N , para valores dados de C y t_{tot} podría aproximarse mediante una distribución normal. Este resultado se deriva del *Central Limit Theorem*. Se llamará $\sigma = \sum_{stoch}$. En esta subsección exploramos qué sucede si deformamos ligeramente la distribución normal para

incluir cierta asimetría. Sin pérdida de generalidad introduciremos un cambio de variables tal que $\langle \bar{N} \rangle = 0$, es decir, consideramos que la densidad de probabilidad para la variable reescalada, $\bar{N} = N/t_{tot}$ viene dado por:

$$p(\bar{N}|T, t_{tot}) \sim \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\varphi \frac{\bar{N}}{\bar{\sigma}_N}\right), \quad (\text{A.61})$$

en lugar de Ec. (A.52b), donde $\varphi \rightarrow 0$, $f(0) = 1$, $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\varphi \frac{\bar{N}}{\bar{\sigma}_N}\right) d\bar{N} = 1$.

Un posible $f(\zeta)$ específico podría ser $2\Phi(\zeta)$ con Φ la función de distribución acumulativa de la distribución normal. De esta forma tendríamos la distribución *skew-normal*. Ahora les recordamos a los lectores que estamos trabajando en el límite de grandes t_{tot} para que $\bar{\sigma}_N \sim \mathcal{O}\left(\frac{1}{\sqrt{t_{total}}}\right)$. La función de densidad de probabilidad real para la variable N , $p(N|T, t_{tot})$, en el límite de Poisson del modelo es la distribución de Poisson que se puede aproximar mediante una distribución normal de la misma varianza. La asimetría (*skewness*) de la distribución de Poisson con varianza, $\sum_N^2 \sim t_{tot}$, como en la aproximación normal de Ec. (A.49) satisface:

$$S = \frac{\langle (N - \langle N \rangle)^3 \rangle}{\sum_N^3} = \frac{1}{\sum_N} \sim \frac{1}{\sqrt{t_{tot}}}. \quad (\text{A.62})$$

Dado que $\sum_N^2 \sim t_{tot}$, Ec. (A.62) implica que $\langle (N - \langle N \rangle)^3 \rangle \sim t_{tot}$ tal que:

$$\langle (\bar{N} - \langle \bar{N} \rangle)^3 \rangle \sim \mathcal{O}\left(\frac{1}{t_{tot}^2}\right). \quad (\text{A.63})$$

Ahora aproximamos la función f en Ec. (A.61) por su expansión de Taylor, $f(x) \approx 1 + a_1x + a_2x^2 + a_3x^3$, y calculamos:

$$\begin{aligned} \langle \bar{N} \rangle & \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\frac{\varphi}{\bar{\sigma}_N} \bar{N}\right) \bar{N} d\bar{N} = \\ & \left(\frac{\bar{\sigma}_N}{\varphi}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\varphi^2}} e^{-\frac{u^2}{\varphi^2}} (1 + a_1u + a_2u^2 + a_3u^3) u du \approx \\ & \left(\frac{\bar{\sigma}_N}{\varphi}\right) [a_1\varphi^2 + 3a_3\varphi^4], \end{aligned} \quad (\text{A.64})$$

$$\begin{aligned} \langle \bar{N}^2 \rangle & \sim \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\frac{\varphi}{\bar{\sigma}_N} \bar{N}\right) \bar{N}^2 d\bar{N} = \\ & \left(\frac{\bar{\sigma}_N}{\varphi}\right)^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\varphi^2}} e^{-\frac{u^2}{\varphi^2}} (1 + a_1u + a_2u^2 + a_3u^3) u^2 du \approx \\ & \left(\frac{\bar{\sigma}_N}{\varphi}\right)^2 [\varphi^2 + 3a_2\varphi^4], \end{aligned} \quad (\text{A.65})$$

$$\begin{aligned}
\langle \bar{N}^3 \rangle & \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\frac{\varphi}{\bar{\sigma}_N}\bar{N}\right) \bar{N}^3 d\bar{N} = \\
& \left(\frac{\bar{\sigma}_N}{\varphi}\right)^3 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\varphi^2}} e^{-\frac{u^2}{2\varphi^2}} (1 + a_1u + a_2u^2 + a_3u^3)u^3 du \approx \\
& \left(\frac{\bar{\sigma}_N}{\varphi}\right)^3 [3a_1\varphi^4 + 15a_3\varphi^6].
\end{aligned} \tag{A.66}$$

La última da:

$$\begin{aligned}
\langle (\bar{N} - \langle \bar{N} \rangle)^3 \rangle & = \langle \bar{N}^3 \rangle - 3\langle \bar{N} \rangle \langle \bar{N}^2 \rangle + 2\langle \bar{N} \rangle^3 \sim \\
& \left(\frac{\bar{\sigma}_N}{\varphi}\right)^3 [(3a_1\varphi^4 + 15a_3\varphi^6) - 3(a_1\varphi^2 + 3a_3\varphi^4)(\varphi^2 + 3a_2\varphi^4) + 2(a_1\varphi^2 + 3a_3\varphi^4)^3] \sim \\
& (\bar{\sigma}_N\varphi)^3 [15a_3 - 9a_1a_3 + 9a_1a_2 + 2a_1^3] \sim \\
& \mathcal{O}((\bar{\sigma}_N\varphi)^3).
\end{aligned} \tag{A.67}$$

Considerando que $\bar{\sigma}_N \sim \mathcal{O}(\frac{1}{\sqrt{t_{tot}}})$, Ecs. (A.63) y (A.67) implican que $\varphi \sim \mathcal{O}(\frac{1}{\sqrt[3]{t_{tot}}})$. Así, cuando calculamos el segundo término en la r.h.s. de la Ec. (A.52a) tendríamos que realizar un cálculo de la forma:

$$\begin{aligned}
& \int \frac{1}{\sqrt{2\pi\bar{\sigma}_N^2}} e^{-\frac{\bar{N}^2}{2\bar{\sigma}_N^2}} f\left(\frac{\varphi}{\bar{\sigma}_N}\bar{N}\right) \log_2\left(f\left(\frac{\varphi}{\bar{\sigma}_N}\bar{N}\right)\right) d\bar{N} = \\
& \int \frac{1}{\sqrt{2\pi\varphi^2}} e^{-\frac{u^2}{2\varphi^2}} (1 + a_1u + a_2u^2 + a_3u^3) \log_2(1 + a_1u + a_2u^2 + a_3u^3) du \sim \\
& \frac{1}{\ln(2)} \int \frac{1}{\sqrt{2\pi\varphi^2}} e^{-\frac{u^2}{2\varphi^2}} \left(a_1u + \frac{a_1^2 + 2a_2}{2}u^2\right) du = \\
& \frac{1}{\ln(2)} \left(\frac{a_1^2 + 2a_2}{2}\varphi^2\right) \sim \mathcal{O}(\varphi^2) \sim \mathcal{O}\left(\frac{1}{\sqrt[3]{t_{tot}}}\right).
\end{aligned} \tag{A.68}$$

Entonces concluimos que el error que se comete al calcular la información mutua, $I(C, N)$, usando la aproximación normal de la función de densidad de probabilidad, $p(N|T, t_{tot})$, es $\sim \mathcal{O}(\frac{1}{\sqrt[3]{t_{tot}}})$. Las simulaciones presentadas en el cuerpo principal del documento confirman esta conclusión.

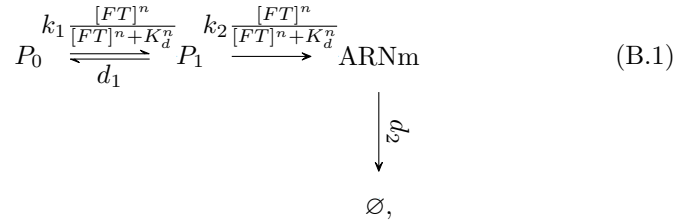
Apéndice B

Apéndice al Capítulo 3

En este Apéndice incluimos cálculos y figuras suplementarias al Capítulo 3. Entre otras cosas, incluimos la descripción mecánica del modelo de transcripción a partir del cual se puede derivar el utilizado en los Caps. 3 y 4.

