



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Efectos de las fluctuaciones intrínsecas en redes de procesos
bioquímicos.

Tesis presentada al

Posgrado en Física Aplicada

como requisito parcial para la obtención del grado de

MAESTRO EN CIENCIAS

por

Lic. Manuel Eduardo Hernández García

Asesorado por

Dr. Jorge Velázquez Castro, Dr. Javier Miguel Hernández López

Puebla Pue.
Enero 2023



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Efectos de las fluctuaciones intrínsecas en redes de procesos
bioquímicos.

Tesis presentada al

Posgrado en Física Aplicada

como requisito parcial para la obtención del grado de

MAESTRO EN CIENCIAS

por

Lic. Manuel Eduardo Hernández García

Asesorado por

Dr. Jorge Velázquez Castro, Dr. Javier Miguel Hernández López

Puebla Pue.
Enero 2023

Título: Efectos de las fluctuaciones intrínsecas en redes de procesos bioquímicos.

Estudiante: LIC. MANUEL EDUARDO HERNÁNDEZ GARCÍA

COMITÉ

Dr. Eduardo Moreno Barbosa
Presidente

Dr. Hugo Adán Cruz Suárez
Secretario

Dra. Beatriz Bonilla Capilla
Vocal

Dra. Mariana Gómez Schiavon
Vocal

Dr. Jorge Velázquez Castro, Dr. Javier Miguel Hernández López
Asesor

Agradecimientos

La culminación de una nueva etapa extraordinaria de mi vida no habría sido posible sin el apoyo incondicional de mi familia, por ello quiero agradecer infinitamente y eternamente a mi padre Cristóbal Hernández García y a mi madre Magdalena García Martínez, a mi hermano Oscar Ediel Hernández García, por confiar en mí cuando decidí estudiar física y perseguir mis sueños; gracias por tantos años de orientación y lecciones de vida, gracias a todo eso me ha llevado a explorar un poco más del mundo, que no les quepa duda que estaré siempre para ustedes así como ustedes lo han estado para mí.

A todos mis amigos de la facultad, Máx, Ime, Joshua, Jorge, Orlando, Laura, Zaida, André, Cristhian, Alexis, Yetla, César, Irandheny y demás compañeros, con quienes he compartido parte de mi aprendizaje, me han motivado, y quienes siempre llevaré en la memoria.

Agradezco al Dr. Jorge Velázquez Castro por su paciencia, apoyo, dirección y mostrarme una de las áreas del conocimiento que más me han fascinado.

Agradezco al Dr. Javier Miguel Hernández López por sus valiosos comentarios.

Finalmente, agradezco al CONACYT por la beca proporcionada durante el transcurso de la maestría.

*“¿Qué es la vida? Un frenesí.
¿Qué es la vida? Una ilusión,
una sombra, una ficción,
y el mayor bien es pequeño;
que toda la vida es sueño,
y los sueños, sueños son”.*
Pedro Calderón de la Barca,
La vida es sueño (1635).

Dedico esta tesis a todo aquel que desea ser un explorador.

Índice general

Resumen	XIX
Introducción	XXI
1. Teoría de control	1
1.1. ¿Qué es retroalimentación?	1
1.2. ¿Qué es control?	2
1.3. Usos de retroalimentación	3
1.4. Análisis de sistemas con retroalimentación y control	4
1.4.1. Marco matemático	4
1.4.2. Ejemplos	6
1.5. Retroalimentación del estado	8
1.5.1. Linealizar sistemas con retroalimentación	8
1.5.2. Estabilización del estado por retroalimentación	10
1.6. Funciones de transferencia	11
1.6.1. Derivación de la función de transferencia	12
1.6.2. Diagramas de bloque y funciones de transferencia	13
1.7. Control PID	14
1.7.1. Funciones básicas de control	14
2. Biología sintética	17
2.1. ¿Qué es la biología sintética?	17
2.2. Módulo de transcripción/traducción	18
2.3. Dinámica de MTT con regulación	21
2.4. Diseño de circuitos genéticos	22
2.4.1. El Represilador	23
2.4.2. El reloj activador-represor	24
2.5. Mas allá de la asunción de modularidad	25
2.5.1. Modelando la retroactividad	27
2.5.2. Cuantificando la retroactividad	28
2.6. Dispositivos de aislamiento para reforzar la modularidad	31
2.7. Retroactividad en procesos bioquímicos	32
2.8. Proceso de fosforilación/desfosforilación	35
3. Análisis estocástico	39
3.1. Introducción	39
3.2. Procesos estocásticos	40
3.3. Proceso de vida y muerte multivariable	42
3.3.1. Método alternativo para determinar la distribución estacionaria	43
3.4. Límite determinista del proceso de vida y muerte multivariable	46
3.5. Métodos de solución de la ecuación maestra	49