B.1. Modelo de transcripción: descripción mecánica

Los cálculos de los Capítulos 3 y 4, se han realizado utilizando el modelo de transcripción incluido en la Fig. 3.1, el que corresponde a la versión sencilla del modelo de [46] dado por la Ec. (1.12). Reproducimos acá la parte del modelo que corresponde exclusivamente a la transcripción y su regulación:



donde $[FT]$ es la concentración de factor de transcripción en el núcleo; P_0 y P_1 son las probabilidades, respectivamente, de que el promotor esté en estado inactivo o activo; y el ARNm se degrada a una velocidad d_2 y se produce a una velocidad instantánea $P_1(t)k_2[FT(t)]^n/(K_d^n + [FT(t)]^n)$. Se supone que la proteína producida hasta cierto tiempo, t , es proporcional a la integral de tiempo del ARNm hasta t . Este modelo se presenta en [46] como una versión simplificada y analíticamente resoluble de la más detallada dada por la Ec. (1.12) que resultó ser el que mejor ajustaba las observaciones experimentales. En [46] los autores argumentaron que “hacer que el paso de transcripción dependiera” de $[FT](t)$, era “esencial”. Esta propiedad, que también caracteriza al modelo simplificado (B.1), ‘significa que la transcripción se interrumpe más o menos

inmediatamente después de que el FT “sale del núcleo”, “que la transcripción requiere” que el FT esté ligado al promotor. Los autores también afirmaron que “una forma alternativa de explicar esto” era introducir otro estado del promotor, transcripcionalmente activo, que se desactivara muy rápidamente, pero que no eligieron esa opción para minimizar el número de parámetros libres del modelo y evitar el sobreajuste. Proporcionamos aquí una descripción mecánica que justifica esta dependencia. Lo hacemos para el modelo simplificado, pero se puede seguir un enfoque similar para el modelo de la Ec. (1.12) con un esquema cinético más complicado.

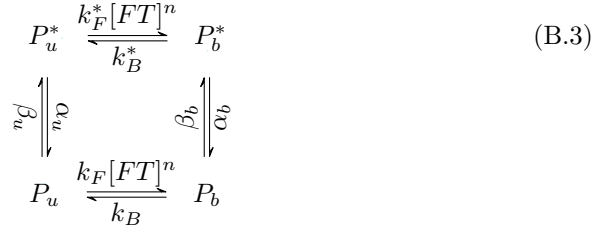
La primera suposición que es evidente en el modelo (B.1) es que el pegado/despegado del FT al promotor ocurre mucho más rápido que su activación/inactivación y mucho más rápido que la tasa de transcripción. En realidad, las moléculas de FT se pegan/despegan una por una, pero si hay una cooperatividad muy fuerte, existen principalmente dos estados “detectables” en términos de la unión: los sitios de unión están todos ocupados o todos vacíos. Esto está representado por la reacción:



donde P_u y P_b representan al promotor con todos sus sitios libres u ocupados, respectivamente. Ahora bien, el promotor (en realidad, la molécula de ADN) también puede estar en diferentes conformaciones. Aquí consideramos que el promotor puede estar en dos que arbitrariamente llamamos “activo” e “inactivo”. Aquí hay dos opciones del modelado: suponer que la transición al estado activo solo ocurre cuando el FT está ligado al promotor o que el promotor puede existir en dos posibles conformaciones y que la unión del FT favorece a una de ellas. Este último es el enfoque introducido por Monod, Wyman y Changeux (MWC) para describir el alosterismo en proteínas [74]. Se han utilizado modelos similares a MWC para describir la transcripción en eucariotas, en particular, para explicar la cooperatividad de la unión de FTs mediada por nucleosomas [72] y para modelar la regulación del gen *hunchback* por el FT, Bicoid, en los embriones de *Drosophila melanogaster* [79], entre otros ejemplos. La justificación del uso de modelos MWC en la regulación transcripcional se deriva del hecho de que el ADN genómico en eucariotas, debido a su interacción con los nucleosomas y las diferentes configuraciones que puede adoptar la cromatina, puede existir en un estado inaccesible para las moléculas FT pero que, una vez que un FT se une por casualidad, se favorece una configuración en la que los sitios de unión de FT se vuelven fácilmente disponibles para otras moléculas [70]. Esto se ilustra en la Fig. 1.3. Así, dentro de este marco podemos pensar en dos estados (“tenso” y “relajado”, en términos de la nomenclatura original del modelo MWC o “activo” e “inactivo”, como hemos estado usando acá) entre los cuales el promotor puede cambiar. Mostramos a continuación cómo se puede derivar el modelo simplificado (B.1) combinando el marco MWC, el esquema cinético aproximado (B.2) y valores de parámetros tales que la transición del estado inactivo al activo ocurre preferentemente cuando las moléculas de FT

están unidas al promotor. Esta última condición está de acuerdo con resultados recientes sobre la transcripción en embriones vivos de *Drosophila* que muestran que la accesibilidad del ADN no es el resultado de fluctuaciones térmicas sino que es catalizada por los FTs [30].

Consideremos un promotor que puede estar en cuatro estados posibles: inactivo con todos sus (n) sitios de ligadura libres, P_u ; activo con todos sus sitios libres, P_u^* ; activo con todos sus sitios unidos a moléculas de FT, P_b^* ; inactivo con todos sus sitios unidos a moléculas de FT, P_b . Por lo tanto, el esquema cinético es:



Este esquema cinético debe combinarse con la ecuación que describe la producción y degradación del ARNm (que, como se explica en el Capítulo 3, modelamos estocásticamente en términos del número de moléculas de ARNm, $X(t) \in \mathbb{N} \cup \{0\}$). Dentro del marco de este modelo mecanístico, consideramos que la producción de ARNm procede con una tasa, $k_T P_b^*$, *i.e.*, solo si el promotor está en el estado P_b^* (“activo” y con FT ligado). Podemos interpretar este condicionamiento como consecuencia de la interacción del ADN con los nucleosomas o de la configuración de la cromatina como se ha descrito anteriormente. Bajo este supuesto, las ecuaciones que rigen el cambio en X son:

$$X = N \xrightarrow{k_T P_b^*(t)} X = N + 1; \quad X = N \xrightarrow{d_2 X} X = N - 1, \tag{B.4}$$

con $N = \{0, 1, 2, \dots\}$. Aunque las transiciones entre los diversos estados del promotor en el esquema (B.3) forman un bucle, debemos considerar que estas transiciones no solo dependen de la interacción del promotor con las moléculas de FT sino también con otros componentes (*p. ej.*, nucleosomas). Por lo tanto, la condición de equilibrio detallada (*i.e.*, que el producto de las tasas alrededor del ciclo en una dirección debe ser igual al producto en la dirección opuesta) no necesariamente se cumple.

Consideremos ahora el esquema cinético (B.3) y supongamos que la unión/desunión del FT ocurre en una escala de tiempo más rápida que la activación/desactivación del promotor y la transcripción. Usando la notación, $P_b(t)$, $P_u(t)$, $P_b^*(t)$ y $P_u^*(t)$, para denotar tanto los estados como la probabilidad de que el promotor esté en cada de ellos al tiempo t , las ecuaciones de evolución en esta escala rápida están dadas aproximadamente por:

$$\begin{aligned}
 \frac{dP_u}{dt} &\approx -k_F[FT]^n P_u + k_B P_b, & \frac{dP_b}{dt} &\approx k_F[FT]^n P_u - k_B P_b, \\
 \frac{dP_u^*}{dt} &\approx -k_F^*[FT]^n P_u^* + k_B^* P_b^*, & \frac{dP_b^*}{dt} &\approx k_F^*[FT]^n P_u^* - k_B^* P_b^*,
 \end{aligned} \tag{B.5}$$

tal que

$$P_0 \equiv P_u + P_b \approx \text{constant}, \quad P_1 \equiv P_u^* + P_b^* \approx \text{constant}, \quad \text{on the fast timescale,} \quad (\text{B.6})$$

$P_0 + P_1 = 1$ y las distintas variables tienden a la (cuasi) solución de equilibrio:

$$P_b(t) \approx \frac{[FT]^n}{K_{d,0}^n + [FT]^n} P_0(t), \quad P_u(t) = P_0(t) - P_b(t) \quad (\text{B.7})$$

$$P_b^*(t) \approx \frac{[FT]^n}{K_{d,1}^n + [FT]^n} P_1(t), \quad P_u^*(t) = P_1(t) - P_b^*(t), \quad (\text{B.8})$$

que en realidad varía en la escala de tiempo lenta. En las Ecs. (B.7)–(B.8), $K_{d,0} \equiv (k_B/k_F)^{1/n}$ y $K_{d,1} \equiv (k_B^*/k_F^*)^{1/n}$ son constantes de disociación efectivas de la reacción de unión/desunión del FT al promotor en sus estados inactivo y activo, respectivamente. Las “constantes” en la Ec. (B.6) evolucionan en una escala de tiempo lenta, manteniendo siempre $P_0 + P_1 = 1$. Cuando se considera la evolución en la escala de tiempo lenta, las Ecs. (B.7)–(B.8) se satisfacen a cada instante (todas las variables ajustan sus valores a un nuevo equilibrio que cambia lentamente, expresado por la dependencia con el tiempo en estas ecuaciones).