3.5.1.	Operaciones matriciales	49
3.5.2.	Transformada de Laplace	50
3.5.3.	Función generatriz	50
3.5.4.	Representación de Poisson	51
3.5.5.	Operadores	52
3.6.	Métodos de simulación: Algoritmo de Gillespie	53
3.6.1.	Método directo	54
3.6.2.	Primera reacción	56
4.	Fluctuaciones intrínsecas en una red bioquímica estocástica	57
4.1.	Ruido vs fluctuación	57
4.2.	Fluctuaciones	58
4.3.	Ley de acción de masas hasta el tercer momento	62
4.4.	Importancia de las fluctuaciones en los circuitos genéticos	65
4.5.	Fluctuaciones extrínsecas e intrínsecas	66
4.6.	Cuantificar fluctuaciones	67
4.6.1.	Toggle switch	67
4.7.	Retroactividad con correcciones estocásticas	72
5.	Fosforilación	75
5.1.	Análisis a partir del algoritmo de Gillespie	75
5.1.1.	Caso I	75
5.1.2.	Caso II	78
5.2.	Análisis alternativo	82
5.2.1.	Caso I	82
5.2.2.	Caso II	83
6.	Función de Hill	87
6.1.	Derivación determinista	87
6.1.1.	Función de Hill usual	87
6.1.2.	Función de Hill con varios tipos de enzimas	89
6.2.	Derivación estocástica	93
6.2.1.	Función de Hill en el límite estacionario	94
6.2.2.	Estados estacionarios a partir de estados coherentes	98
6.2.3.	Variantes de la función de Hill	99
6.2.4.	Proceso estocástico con una función de Hill	103
6.2.5.	Toggle switch	104
6.3.	Función de Hill determinista con correcciones	109
6.4.	Coficiente de Hill decimal	112
6.4.1.	Caso determinista	113
6.4.2.	Caso estocástico	116
7.	Redes estocásticas de regulación genética	123
7.1.	Introducción	123
7.2.	Red estocástica de regulación genética I	124
7.2.1.	Elementos para la simulación	127
7.2.2.	Represilador	129
7.3.	Red estocástica de regulación genética general	133
7.3.1.	Elementos para simulación	135
7.3.2.	Reloj activador-represor	136
8.	Conclusiones y resultados	141

<i>ÍNDICE GENERAL</i>	XIII
A. Teorema de fluctuación disipación	145
B. Función de Hill general	147
C. Cálculo de la deficiencia	151
C.1. Fosforilación	151
C.1.1. Caso I	151
C.2. Redes de regulación genética	152
D. Proceso de vida y muerte	153
E. Modelo estocástico de un MTT	159
E.1. Cálculos para el módulo de transcripción	164
Bibliografía	167

Índice de figuras

1.1. Sistema abierto (izquierda) y cerrado(derecha).	2
1.2. Sistema con control [6].	3
1.3. Sistema con retroalimentación.	5
1.4. Esquematación del proceso de transcripción y traducción [6].	7
1.5. Depredador-presa [6]. En la figura de la izquierda se observa como un depredador atrapa a su presa, en la izquierda la dinámica discreta del modelo de depredador-presa.	7
1.6. Control por retroalimentación [6].	10
1.7. Bloques y funciones de transferencia [6].	13
1.8. Módulo de control.	14
1.9. Control PID [6].	15
1.10. Diagramas de bloque de control PDI [6].	16
1.11. Respuesta de control PDI [6].	16
2.1. Circuitos de genéticos [8].	18
2.2. Proceso de transcripción y traducción [15].	18
2.3. Proceso de transcripción [15].	19
2.4. Regulación genética [15].	20
2.5. Grafo.	21
2.6. Módulo de transcripción/traducción.	22
2.7. Circuitos de genéticos [8].	23
2.8. Regiones de oscilación del represilador [8].	24
2.9. Regiones de oscilación del reloj activador-represor [16].	25
2.10. Perdida del reloj activador-represor [16].	26
2.11. Sistema DT [9].	27
2.12. Desfase por retroactividad [9].	28
2.13. Sistema con retroactividad [9].	29
2.14. Sistemas con retroactividad.	29
2.15. Sistema con retroalimentación [9].	31
2.16. Sistema modular. Separación de un sistema en módulos y los términos de interacción.	32
2.17. Componente aislado [9].	35
2.18. Proceso de fosforilación. Se modela la dinámica de las concentraciones de Z y X_p cuando hay retroactividad (líneas punteadas) y cuando no hay retroactividad (línea verde y roja respectivamente). Se han usado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$ [9].	37
2.19. Proceso de fosforilación. Se modela la dinámica de las concentraciones de Z y X_p cuando hay retroactividad (líneas punteadas) y cuando no hay retroactividad (línea verde y roja respectivamente). Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 0,01$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$ [9].	38