Como ya se mencionó, suponemos que la producción de ARNm avanza a un ritmo más lento que la unión/desunión, por lo que puede aproximarse a:

$$\text{tasa de producción de ARNm} = k_T P_b^*(t) \approx k_T \frac{[FT]^n}{K_{d,1}^n + [FT]^n} P_1(t). \quad (\text{B.9})$$

Luego necesitamos determinar cómo varía $P_1(t)$ en la escala de tiempo “lenta” (donde “lenta” significa “más lento” que el pegado/despegado del FT al promotor).

De acuerdo con el esquema, (B.3), la evolución de P_1 está dada por:

$$\frac{dP_1}{dt} = \frac{d}{dt} (P_u^* + P_b^*) = -\alpha_b P_b^* + \beta_b P_b - \alpha_u P_u^* + \beta_u P_u. \quad (\text{B.10})$$

Insertando las Ecs. (B.7)–(B.8) obtenemos la ecuación de evolución de P_1 en la escala de tiempo lenta:

$$\frac{dP_1}{dt} = -\alpha_u P_1 - (\alpha_b - \alpha_u) \frac{[FT]^n}{K_{d,1}^n + [FT]^n} P_1 + \beta_b \frac{[FT]^n}{K_{d,0}^n + [FT]^n} P_0 + \beta_u \frac{K_{d,0}^n}{K_{d,0}^n + [FT]^n} P_0, \quad (\text{B.11})$$

que se puede resolver dado que $P_0 = 1 - P_1$. La Ec. (B.11) describe las transiciones del primer paso del modelo simplificado (B.1) siempre que $|\alpha_b - \alpha_u| \ll \alpha_u$ y $\beta_u \ll \beta_b$. La última desigualdad se mantiene bajo la suposición muy razonable de que la transición del estado inactivo al activo del promotor ocurre con una probabilidad mucho mayor por unidad de tiempo si el promotor está unido a

moléculas de FT que si no lo está (*i.e.*, que la transición $P_b \rightarrow P_b^*$ ocurre con más frecuencia que la $P_u \rightarrow P_u^*$). Como ya se mencionó, esto está de acuerdo con estudios recientes de transcripción en embriones vivos de *Drosophila* que concluyeron que la accesibilidad del ADN es catalizada por los FT [30]. La condición $|\alpha_b - \alpha_u| \ll \alpha_u$ se cumple si suponemos que la relajación del estado activo al inactivo ocurre aproximadamente a la misma velocidad, independientemente de si las moléculas de FT están unidas al promotor o no. Bajo estos supuestos, que podrían justificarse también en la capacidad del modelo (B.1) para reproducir las principales observaciones experimentales con un número mínimo de parámetros [46], la Ec. (B.11) se convierte en:

$$\frac{dP_1}{dt} = -\alpha_u P_1 + \beta_b \frac{[FT]^n}{K_{d,0}^n + [FT]^n} P_0, \quad (\text{B.12})$$

que corresponde a las transiciones del primer paso de (B.1) si identificamos $\alpha_u \approx \alpha_b = d_1$ y $\beta_b = k_1$. Las ecuaciones (B.12) y (B.9) justifican la “doble dependencia” de la tasa de transcripción del modelo simple (B.1) con la $[FT]$ (a través de la transición de P_0 a P_1 y a través de la tasa de transcripción, $k_T \frac{[FT]^n}{K_{d,1}^n + [FT]^n} P_1(t)$). La descripción mecanística que proporcionamos aquí muestra que, para que esta doble dependencia tenga sentido, P_0 y P_1 deben considerarse como la probabilidad de que el promotor esté inactivo o activo, respectivamente, incluidas tanto las situaciones con el FT ligado como sin ligar. La dependencia de la tasa de transición de P_0 a P_1 con $[FT]$ surge porque las transiciones elementales del modelo mecanístico (B.3) ocurren entre subestados de P_0 y P_1 cuyas probabilidades pueden aproximarse como fracciones dependientes de $[FT]$ de P_0 y P_1 . Podemos notar que tanto en (B.1) como en la descripción aproximada del modelo mecanístico, (B.3), dado por las Ecs. (B.9) y (B.12)), el índice de cooperatividad, n , es el mismo para la probabilidad de transición de P_0 a P_1 que para la tasa de transcripción. Sin embargo, la constante de disociación, que también es la misma en el modelo (B.1) no es necesariamente igual según la descripción mecanística. Siempre podemos argumentar que el mejor ajuste a los experimentos de [46] es tal que $K_{d,0} \approx K_{d,1} = K_d$. Sin embargo, también podemos razonar que incluso si k_F y k_F^* o k_B y k_B^* fueran diferentes, no es improbable que la diferencia de energía entre los estados ligados y no ligados sea semajante cuando el promotor está en su conformación activa o cuando no lo está. Esto implica que las constantes de disociación que dependen de esta diferencia satisfacen $K_{d,1} \approx K_{d,0} = K_d$. Bajo este último supuesto, la descripción aproximada del modelo mecanístico (B.3) dada por las Ecs. (B.9) y (B.12) es equivalente a la del modelo simple (B.1).

B.2. Dinámica temporal en la modulación por amplitud y frecuencia

En esta sección se muestran ejemplos donde se varía la duración de un pulso o la frecuencia de pulsos cortos. Se muestra cómo los distintos sets de parámetros

pueden provocar una gran cantidad de información o una más pequeña en una dinámica de modulación por duración o frecuencia, al disminuir el *threshold* o al aumentar el el tiempo característico del sistema. A su vez, se muestra que la MI es no monótona para los cortes temporales en el caso de la modulación por Duración, mientras que crece constantemente en el caso de modulación por Frecuencia.

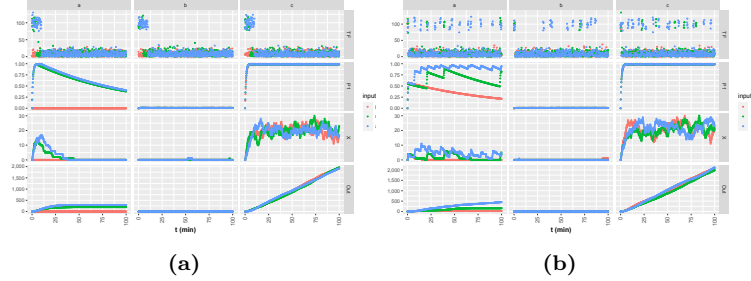


Figura B.1: La dinámica temporal de $[FT]$, P_1 , X y $Out(time)$ para 3 sets de valores de parámetros y 3 elecciones de $FT(t)$ (a-c) y el curso del tiempo para MI para la modulación por amplitud, el tiempo de corte T_c y los parámetros cinéticos del ejemplo en a. (a-c) En todas las simulaciones, un único pulso de FT de $10min$ de duración fue considerado. Las amplitudes fueron de 0 (rojo), 10 (verde) y 100 (azul) con un ruido aditivo superimpuesto de distribución Gaussiana, de desviación estándar 10. Los parámetros usados en las simulaciones de la primer columna son $k_1 = 1/min$, $k_2 = 100/min$, $d_1 = 0,01/min$, $Kd = 70$, $n = 10$ para quienes la información mutua, computada para el input de tipo de modulación amplitud y el corte temporal del output T_c es (B.1a) $MI = 1,47$ (i.e. 97% de MI_M) y (B.1b) $MI = 1,76$ (i.e. 94% de MI_M). Aquellos de la segunda columna son $k_1 = 0,01/min$, $k_2 = 1/min$, $d_1 = 1/min$, $Kd = 10$, $n = 1$ para quienes (B.1a) $MI = 0,03$ y (B.1b) $MI = 0,03$, y aquellos de la tercera: $k_1 = 1/min$, $k_2 = 100/min$, $d_1 = 0,01/min$, $Kd = 10$, $n = 1$ para quienes (B.1a) $MI = 0,02$ y (B.1b) $MI = 0,02$. (B.1a) En la modulación por duración, para el caso a, podemos ver que el FT tiene una duración alta para la entrada alta y una duración baja para la entrada baja. Como consecuencia, la cantidad máxima del promotor se activa durante un tiempo más largo, que es mayor en el caso de que los parámetros estén configurados para transmitir más información, y se logra más rápido. El ARNm alcanza su punto máximo en un valor más alto y en un tiempo posterior. La cantidad de proteína es la integral de la curva de ARNm. Para el caso b, hay baja información porque hay baja señal. Para el caso c, hay baja información porque satura. (B.1b) Para la modulación por frecuencia, para el caso a, a mayor nivel de entrada, la frecuencia es mayor (el periodo medio es menor). El promotor alcanza un valor relativamente estacionario, pero la cantidad de mRNA oscila a un nivel más alto para la entrada más alta. La cantidad de proteína es la integral del ARNm, que alcanza un nivel más alto y más rápido, para una mayor entrada. Para el caso b, hay baja información porque hay baja señal. Para el caso c, hay baja información porque satura.

B.2.1. Simulación del modelo sin límite a la tasa de producción de ARNm

Los resultados que se muestran en el Capítulo 3 fueron obtenidos usando el modelo de la Fig. 3.1 poniendo un límite a la tasa de producción del ARNm de modo que a lo sumo se produce una molécula de ARNm cada 0.2 minutos. Como argumentamos en la Discusión de ese Capítulo, limitar la cantidad de moléculas de ARNm producidas es una forma sencilla de lograr que las fluctuaciones en el reportero de la salida usado en nuestro modelo (el ARNm acumulado) sean más similares a las de las proteínas producidas (la salida en los experimentos) lo que nos permitió obtener valores de MI similares a los experimentales. En esta subsección analizamos el comportamiento del modelo si no se impone este límite pero si restringe la búsqueda de parámetros para maximizar MI a valores tales que $k_2 \leq 10/min$. Lo que encontramos en este sistema es que casi todos los resultados obtenidos coinciden cualitativamente con lo descritos en el Capítulo 3. La principal diferencia radica en los valores que toma la MI que resultan mayores que en el caso anterior debido a la menor incidencia relativa de las fluctuaciones en el número acumulado de moléculas de ARNm producidas.

Como en el Capítulo 3 se muestran los resultados de la ejecución de la optimización para cada combinación de tipo de modulación de entrada y corte de tiempo de salida cuando los parámetros, d_1 , k_1 , k_2 , K_d y n , son variados y $d_2 = 0,12/min$ (ver Sec. 3.1). La Tabla B.1 muestra los conjuntos de parámetros que maximizan MI en cada caso y el rango de valores sobre los cuales se realizó la optimización. Se pueden observar los mayores valores de MI en comparación con los de la Tabla 3.2. A pesar de esto, observamos también en este caso que el óptimo se alcanza cuando $k_2 \gg k_1, d_1, d_2$, lo que implica que la escala de tiempo de la transcripción debe ser tan rápida como sea posible para garantizar una buena transmisión de la información. Análogamente, observamos que si dejamos de lado el corte T_a para la modulación por duración, resulta $n \geq 3$ y la constante de disociación del pegado/despegado del FT resulta $K_d \gtrsim 50$, del mismo orden de magnitud pero menor que la concentración máxima de FT (100).

Presentamos en lo que sigue los mismos análisis realizados para el modelo donde limitamos la tasa de producción de ARNm presentados en el Capítulo TAL. Mostramos en la Fig. B.2 los resultados del análisis de sensibilidad de la MI al variar cada parámetro manteniendo los demás fijos en los valores que dieron MI_M . En todos los casos encontramos comportamientos similares a los descritos en relación a la Fig. 3.5 La principal diferencia entre ambos modelos es que en este caso en que no se limita la tasa de producción del ARNm, MI sigue subiendo para valores altos de k_2 .

Mostramos en la Fig. B.3 un análisis similar al de la Fig. 3.6 donde se muestra la proyección de las seis regiones $Minmax$, para el corte de tiempo T_c , en cada uno de los cinco ejes espaciales de parámetros del modelo de transcripción. Podemos observar que en este caso el parámetro d_1 concentra las diferencias en sensibilidad entre duración y frecuencia, o entre amplitud y frecuencia. En cambio, la diferencia de sensibilidad entre duración y amplitud se encuentra concentrada en los parámetros K_d y n . Es decir, al aumentar el valor del parámetro

Input Modulación	Output Cut	k_1 [1/min] (0.01,1)	k_2 [1/min] (1,10)	n [a.u.] (1,10)	K_d [a.u.] (10,100)	d_1 [1/min] (0.01,1)	MI_M [bits]
Duración	T_a	0.75	9.7	3.4	10	0.011	3.39
Duración	T_b	0.63	9.8	7.6	48.8	0.024	2.62
Duración	T_c	0.86	9.7	8.4	45	0.026	2.58
Amplitud	T_a	0.86	9.7	8.4	45	0.026	1.02
Amplitud	T_b	0.81	9.0	4.1	49	0.03	2.69
Amplitud	T_c	0.73	9.7	4.4	48	0.01	2.52
Frecuencia	T_a	0.72	9.2	9.2	29	0.01	1.71
Frecuencia	T_b	0.72	9.2	9.2	29	0.01	2.21
Frecuencia	T_c	0.73	9.7	4.4	48	0.01	2.54

Tabla B.1: Conjuntos de parámetros que maximizan la información mutua, MI , para cada combinación de modulación de la entrada, el tiempo de integración de la salida, y el valor correspondiente, MI_M , obtenido en cada caso. La tabla muestra el rango en el que cada valor del parámetro fue variado al buscar este máximo. El parámetro, d_2 , siempre se fijó siempre en $d_2 = 0,12/min$. Recordamos que las unidades de concentración arbitrarias son tales que 100 es la concentración de FT nuclear (activa) que se utilizó para calcular MI en el caso de modulación de frecuencia o duración y la concentración máxima en el caso de modulación por amplitud.

d_1 , la información modulada por frecuencia disminuye drásticamente, lo cual no ocurre con la modulación por duración o por amplitud.

Mostramos en la Fig. B.4 los resultados de una exploración similar a la de la Fig. 3.7. Vemos que en este caso también es posible encontrar promotores que sean sensibles a una modulación, pero “ciegos” a otra. No sólo esto, sino que la jerarquía de sensibilidad coincide con la del modelo anterior. En particular, la modulación por duración tiene $MI \sim 2,3bits$ para todos los casos, mientras que la modulación por frecuencia es la más sensible ante el cambio de los parámetros. Los resultados son similares a la figura mencionada. Sin embargo, los resultados son incluso más agudos: en particular, se puede encontrar más claramente que la sensibilidad de la modulación por Amplitud es intermedia entre la modulación por Duración y por Frecuencia.

De acuerdo con la discusión anterior (Fig. B.4a), es la codificación de frecuencia (Inputs F) la que produce mayores diferencias dependiendo del conjunto de parámetros. No solo la cantidad de ARNm producido para las Inputs F es mucho mayor para los dos conjuntos en los que se maximizó la transmisión por modulación de frecuencia (regiones de *Minmax* 20 y 21) que en aquellos en los que se minimizó condicionalmente (02 y 12), sino también los tres cursos de tiempo FT dentro de las Inputs F son más fácilmente distinguibles en el primero que en los dos últimos casos. En el caso de la codificación de duración (Inputs D), las cantidades de ARNm acumulado son ligeramente mayores para el conjunto que pertenece a la región 01 *Minmax* que para el que pertenece a la 10 y ligeramente mayores para este último en comparación con los de las regiones 02 o 20 *Minmax*. La capacidad de distinguir las 3 duraciones no parece variar

mucho con el conjunto de parámetros. En el caso de la codificación de amplitud (Inputs A), la cantidad de ARNm acumulado es mayor para el conjunto que se determinó minimizando condicionalmente la transmisión de información para la modulación de amplitud. Para este tipo de Inputs, la falta de transmisión de información suficiente para el parámetro establecido en la región 01 *Minmax* en comparación con el establecido en la región 10 (ver *vs* cruz marrón marrón en la Fig. B.4 a) parece estar relacionado con la producción de cantidades similares de ARNm independientemente de la amplitud de FT en el primero (consultar las curvas marrón *vs* verde para las Inputs A). Algo similar ocurre cuando se compara el comportamiento obtenido para las Inputs A y los conjuntos en las regiones 12 y 21 *Minmax* (curvas azul y cyan de las Inputs B, respectivamente). La sensibilidad intermedia de la codificación por Amplitud (Inputs A) se puede encontrar en el modelo dado por Fig. B.4, pero no la Fig. 3.7.