3.1.	Clasificación de procesos estocásticos [10].	41
3.2.	Proceso de Hill. Se modela la dinámica estocástica de las reacciones de donde se deriva la función de Hill.	55
4.1.	Ejemplo de un sistema con control [6].	58
4.2.	Toggle switch.	67
4.3.	Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de simulaciones estocásticas, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_1(0) = m_2(0) = 0$, $n = 2$, $\Omega = 80$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$, tamaño del ensamble= 10000.	69
4.4.	Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_1(0) = m_2(0) = 0$, $n = 2$, $\Omega = 80$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$	70
4.5.	Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, se ha elegido una condición inicial diferente, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_2(0) = 0$, $m_1(0) = 1$, $n = 2$, $\Omega = 80000$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$	71
4.6.	Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = 0$, $n = 2$, $\Omega = 80000$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$	72
5.1.	Componente aislado [9]. En la figura se observa como está aislada la región en donde ocurre el proceso de fosforilación/desfosforilación.	76
5.2.	Dinámica estocástica del proceso de fosforilación y desfosforilación, con $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$	78
5.3.	Dinámica estocástica del proceso de fosforilación y desfosforilación con procesos intermedios, con $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$	81
5.4.	Fosforilación con fluctuaciones. Se modela la dinámica del proceso de fosforilación/desfosforilación a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$	83
5.5.	Fosforilación con procesos intermedios y fluctuaciones. Se modela la dinámica del proceso de fosforilación/desfosforilación con procesos intermedios a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$	85

5.6. Fosforilación con procesos intermedios, mayor retroactividad y fluctuaciones. Se modela la dinámica del proceso de fosforilación/desfosforilación con procesos intermedios a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 0,01$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$ 86

6.1. Función de Hill determinista. En la figura se gráfica a la función de Hill determinista para un activador a diferentes valores de n y con $K = 1$ 88

6.2. Función de Hill determinista. En la figura se gráfica a la función de Hill determinista para un supresor a diferentes valores de n y con $K = 1$ 89

6.3. Caso 1: función de Hill determinista. Se gráfica el valor de la función de Hill cuando las dos enzimas son activadoras, con $K_1 = K_2 = 1$ y $n_1 = n_2 = 1$ 90

6.4. Caso 2: función de Hill determinista. Se gráfica el valor de la función de Hill cuando las dos enzimas son supresoras, con $K_1 = K_2 = 1$ y $n_1 = n_2 = 1$ 91

6.5. Caso 3: función de Hill determinista. Se gráfica el valor de la función de Hill cuando una de las enzimas es activadora y la otra supresora, con $K_1 = K_2 = 1$ y $n_1 = n_2 = 1$. 92

6.6. Función de Hill estocástica. Se realizó la simulación estocástica de las reacciones de donde se obtiene la función de Hill para diferentes tamaños de sistema $\Omega = \{60, 80\}$ y con los siguientes parámetros: $N_0 = 2\Omega$, $k_+ = k_- = \lambda_- = \lambda_+ = 1$, $n = 2$ 97

6.7. Funciones de Hill. En esta figura se gráfica la función de Hill determinista, semi-determinista y semi-estocástica para diferentes valores de n , con $K = 0,01$ y $\Omega = 50$. 103

6.8. Toggle switch. Se realizó la simulación estocástica del toggle switch usando la función de Hill semi-determinista, con $K = 1$, $r_0(0) = r_1(0) = 0$ y $\Omega = 80$ 107

6.9. Gráfica de la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$, $\Omega = 50$ 111

6.10. Esquematización de procesos de unión de proteína-enzima tipo secuencial e independiente. 112

6.11. Función de Hill con unión de proteína-enzima secuencial, se evalúa la función de Hill para distintos valores de n y con $K = 1$ 114

6.12. Función de Hill con unión de proteína-enzima independiente, se evalúa la función de Hill para distintos valores de n y con $K = 1$ 115

6.13. Función de Hill con unión de proteína-enzima secuencial. Se grafica la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$ 119

6.14. Función de Hill con unión de proteína-enzima independiente. Se grafica la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$ 120

7.1. Módulo de transcripción/traducción [8]. 124

7.2. Red genética de la E. Coli [4]. 125

7.3. Represilador. 129

7.4. Represilador. Simulación estocástica usando al algoritmo de Gillespie y la función de Hill determinista, en la primera figura se observa la concentración de ARm, en la segunda la concentración de las proteínas. Se han usado los siguientes parámetros: $n = 2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$ 130