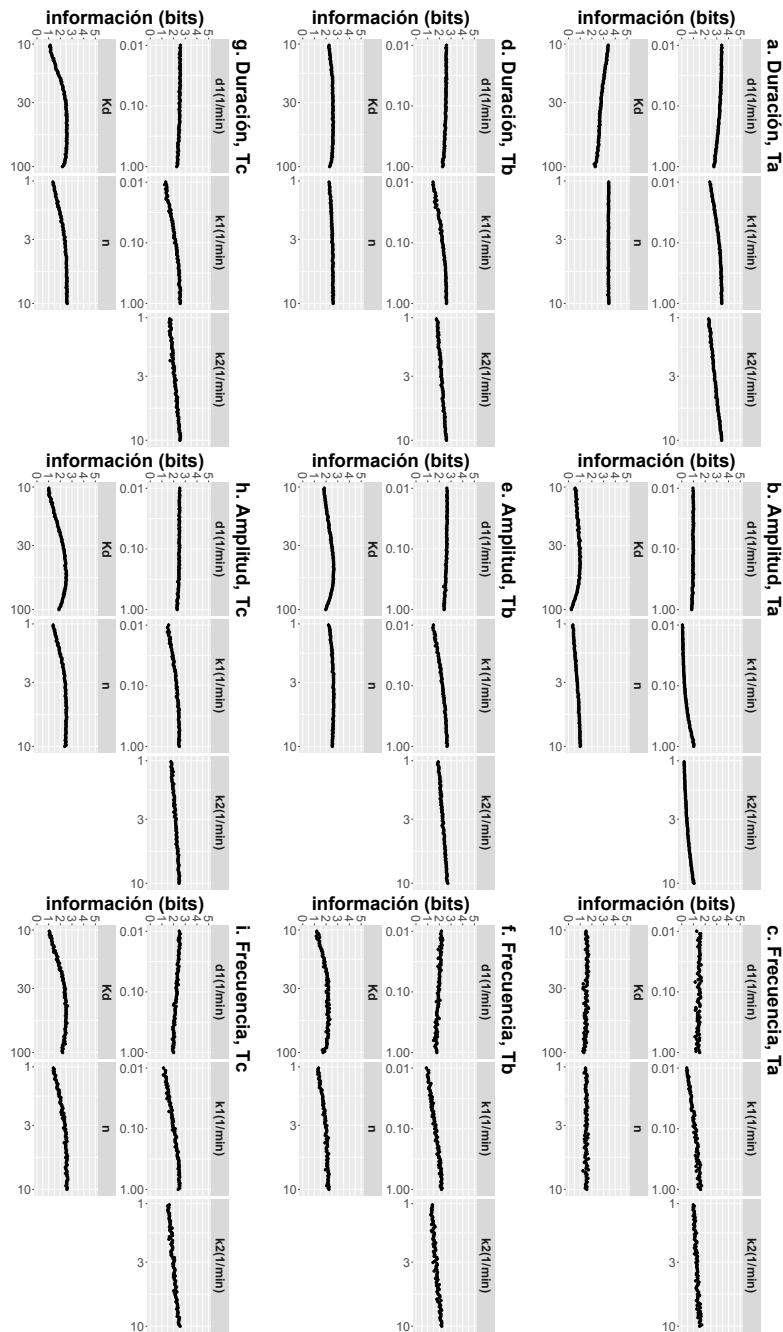


Figura B.2: (Figura rotada). Comportamiento de la información mutua, MI , alrededor de su máximo valor, MI_M , cuando un parámetro cinético es variado. Mostramos MI como función del correspondiente parámetro para cada tipo de modulación de input y corte temporal del output (a-c, d-f y g-i corresponden, respectivamente, al tiempo de corte T_a , T_b y T_c ; a-d-g, b-e-h y c-f-i corresponden, respectivamente, al tipo de modulación de input de duración, amplitud y frecuencia). Excepto por la modulación por duración y el tiempo de corte de T_a , el resto de los comportamientos son muy similares a través de las modulaciones y tiempo de cortes. En particular, la mayor información es obtenida si d_1 es baja, k_1 , k_2 y n son altas y K_d está en el rango medio-alto, pero no muy alto. Para el corte temporal T_a y el tipo de modulación K_d necesita ser baja para maximizar MI .

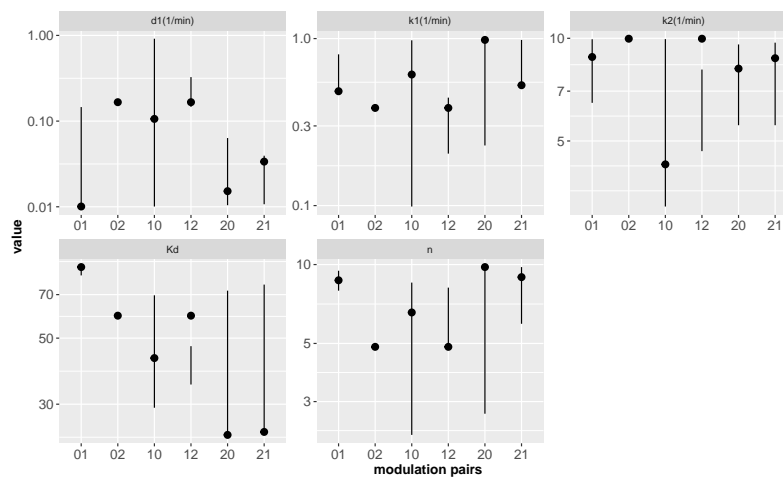


Figura B.3: La proyección en el eje de cada parámetro de las regiones del *Minmax* obtenidas para cada par de tipos de modulación y el corte temporal T_c . Los pares de dígitos en los ejes horizontales identifican los modos para los cuales la transmisión de información fue maximizada (primer dígito) y subsecuentemente minimizada (segundo dígito) para determinar las regiones *Minmax*, con 0: duración, 1: amplitud y 2: frecuencia. Los símbolos corresponden a los valores que minimizan MI para la modulación del segundo dígito que MI permanecer dentro del 90% del MI_M para la modulación del primer dígito.

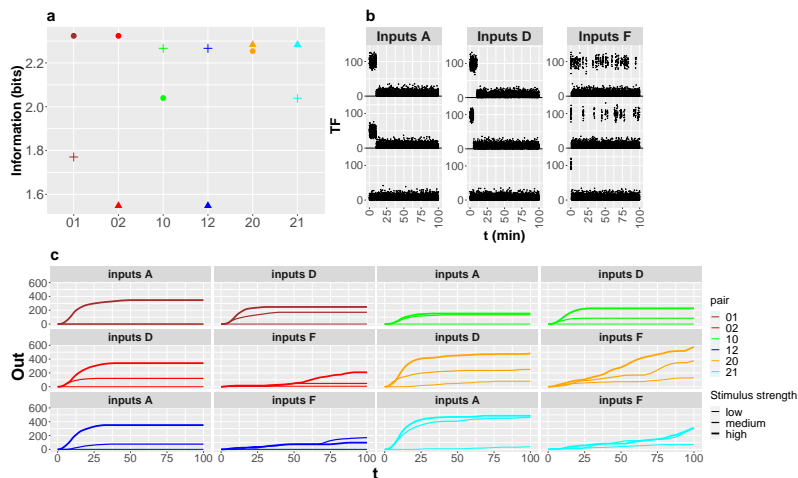


Figura B.4: La información mutua y los comportamientos dinámicos obtenidos usando el conjunto de cada región *Minmax* que minimizan *MI* para el tipo de modulación “segundo”. **a:** *MI* computado usando el set seleccionado en cada región *Minmax* para los modos de modulación de los inputs y condicionalmente minimizados para determinar la región (círculos: duración, cruces: amplitud, triángulos: frecuencia) y el corte temporal T_c . Los labels en el eje horizontal identifican las regiones *Minmax* regions con el orden de dígitos como en la Fig. 3.6. Cada conjunto de parámetros está identificado por un color diferente. Los valores de los parámetros en cada set están mostrados con símbolos en Fig 3.6. **b-c:** mRNA acumulado (*Out*(tiempo)) como función del tiempo (c) obtenidos de simulaciones del modelo usando las trayectorias temporales FT moderadas en **b** (curvas en **c** son mostradas con creciente grosor en relación a la intensidad del estímulo) y el mismo conjunto de parámetros (identificado con el mismo color) que **a**.

Bibliografía

- [1] *Albeck John G, Mills Gordon B, Brugge Joan S.* Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals // *Molecular Cell.* 2013. 49. 249–261.
- [2] *Alberts Bruce, Bray Dennis, Hopkin Karen, Johnson Alexander, Lewis Julian, Raff Martin, Roberts Keith, Walter Peter.* *Essential Cell Biology.* New York: Garland Science, 2014.
- [3] *Ashall Louise, Horton Caroline A, Nelson David E, Paszek Pawel, Harper Claire V, Sillitoe Kate, Ryan Sheila, Spiller David G, Unitt John F, Broomhead David S, Kell Douglas B, Rand David A, Sée Violaine, White Michael R H.* Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription // *Science (New York, N.Y.).* IV 2009. 324, 5924. 242–246.
- [4] *Bardwell Lee.* A walk-through of the yeast mating pheromone response pathway // *Peptides.* 2005. 26, 2. 339–350.
- [5] *Batchelor Eric, Loewer Alexander, Mock Caroline, Lahav Galit.* Stimulus-dependent dynamics of p53 in single cells // *Mol Syst Biol.* V 2011. 7. 488.
- [6] *Batchelor Eric, Mock Caroline S., Bhan Irun, Loewer Alexander, Lahav Galit.* Recurrent Initiation: A Mechanism for Triggering p53 Pulses in Response to DNA Damage // *Molecular Cell.* 2008. 30, 3. 277–289.
- [7] *Batchelor Eric, Mock Caroline S, Bhan Irun, Loewer Alexander, Lahav Galit.* Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage // *Mol Cell.* V 2008. 30, 3. 277–289.
- [8] *Beaudry Normand J., Renner Renato.* An Intuitive Proof of the Data Processing Inequality // *Quantum Info. Comput.* may 2012. 12, 5–6. 432–441.
- [9] *Berg H C, Purcell E M.* Physics of chemoreception // *Biophys J.* XI 1977. 20, 2. 193–219.

-
- [10] *Berridge Michael, Bootman Martin, Lipp Peter*. Calcium - a life and death signal // *Nature*. 1998. 395. 645–648.
- [11] *Berridge Michael J, Lipp Peter, Bootman Martin D*. The versatility and universality of calcium signalling // *Nature reviews Molecular cell biology*. 2000. 1, 1. 11–21.
- [12] *Bialek W, Rieke F, Steveninck R R de Ruyter van, Warland D*. Reading a neural code // *Science*. VI 1991. 252, 5014. 1854–1857.
- [13] *Bialek William*. *Biophysics. Searching for Principles*. Princeton & Oxford: Princeton University Press, 2012.
- [14] *Bialek William*. Perspectives on theory at the interface of physics and biology // *Reports on Progress in Physics*. dec 2017. 81, 1. 012601.
- [15] *Bruno Luciana, Solovey Guillermo, Ventura Alejandra C., Dargan Sheila, Dawson Silvina Ponce*. Quantifying calcium fluxes underlying calcium puffs in *Xenopus laevis* oocytes // *Cell Calcium*. 2010. 47, 3. 273–286.
- [16] *Bush Alan, Vasen Gustavo, Constantinou Andreas, Dunayevich Paula, Patop Inés Lucía, Blaustein Matías, Colman-Lerner Alejandro*. Yeast GPCR signaling reflects the fraction of occupied receptors, not the number // *Molecular Systems Biology*. 2016. 12, 12. 898.
- [17] *Cai Long, Dalal Chiraj K., Elowitz Michael B*. Frequency-modulated nuclear localization bursts coordinate gene regulation // *Nature*. 2008. 455. 485–490.
- [18] *Callamaras Nick, Parker Ian*. Phasic characteristic of elementary Ca²⁺ release sites underlies quantal responses to IP₃ // *The EMBO Journal*. 2000. 19, 14. 3608–3617.
- [19] *Carbó Natalia, Tarkowski Nahuel, Ipiña Emiliano Perez, Dawson Silvina Ponce, Aguilar Pablo S*. Sexual pheromone modulates the frequency of cytosolic Ca²⁺ bursts in *Saccharomyces cerevisiae* // *Molecular Biology of the Cell*. 2017. 28, 4. 501–510.
- [20] *Cheong Raymond, Rhee Alex, Wang Chiao-chun Joanne, Nemenman Ilya, Levchenko Andre*. Information Transduction Capacity of Noisy Biochemical Signaling Networks // *Science*. 2011. 334, 6054. 354–358.
- [21] *Colman-Lerner Alejandro, Gordon Andrew, Serra Eduard, Chin Tina, Resnekov Orna, Endy Drew, Gustavo Pesce C., Brent Roger*. Regulated cell-to-cell variation in a cell-fate decision system // *Nature*. 2005. 437. 699–706.
- [22] *Dalal Chiraj K., Cai Long, Lin Yihan, Rahbar Kasra, Elowitz Michael B*. Pulsatile Dynamics in the Yeast Proteome // *Current Biology*. 2014. 24, 18. 2189–2194.

- [23] *De Koninck Paul, Schulman Howard.* Sensitivity of CaM Kinase II to the Frequency of Ca²⁺ Oscillations // *Science*. 1998. 279, 5348. 227–230.
- [24] *De Young G W, Keizer J.* A single-pool inositol 1,4,5-trisphosphate-receptor-based model for agonist-stimulated oscillations in Ca²⁺ concentration. // *Proceedings of the National Academy of Sciences*. 1992. 89, 20. 9895–9899.
- [25] *Dolmetsch Ricardo E., Xu Keli, Lewis Richard S.* Calcium oscillations increase the efficiency and specificity of gene expression // *Nature*. 1998. 392. 933–936.
- [26] *Dubuis Julien O., Tkacik Gasper, Wieschaus Eric F., Gregor Thomas, Bialek William.* Positional information, in bits // *Proceedings of the National Academy of Sciences*. 2013. 110, 41. 16301–16308.
- [27] *Dupont G, Goldbeter A.* One-pool model for Ca²⁺ oscillations involving Ca²⁺ and inositol 1,4,5-trisphosphate as co-agonists for Ca²⁺ release // *Cell Calcium*. IV 1993. 14, 4. 311–322.
- [28] *Dupont Genevieve, Combettes Laurent, Bird Gary S., Putney James W.* Calcium Oscillations // *Cold Spring Harbor Perspectives in Biology*. 2011. 3, 3.
- [29] *Dupont Geneviève, Goldbeter Albert.* CaM kinase II as frequency decoder of Ca²⁺ oscillations // *BioEssays*. 1998. 20, 8. 607–610.
- [30] *Eck Elizabeth, Liu Jonathan, Kazemzadeh-Atoufi Maryam, Ghoreishi Sydney, Blythe Shelby A, Garcia Hernan G.* Quantitative dissection of transcription in development yields evidence for transcription-factor-driven chromatin accessibility // *eLife*. oct 2020. 9. e56429.
- [31] *Eguia M C, Mindlin G B.* Distribution of interspike times in noise-driven excitable systems // *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*. VI 2000. 61, 6 Pt A. 6490–6499.
- [32] *Eldar Avigdor, Elowitz Michael B.* Functional roles for noise in genetic circuits // *Nature*. 09 2019. 467, 7312. 167–173.
- [33] *Elowitz Michael B., Levine Arnold J., Siggia Eric D., Swain Peter S.* Stochastic Gene Expression in a Single Cell // *Science*. 2002. 297, 5584. 1183–1186.
- [34] *Fabiato A.* Calcium-induced release of calcium from the cardiac sarcoplasmic reticulum // *Am J Physiol*. 1983. 245, 1. 1–15.
- [35] *Ferrell James E Jr, Ha Sang Hoon.* Ultrasensitivity part II: multisite phosphorylation, stoichiometric inhibitors, and positive feedback // *Trends Biochem. Sci*. XI 2014. 39, 11. 556–569.