7.5.	Represilador. Dinámica de las concentraciones deterministas y las fluctuaciones a partir de simulaciones estocásticas con tamaño de ensamble= 10000, en las figuras de la parte superior se observa la dinámica de los ARm 's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $Arm_2(0) = r_1(0) = r_2(0) = 0$, $Arm_1(0) = \Omega$, $n=2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$	131
7.6.	Represilador. Se ha modela la dinámica y las fluctuaciones del sistema usando la aproximación determinista y el TFD, en la figura (a) se ha usado la función de Hill determinista, en (b) la función de Hill con correcciones estocásticas. Tanto para (a) y (b) las figuras de la parte superior se observa la dinámica de los ARm 's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $Arm_2(0) = r_1(0) = r_2(0) = 0$, $Arm_1(0) = 1$, $n=2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$	132
7.7.	Módulo de transcripción/traducción [8].	133
7.8.	Reloj activador-represor.	137
7.9.	Reloj activador-represor. Simulación estocástica usando al algoritmo de Gillespie y la función de Hill determinista, en la primera figura se observa la concentración de ARm, en la segunda la concentración de las proteínas. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$ $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{21} = 3$	138
7.10.	Reloj activador-represor. Dinámica de las concentraciones deterministas y las fluctuaciones a partir de simulaciones estocásticas con tamaño de ensamble= 10000, en las figuras de la parte superior se observa la dinámica de los ARm 's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$ $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{21} = 3$	139
7.11.	Reloj activador-represor. Se ha modela la dinámica y las fluctuaciones del sistema usando la aproximación determinista y el TFD, en la figura (a) se ha usado la función de Hill determinista, en (b) la función de Hill con correcciones estocásticas. Tanto para (a) y (b) las figuras de la parte superior se observa la dinámica de los ARm 's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$ $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{21} = 3$	140
B.1.	Proceso de transcripción [15].	147
B.2.	Regulación genética. En la primera figura se observa la síntesis del ARm sin que participen factores de transcripción. En la segunda figura participan factores de transcripción de activación que propician la síntesis del ARm. En la tercera figura participan factores de transcripción de supresión que inhiben la síntesis del ARm.	148
D.1.	Distribución del proceso de vida y muerte de una variable a lo largo del tiempo.	158
E.1.	Módulo de transcripción/traducción [8].	159

Resumen

En este trabajo se analizará el efecto y cambio en la eficiencia de circuitos biológicos debido a las fluctuaciones (estocásticas) intrínsecas. Nos enfocaremos en circuitos biológicos que presenten retroalimentación como su principal mecanismo de funcionamiento. Para ello primero analizaremos y estudiaremos sistemas estocásticos, enfocados principalmente en las redes bioquímicas, además revisaremos algunas técnicas y métodos para simular procesos estocásticos cuantificando las fluctuaciones durante toda su dinámica.

Palabras clave: *Fluctuaciones, redes genéticas, función de Hill, procesos estocásticos, retroactividad, correcciones estocásticas.*

Introducción

La ciencia ayuda a entender el universo y ha propuesto diversos modelos matemáticos, así como varias técnicas de simulación [1] para lograrlo. Los sistemas biológicos siguen siendo parte del universo por lo que estos no escapan de la ciencia, es relevante entender a estos sistemas, principalmente para la conservación del ambiente y el cuidado de la salud, esto último estrechamente relacionado con el desarrollo de fármacos y tratamientos para distintos padecimientos. Sorprendentemente muchos de los modelos que se han desarrollado describen muy bien los sistemas biológicos, sin embargo, aún estamos lejos de comprenderlos completamente.

El estudio de los sistemas biológicos se hace en diferentes estratos, desde el estudio del ecosistema, los individuos o comunidades que los componen hasta los componentes celulares y genéticos más elementales [2], cada una de estas áreas tiene sus propios métodos y técnicas de estudio. En este trabajo nos enfocaremos en las redes bioquímicas, estas entran dentro un área conocida como biología sintética [3], que se encarga de diseñar y rediseñar circuitos biológicos, recientemente ha cobrado gran importancia debido al auge de la edición genética y el diseño de fármacos, por lo que esta nueva área intenta dar luz a los procedimientos experimentales ya hechos, además de idear nuevas técnicas. Uno de los trabajos que empezó a formalizar esta área es el diseño de un circuito genético con oscilaciones [4], las oscilaciones son importantes debido a que estos pueden llegar a jugar un papel similar a un reloj, un ejemplo de esto es el ciclo circadiano [5].

Uno puede llegar a asumir que los sistemas se pueden describir a través de módulos, de tal manera que la dinámica de los sistemas se puede describir a partir de estas componentes. Imagine una computadora compuesta de varios módulos tales como la pantalla, el teclado, la memoria RAM, etc., en estos sistemas puede que exista retroalimentación entre los diferentes componentes o con el entorno. La retroalimentación se puede llegar a dar a través de diferentes señales, en el caso de la computadora serían señales eléctricas. Es común que la retroalimentación juegue un papel de controlador. Es decir, en algunos sistemas las perturbaciones debido a la retroalimentación podrían afectar la dinámica del sistema, pero dentro del sistema existen módulos que cancelan o minimizan estas perturbaciones. Existen diversos ejemplos de control, como por ejemplo el cultivo de algunas bacterias en el cual se tiene que controlar la temperatura. El uso de control por retroalimentación tiene sus ventajas como: robustez ante perturbaciones, establecer una dinámica en el sistema, además de altos niveles de automatización, tal como se ha visto últimamente en automóviles autónomos [6]. Los sistemas biológicos también pueden llegar a tener retroalimentación [7, 8] y se han ideado técnicas para entender e intentar reducir el efecto de las perturbaciones en el sistema, así como intentar dar una medida de esta cuando se conectan dos diferentes módulos [9], de tal manera que se recupere la asunción de modularidad, con el fin de que estos sean más manejables. Este tipo de circuitos se han estudiado principalmente en un enfoque determinista dando buenos resultados. Sin embargo, debido a la baja concentración de proteínas dentro de la célula, las fluctuaciones estocásticas juegan un papel relevante. Es por este motivo que el área de bioquímica estocástica ha cobrado importancia [7, 10, 11].

En este trabajo se analizará el efecto y cambio en la eficiencia de circuitos biológicos debido a las fluctuaciones (estocásticas) intrínsecas. Nos enfocaremos en circuitos biológicos que presenten retroalimentación como su principal mecanismo de funcionamiento. La comprensión del efecto de las fluctuaciones en los circuitos biológicos permitirá avanzar en técnicas para el eventual diseño de computadoras biológicas [4].

Nuestro trabajo está estructurado de la siguiente manera:

En el capítulo 1 hablaremos sobre teoría de control, con el fin de entender cómo funcionan los sistemas modulares, principalmente enfocados en las técnicas de retroalimentación por control, además describiremos los conceptos de control y retroalimentación, así como las ecuaciones que describen estos sistemas, y finalmente hablaremos de algunas técnicas de control que pueden ser aplicados a algunos sistemas.

En el capítulo 2 haremos una introducción a la biología sintética, nos enfocaremos principalmente en el diseño y descripción de circuitos genéticos (principalmente redes de regulación genética), además analizaremos un artículo que describa la forma de cuantificar la retroalimentación en las redes bioquímicas [9].

En el capítulo 3 daremos algunas herramientas matemáticas para poder estudiar las fluctuaciones presente en los sistemas estocásticos, nos enfocaremos en la ecuación química maestra que tendrá un papel relevante en el trabajo, deseamos que nuestros modelos estocásticos converjan a uno determinista entonces explicamos una manera de obtener el límite determinista basándonos de la idea de ensamble, finalmente describiremos algunos algoritmos de simulación que se basan en procesos estocásticos, enfocados principalmente en el algoritmo de Gillespie.

En el capítulo 4 definiremos a las fluctuaciones y su uso en sistemas estocásticos, aquí también obtendremos una expresión que es una variante del teorema de fluctuación-disipación (TFD), lo que nos permitiría cuantificar las fluctuaciones de nuestros sistemas. También explicaremos que son las fluctuaciones intrínsecas, debido a que los modelos que analizaremos son no lineales, nos apoyaremos del algoritmo de Gillespie para realizar las simulaciones y a partir de estas inferir las fluctuaciones en el sistema, para ello escribiremos un programa en Python. Además, haremos un programa de Python para inferir las fluctuaciones con ayuda del TFD.

En el capítulo 5 analizaremos el proceso de fosforilación/desfosforilación que aparece en el artículo de Del Vecchio [9], es posible que este sistema sea pequeño, entonces se requeriría de una descripción estocástica, para esto construimos sus correspondientes ecuaciones maestras y de esta manera simular al sistema empleando el algoritmo de Gillespie, finalmente usaremos a la aproximación determinista y a la variante del TFD para modelar y simular este sistema.

En el capítulo 6 nos enfocaremos en la función de Hill, la cual resulta ser de suma importancia al estudiar varios sistemas químicos y biológicos, debido a que describe la unión de proteínas y enzimas (o procesos similares), particularmente aparece en las redes de regulación genética, pero estos sistemas son pequeños por lo que se necesita una descripción estocástica, sin embargo, cuando se usa la función de Hill en los modelos estocásticos no queda claro el cómo se introducen. Primero derivaremos esta función a partir de una perspectiva determinista cuando actúa solo una enzima, así como en el caso que participen varios tipos, posteriormente daremos una nueva derivación con ayuda de los procesos estocásticos y algunas variaciones que se pueden encontrar. También mostraremos como aparece esta función dentro de una reacción bioquímica, además discutiremos y analizaremos algunos elementos importantes de esta función como es el caso del coeficiente de Hill.

En el capítulo 7 presentamos y desarrollamos algunos modelos de redes estocásticas de regulación genética, debido a que muchos de estos sistemas son pequeños y presentan fluctuaciones, es necesaria una descripción estocástica, por lo que para construirlos nos apoyamos de la ecuación maestra y de la función de Hill desarrollada en el capítulo 6, además desarrollaremos un modelo computacional para realizar simulaciones de este sistema, y realizaremos las simulaciones correspondientes. También nos interesa que el modelo estocástico converja al determinista, por lo que implementamos el límite determinista descrito en el capítulo 3, y de paso cuantificamos las fluctuaciones del sistema, posteriormente compararemos lo que se obtendría de la función de Hill usual y la desarrollada en este trabajo.