- [36] *Ferrell James E Jr, Ha Sang Hoon*. Ultrasensitivity part III: cascades, bistable switches, and oscillators // Trends Biochem. Sci. XII 2014. 39, 12. 612–618.
- [37] *Fitzhugh R*. Impulses and Physiological States in Theoretical Models of Nerve Membrane // Biophys J. VII 1961. 1, 6. 445–466.
- [38] *Giorgetti Luca, Siggers Trevor, Tiana Guido, Caprara Greta, Notarbartolo Samuele, Corona Teresa, Pasparakis Manolis, Milani Paolo, Bulyk Martha L, Natoli Gioacchino*. Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs // Mol Cell. II 2010. 37, 3. 418–428.
- [39] *Givré Alan, Colman-Lerner Alejandro, Ponce Dawson Silvina*. Modulation of transcription factor dynamics allows versatile information transmission // Scientific Reports. 2023. 13, 1. 2652.
- [40] *Givré Alan, Colman-Lerner Alejandro, Dawson Silvina Ponce*. Amplitude and Frequency encodings give cells a different lens to sense the environment. 2024.
- [41] *Givré Alan, Dawson Silvina Ponce*. Cell information processing via frequency encoding and excitability. 2023.
- [42] *Givré Alan, Ponce Dawson Silvina*. Information Content in Stochastic Pulse Sequences of Intracellular Messengers // Frontiers in Physics. 2018. 6.
- [43] *Gregor Thomas, Tank David W, Wieschaus Eric F, Bialek William*. Probing the limits to positional information // Cell. VII 2007. 130, 1. 153–164.
- [44] *Gutkin B S, Ermentrout G B*. Dynamics of membrane excitability determine interspike interval variability: a link between spike generation mechanisms and cortical spike train statistics // Neural Comput. VII 1998. 10, 5. 1047–1065.
- [45] *Hänggi Peter, Talkner Peter, Borkovec Michal*. Reaction-rate theory: fifty years after Kramers // Rev. Mod. Phys. Apr 1990. 62. 251–341.
- [46] *Hansen Anders S, O’Shea Erin K*. Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression // Mol Syst Biol. XI 2013. 9. 704.
- [47] *Hansen Anders S, O’Shea Erin K*. Limits on information transduction through amplitude and frequency regulation of transcription factor activity // eLife. may 2015. 4. e06559.
- [48] *Hao Nan, O’Shea Erin K*. Signal- dependent dynamics of transcription factor translocation controls gene expression // Nature Structural & Molecular Biology. I 2012. 19, 1. 31–39.

- [49] *Hao Nan, Zeng Yaxue, Elston Timothy C, Dohlman Henrik G.* Control of MAPK specificity by feedback phosphorylation of shared adaptor protein Ste50 // *The Journal of biological chemistry.* XII 2008. 283, 49. 33798–33802.
- [50] *Hill A. V., Hill A., Paganini-Hill A.* The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves // *The Journal of Physiology.* 1909. 40. 4–7.
- [51] *Hodgkin A L, Huxley A F.* A quantitative description of membrane current and its application to conduction and excitation in nerve // *J Physiol.* VIII 1952. 117, 4. 500–544.
- [52] *Hughey Jacob J., Gutschow Miriam V., Bajar Bryce T., Covert Markus W.* Single-cell variation leads to population invariance in NF- κ B signaling dynamics // *Molecular Biology of the Cell.* 2015. 26, 3. 583–590. PMID: 25473117.
- [53] *IZHIKEVICH EUGENE M.* NEURAL EXCITABILITY, SPIKING AND BURSTING // *International Journal of Bifurcation and Chaos.* 2000. 10, 06. 1171–1266.
- [54] *Iman. R. L.* Appendix A : Latin Hypercube Sampling 1, *Encyclopedia of Statistical Sciences.* London: Kotz, 2001.
- [55] *Jacquet Michel, Renault Georges, Lallet Sylvie, De Mey Jan, Goldbeter Albert.* Oscillatory nucleocytoplasmic shuttling of the general stress response transcriptional activators Msn2 and Msn4 in *Saccharomyces cerevisiae* // *The Journal of cell biology.* May 2003. 161, 3. 497–505. 12732613[pmid].
- [56] *Jones R.* *Soft Condensed Matter.* New York: Oxford U. Press, 2002.
- [57] *Keener J, Sneyd J.* *Mathematical Physiology I: Cellular Physiology.* Berlin: Springer, 2009.
- [58] *Kribelbauer Judith F, Rastogi Chaitanya, Bussemaker Harmen J, Mann Richard S.* Low-affinity binding sites and the transcription factor specificity paradox in eukaryotes // *Annu. Rev. Cell Dev. Biol.* X 2019. 35, 1. 357–379.
- [59] *Laughlin Simon.* A Simple Coding Procedure Enhances a Neuron's Information Capacity // *Zeitschrift für Naturforschung C.* 1981. 36, 9-10. 910–912.
- [60] *Lee DeVille R. E., Vanden-Eijnden Eric, Muratov Cyrill B.* Two distinct mechanisms of coherence in randomly perturbed dynamical systems // *Phys. Rev. E.* Sep 2005. 72. 031105.
- [61] *Levine Joe H, Lin Yihan, Elowitz Michael B.* Functional roles of pulsing in genetic circuits // *Science.* XII 2013. 342, 6163. 1193–1200.

- [62] *Li Yue-Xian, Rinzel John.* Equations for InsP3 Receptor-mediated [Ca²⁺]_i Oscillations Derived from a Detailed Kinetic Model: A Hodgkin-Huxley Like Formalism // *Journal of Theoretical Biology.* 1994. 166, 4. 461–473.
- [63] *Lindner Benjamin, Longtin André, Bulsara Adi.* Analytic expressions for rate and CV of a type I neuron driven by white gaussian noise // *Neural Comput.* VIII 2003. 15, 8. 1760–1787.
- [64] *Loewer Alexander, Batchelor Eric, Gaglia Giorgio, Lahav Galit.* Basal dynamics of p53 reveal transcriptionally attenuated pulses in cycling cells // *Cell.* VII 2010. 142, 1. 89–100.
- [65] *Lopez Lucía, Piegari Estefanía, Sigaut Lorena, Ponce Dawson Silvina.* Intracellular calcium signals display an avalanche-like behavior over multiple length scales. // *Frontiers in Physiology.* 2012. 3, 350.
- [66] *Lopez Lucia F, Dawson Silvina Ponce.* Luminal Ca²⁺ dynamics during IP3R mediated signals // *Physical Biology.* may 2016. 13, 3. 036006.
- [67] *Luo A.* Bifurcation and Stability in Nonlinear Dynamical Systems. 2019.
- [68] *Luzzi Veronica, Sims Christopher E., Soughayer Joseph S., Allbritton Nancy L.* The Physiologic Concentration of Inositol 1,4,5-Trisphosphate in the Oocytes of *Xenopus laevis* * // *Journal of Biological Chemistry.* 1998. 273, 44. 28657–28662.
- [69] *Marchant Jonathan S, Parker Ian.* Role of elementary Ca²⁺ puffs in generating repetitive Ca²⁺ oscillations // *EMBO Journal.* 2001. 20, 1. 65–76.
- [70] *Marzen Sarah, Garcia Hernan G., Phillips Rob.* Statistical Mechanics of Monod–Wyman–Changeux (MWC) Models // *Journal of Molecular Biology.* 2013. 425, 9. 1433–1460. Allosteric Interactions and Biological Regulation (Part I).
- [71] *Minas Giorgos, Woodcock Dan J., Ashall Louise, Harper Claire V., White Michael R. H., Rand David A.* Multiplexing information flow through dynamic signalling systems // *PLOS Computational Biology.* 08 2020. 16, 8. 1–18.
- [72] *Mirny Leonid A.* Nucleosome-mediated cooperativity between transcription factors // *Proceedings of the National Academy of Sciences.* 2010. 107, 52. 22534–22539.
- [73] *Mönke Gregor, Cristiano Elena, Finzel Ana, Friedrich Dhana, Herzl Hanspeter, Falcke Martin, Loewer Alexander.* Excitability in the p53 network mediates robust signaling with tunable activation thresholds in single cells // *Scientific Reports.* 2017. 7, 1. 46571.

- [74] *Monod Jacques, Wyman Jeffries, Changeux Jean-Pierre*. On the nature of allosteric transitions: A plausible model // *Journal of Molecular Biology*. 1965. 12, 1. 88–118.
- [75] *Mora Thierry, Bialek William*. Are Biological Systems Poised at Criticality? // *Journal of Statistical Physics*. Jul 2011. 144, 2. 268–302.
- [76] *Muller Nicolas, Piel Matthieu, Calvez Vincent, Voituriez Raphaël, Gonçalves-Sá Joana, Guo Chin-Lin, Jiang Xingyu, Murray Andrew, Meunier Nicolas*. A Predictive Model for Yeast Cell Polarization in Pheromone Gradients // *PLOS Computational Biology*. 04 2016. 12, 4. 1–21.
- [77] *Muratov Cyrill B., Vanden-Eijnden Eric, E. Weinan*. Self-induced stochastic resonance in excitable systems // *Physica D: Nonlinear Phenomena*. 2005. 210, 3. 227–240.
- [78] *Petrenko Natalia, Chereji Razvan V, McClean Megan N, Morozov Alexandre V, Broach James R*. Noise and interlocking signaling pathways promote distinct transcription factor dynamics in response to different stresses // *Mol Biol Cell*. IV 2013. 24, 12. 2045–2057.
- [79] *Phillips Rob, Kondev Jane, Theriot Julie, Garcia Hernan*. *Physical Biology of the Cell*, 2nd ed. New York: Garland Science, Taylor & Francis Group, November 2012.
- [80] *Piegari Estefania, López Lucía F., Ponce Dawson Silvina*. Using two dyes to observe the competition of Ca^{2+} trapping mechanisms and their effect on intracellular Ca^{2+} signals. to be published. 2018.
- [81] *Piegari Estefania, Lopez Lucia, Perez Ipiña Emiliano, Ponce Dawson Silvina*. Fluorescence Fluctuations and Equivalence Classes of Ca^{2+} Imaging Experiments // *PLOS ONE*. 04 2014. 9, 4. 1–18.
- [82] *Piegari Estefanía, Sigaut Lorena, Ponce Dawson Silvina*. Ca^{2+} images obtained in different experimental conditions shed light on the spatial distribution of IP3 receptors that underlie Ca^{2+} puffs // *Cell Calcium*. 2015. 57, 2. 109–119.
- [83] *Purvis Jeremy E, Karhohs Kyle W, Mock Caroline, Batchelor Eric, Loefer Alexander, Lahav Galit*. p53 dynamics control cell fate // *Science*. VI 2012. 336, 6087. 1440–1444.
- [84] *Ross Sheldon M*. *Introduction to Probability Models*, 11th ed. Waltham, MA: Academic Press, 2014.
- [85] *Schuster Stefan, Marhl Marko, Höfer Thomas*. Modelling of simple and complex calcium oscillations // *European Journal of Biochemistry*. 2002. 269, 5. 1333–1355.

- [86] *Shannon C. E.* A mathematical theory of communication // The Bell System Technical Journal. 1948. 27, 3. 379–423.
- [87] *Shaywitz A J, Greenberg M E.* CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals // Annu Rev Biochem. 1999. 68. 821–861.
- [88] *Skalhegg B S, Tasken K.* Specificity in the cAMP/PKA signaling pathway. Differential expression, regulation, and subcellular localization of subunits of PKA // Front Biosci. VIII 2000. 5. D678–93.
- [89] *Skataric M.* Fold change detection in 3-node enzymatic networks // New Brunswick Electronic Theses and Dissertations. 2011.
- [90] *Skupin Alexander, Falcke Martin.* Statistical properties and information content of calcium oscillations // Genome Inform Ser. 2007. 18. 44–53.
- [91] *Skupin Alexander, Kettenmann Helmut, Falcke Martin.* Calcium signals driven by single channel noise // PLoS Comput Biol. VIII 2010. 6, 8.
- [92] *Skupin Alexander, Kettenmann Helmut, Winkler Ulrike, Wartenberg Maria, Sauer Heinrich, Tovey Stephen C., Taylor Colin W., Falcke Martin.* How Does Intracellular Ca²⁺ Oscillate: By Chance or by the Clock? // Biophys. J. 2008. 94. 2404–2411.
- [93] *Slater Lucy Joan.* Generalized Hypergeometric Functions. Cambridge: Cambridge University Press, 1966.
- [94] *Smith Ian F., Parker Ian.* Imaging the quantal substructure of single IP₃R channel activity during Ca²⁺ puffs in intact mammalian cells // Proceedings of the National Academy of Sciences. 2009. 106, 15. 6404–6409.
- [95] *Softky W R, Koch C.* The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs // J Neurosci. I 1993. 13, 1. 334–350.
- [96] *Solovey Guillermo, Dawson Silvina Ponce.* Intra-Cluster Percolation of Calcium Signals // PLoS ONE. 02 2010. 5, 2. e8997.
- [97] *Strong S., Koberle Roland, Steveninck Rob, Bialek William.* Entropy and Information in Neural Spike Trains // Physical Review Letters. 04 1996. 80.
- [98] *Süel Gürol M., Garcia-Ojalvo Jordi, Liberman Louisa M., Elowitz Michael B.* An excitable gene regulatory circuit induces transient cellular differentiation // Nature. Mar 2006. 440, 7083. 545–550.
- [99] *Sun Xiao-Ping, Nicholas Callamaras, Marchant Jonathan S, Parker Ian.* A continuum of InsP₃-mediated elementary Ca²⁺ signalling events in *Xenopus* oocyte // The Journal of Physiology. 1998. 509, 1. 67–80.

- [100] *Tang Y, Othmer H G.* Frequency encoding in excitable systems with applications to calcium oscillations // Proceedings of the National Academy of Sciences. 1995. 92, 17. 7869–7873.
- [101] *Tay Savaş, Hughey Jacob J., Lee Timothy K., Lipniacki Tomasz, Quake Stephen R., Covert Markus W.* Single-cell NF- κ B dynamics reveal digital activation and analogue information processing // Nature. Jul 2010. 466, 7303. 267–271.
- [102] *Thurley Kevin, Falcke Martin.* Derivation of Ca²⁺ signals from puff properties reveals that pathway function is robust against cell variability but sensitive for control // Proceedings of the National Academy of Sciences. 2011. 108, 1. 427–432.
- [103] *Thurley Kevin, Tovey Stephen C., Moenke Gregor, Prince Victoria L., Meena Abha, Thomas Andrew P., Skupin Alexander, Taylor Colin W., Falcke Martin.* Reliable Encoding of Stimulus Intensities Within Random Sequences of Intracellular Ca²⁺ Spikes // Science Signaling. 2014. 7, 331. ra59–ra59.
- [104] *Tkacik Gasper, Callan Curtis G., Bialek William.* Information flow and optimization in transcriptional regulation // Proceedings of the National Academy of Sciences. 2008. 105, 34. 12265–12270.
- [105] *Tkacik Gasper, Gregor Thomas, Bialek William.* The Role of Input Noise in Transcriptional Regulation // PLoS ONE. 2008. 3.
- [106] *Tkacik Gasper, Walczak Aleksandra M, Bialek William.* Optimizing information flow in small genetic networks. III. A self-interacting gene // Phys Rev E Stat Nonlin Soft Matter Phys. IV 2012. 85, 4 Pt 1. 041903.
- [107] *Wang S S, Alousi A A, Thompson S H.* The lifetime of inositol 1,4,5-trisphosphate in single cells. // Journal of General Physiology. 01 1995. 105, 1. 149–171.
- [108] *Wu Mengqiu, Ye Hui, Tang Zhiyuan, Shao Chang, Lu Gaoyuan, Chen Baoqiang, Yang Yuyu, Wang Guangji, Hao Haiping.* P53 dynamics orchestrates with binding affinity to target genes for cell fate decision // Cell Death Dis. X 2017. 8, 10. e3130–e3130.
- [109] *Yao Y, Choi J, Parker I.* Quantal puffs of intracellular Ca²⁺ evoked by inositol trisphosphate in Xenopus oocytes // J Physiol (Lond). 1995. 482, 3. 533–553.
- [110] *Yi Tau-Mu, Kitano Hiroaki, Simon Melvin I.* A quantitative characterization of the yeast heterotrimeric G protein cycle // Proceedings of the National Academy of Sciences. 2003. 100, 19. 10764–10769.

-
- [111] *Yissachar Nissan, Sharar Fischler Tali, Cohen Ariel A, Reich-Zeliger Shlomit, Russ Dor, Shifrut Eric, Porat Ziv, Friedman Nir.* Dynamic response diversity of NFAT isoforms in individual living cells // *Mol Cell*. XII 2012. 49, 2. 322–330.
- [112] *Yu Richard C, Pesce C Gustavo, Colman-Lerner Alejandro, Lok Larry, Pincus David, Serra Eduard, Holl Mark, Benjamin , AU-Gordon Kirsten, Andrew , Brent Roger.* Negative feedback that improves information transmission in yeast signalling // *Nature*. XII 2008. 456, 7223. 755–761.