En el capítulo 8 discutiremos con más detalle los resultados obtenidos a lo largo del documento, daremos las conclusiones y posibles trabajos futuros.

En el apéndice hemos resuelto de manera analítica la ecuación maestra asociada al proceso de vida y muerte de una variable haciendo uso del formalismo de operadores, donde hemos encontrado la función generatriz, el cual resulta ser un elemento muy importante al momento de calcular los promedios y momentos de la variable involucrada. Además, hay una deducción del teorema de fluctuación disipación a partir de la aproximación lineal de ruido. Explicamos también el cómo aparece la función de Hill en las redes de regulación genética a partir de un conjunto de reacciones bioquímicas.

A lo largo del documento se emplea Python para simular la dinámica determinista o estocástica de algunos sistemas, por lo que hemos creado una carpeta para que se pueda acceder a los archivos usados (https://drive.google.com/drive/folders/1DcDw7D0B-hq-9gxt-8z4P3_K63m1eU01?usp=sharing), para realizar las simulaciones hemos empleado una PC con un procesador Core i5-10300H y Python 3.8.12, recomendamos usar Visual Studio Code para ejecutar los programas desarrollados.

Capítulo 1

Teoría de control

En este primer capítulo daremos algunos conceptos básicos que conciernen al estudio de sistemas con retroalimentación y teoría de control, algunas de las ideas y estructura del desarrollo de los temas así como imágenes fueron tomados de [6]. Esta área tiene un papel sumamente importante en nuestro mundo moderno. Los desarrollos que conciernen a la teoría de retroalimentación y de control abarcan a sistemas relativamente simples como el encender un televisor hasta tareas más complejas como el control de vehículos autónomos. Nos enfocamos en revisar los temas relacionados a control para entender mejor los sistemas modulares y como se construyen las ecuaciones que los describen (ecuaciones diferenciales), para poder aplicarlo y entender mejor a las redes de procesos bioquímicos.

1.1. ¿Qué es retroalimentación?

Para empezar, es necesario dar el concepto de sistema complejo, sin embargo, una definición precisa es complicada, para los propósitos del trabajo es suficiente lo siguiente.

Los sistemas complejos son redes formadas por una serie de componentes que interactúan entre sí, normalmente de forma no lineal. Los sistemas complejos pueden surgir y evolucionar a través de la autoorganización, de modo que no sean completamente regulares ni completamente aleatorios, lo que permite el desarrollo de un comportamiento emergente a escalas macroscópicas [1].

Anteriormente, los estudios realizados en sistemas complejos, casi siempre se asumía que un determinado sistema complejo puede ser descompuesto en componentes más pequeños para analizarlos de manera individual y posteriormente juntar todo nuevamente, de esta manera uno puede llegar a entender por completo el sistema, aunque como hoy es bien sabido este enfoque es insuficiente, un problema donde se puede realizar este enfoque es el caso de los sistemas eléctricos compuestos de múltiples partículas cargadas. Con la creación de las maquinas uno fue pensando que los sistemas están formados por componentes (quizá engranajes), estos funcionan por medio de la interacción de diversos componentes al cual se le ha asignado alguna determinada función. Por ejemplo, puede pensar en un televisor, que tiene diferentes componentes y cada uno de estos ejecuta alguna determinada función (procesar imagen, recibir información), a estos componentes de un sistema se les nombra módulos, como veremos más adelante, los módulos tomaran un papel importante en nuestro trabajo. Una definición más precisa de modulo es:

modulo: elemento con función propia concebido para poder ser agrupado de distintas maneras con otros elementos constituyendo una entidad mayor [12].

Con este concepto podemos suponer que un sistema complejo está compuesto de varios módulos, los módulos pueden englobar diferentes componentes del sistema complejo (esta idea es suficiente para los propósitos de nuestro trabajo) o simplemente ser un componente, los módulos también pueden ser sistemas complejos, además, es posible que exista una comunicación entre estas y/o su entorno a través de diversos canales. Para continuar es necesaria la definición de retroalimentación, como a continuación.

Retroalimentación: Es la comunicación a través de estímulos entre los módulos y/o con su entorno.

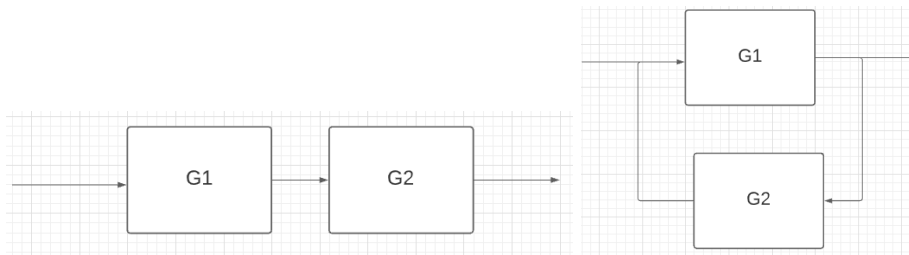


Figura 1.1: Sistema abierto (izquierda) y cerrado(derecha).

También se puede entender a la retroalimentación como interacción, aunque este punto de vista es insuficiente puede llegar a ser didáctico para entender mejor este concepto. Los estímulos pueden llegar a ser de diversa índole, tal como señales eléctricas, presión, entre otros. La conexión entre los diversos módulos puede llegar a ser física, por ejemplo: cables, o por medio de otros mecanismos (en el caso de imágenes se transmitirían ondas electromagnéticas a través del espacio-tiempo), la conexión entre estos es el canal a través del cual el estímulo es transmitido. Dentro de los sistemas con retroalimentación podemos clasificarlos en dos grandes grupos, los abiertos y cerrados (algunos sistemas pueden contener ambas), como en la figura 1.1. En algunos sistemas con retroalimentación es complicado describir la causalidad, principalmente en los sistemas circulares, ya que los módulos son causa y consecuencia, como en la figura 1.1, ya que $G1$ y $G2$ se envían señales el uno al otro y es difícil saber quien causa a quien.

Existen diferentes ejemplos en donde podemos encontrar sistemas con retroalimentación, pero como nos enfocamos en sistemas biológicos, entonces un buen ejemplo seria la regulación de glucosa en la sangre a través de la producción de insulina y glucosa en el páncreas. Existen otros ejemplos más cotidianos, imagine que está usando un telescopio, entonces al ver a través de la lente ve una imagen borrosa, por lo que ajustará la imagen hasta que quede lo más clara posible. Con este ejemplo cotidiano nos damos cuenta de que en realidad gran parte de las actividades que realizamos y las maquinas que hemos construido funcionan a través de retroalimentación.

1.2. ¿Qué es control?

El termino control tiene diversas definiciones según sea el campo de estudio. Nosotros nos conformaremos con la siguiente definición

Control: el uso de algoritmos y/o retroalimentación en un sistema dinámico para cumplir una determinada función.

Existen diversos ejemplos de control, como por ejemplo el cultivo de algunas bacterias en cual se tiene que controlar la temperatura. El termino módulo de control se refiere a insertar un

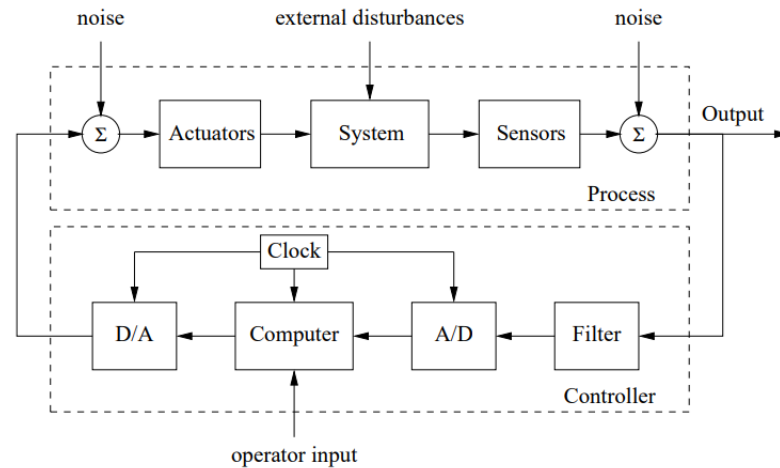


Figura 1.2: Sistema con control [6].

nuevo módulo a algún sistema, donde este tiene la función de controlar alguna variable o varias de estas. Aplicaciones novedosas de esto son los vehículos autónomos y los robots, estos sistemas reciben alguna retroalimentación del entorno o el ambiente. Actualmente se hace mucho uso de la retroalimentación cerrada para control, como el que se presenta en la figura 1.2, en esta se presenta a una computadora rudimentaria. Usar este tipo de retroalimentación en los sistemas ha resultado ventajoso, debido a que estos sistemas tienden a ser estables bajo perturbaciones. En la definición de control se hace uso de retroalimentación, por lo que a veces es complicado discernir si un sistema tiene retroalimentación o control (principalmente en retroalimentación cerrada), por lo que solo nos referiremos a módulo de control a aquellos que mantienen estables a un sistema, aunque estas cuestiones no nos preocuparan en lo que sigue del trabajo.

1.3. Usos de retroalimentación

El uso y/o existencia de retroalimentación en un sistema hace que este posea ciertas propiedades que los hacen muy útiles, tales como las siguientes:

1. **Robustez ante perturbaciones:** Midiendo la diferencia entre el valor detectado de una señal regulada y su valor deseado, podemos proporcionar una acción correctiva.
2. **Diseño de la dinámica:** A través de la retroalimentación es posible cambiar la dinámica de un sistema. Mediante la retroalimentación, podemos alterar el comportamiento de un sistema, por ejemplo: los sistemas que son inestables se pueden estabilizar, los sistemas que son lentos se pueden hacer que respondan rápido.
3. **Altos niveles de automatización:** Una tendencia importante en el uso de la retroalimentación es su aplicación a niveles más altos de conciencia situacional y toma de decisiones. Esto incluye no solo la ramificación lógica tradicional basada en las condiciones del sistema, sino también la optimización, la adaptación, el aprendizaje e incluso niveles más altos de razonamiento abstracto. Estos problemas están en el dominio de la comunidad de inteligencia artificial.

Las propiedades anteriores son ventajas, pero también pueden llegar a existir ciertas desventajas en la retroalimentación, como las que enumeramos a continuación,

1. **Retroalimentación positiva y negativa:** La retroalimentación negativa, es cuando intentamos regular el sistema ante las perturbaciones de una manera que disminuya el efecto de esas perturbaciones. En un sistema con retroalimentación positiva, el aumento de alguna variable o señal conduce a una situación en la que esa cantidad se incrementa aún más a través de su dinámica. Esto tiene un efecto desestabilizador, aunque a menudo se considera indeseable, este comportamiento se utiliza en sistemas biológicos (y de ingeniería) para obtener una respuesta muy rápida a una condición o señal.
2. **Inestabilidad:** Si el sistema no está diseñado correctamente puede ser inestable. Esto es complicado porque debemos diseñar el sistema no solo para que sea estable en condiciones nominales, sino también para que permanezca estable en todas las posibles perturbaciones de la dinámica.
3. **Feedforward:** En algunas circunstancias es posible medir una perturbación antes de que ingrese al sistema, y esta información se puede utilizar para tomar medidas correctivas antes de que la perturbación haya influido en el sistema. El efecto de la perturbación se reduce así midiéndola y generando una señal de control que la contrarreste. Este tipo de control en un sistema se denomina feedforward. Dado que el feedforward intenta hacer coincidir dos señales, se requieren buenos modelos de procesos; de lo contrario, las correcciones pueden tener un tamaño incorrecto o estar mal sincronizadas.

1.4. Análisis de sistemas con retroalimentación y control

El uso de los sistemas con retroalimentación tiene gran importancia en el diseño de sistemas físicos, debido a algunas propiedades que estos tengan, tal como se vio en la sección anterior. Por lo que a continuación daremos un marco matemático para su estudio, así como algunos ejemplos o aplicaciones.

1.4.1. Marco matemático

A continuación expondremos dos métodos para analizar los sistemas con retroalimentación, uno continuo y uno discreto (respecto al tiempo), con ayuda de ecuaciones diferenciales y ecuaciones en diferencias respectivamente, los dos sistemas hacen uso de conceptos tales como (con ayuda de la figura 1.3 daremos estos conceptos):

1. **Entrada:** Se refiere al estímulo de entrada, y se caracteriza por una función u . En el caso que existan r estímulos de entrada vamos a caracterizar a cada uno por u_j (con $j \in \{1, 2, \dots, r\}$).
2. **Estado:** se refiere al estado en la que se encuentra una variable x , que caracteriza a nuestro sistema. En el caso que nuestro sistema este caracterizado por n variables vamos a caracterizar a cada uno por x_i (con $i \in \{1, 2, \dots, n\}$).
3. **Dinámica:** Se refiere a cómo cambia nuestra variable x , con respecto a sí misma y a la entrada, comúnmente asociada a una función $f(u, x)$. Si nuestro sistema esta caracterizado por n variables entonces cada una va evolucionar con respecto a una función $f_k(x_i, x_j)$ (con $k \in \{1, 2, \dots, n\}$).
4. **Salida:** se refiere a que el sistema da un determinado valor de x , posteriormente medida (detectada por el experimento), esta medición se caracteriza por una función $y(x, u)$. Si el estado de nuestro sistema esta descrito por n variables, por lo general solo vamos a detectar n' de ellas ($n' \leq n$), cada una de estas variables y_l tendrá asociada una medición $h_l(x_i, u_j)$ (con $l \in \{1, 2, \dots, n\}$).

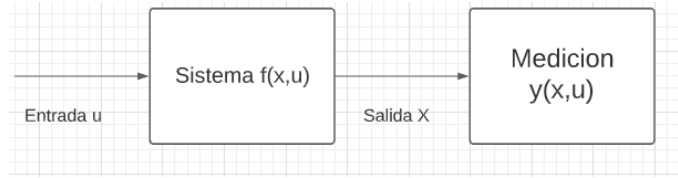


Figura 1.3: Sistema con retroalimentación.

Ecuaciones diferenciales

Ahora las palabras que dimos al principio de esta sección las convertiremos en ecuaciones, para el caso continuo tenemos lo siguiente:

$$\begin{aligned}\frac{dx_i}{dt} &= f_i(x_k, u_j), \\ y_l &= h_l(x_i, u_j),\end{aligned}\tag{1.1}$$

lo que nos dicen estas ecuaciones es que nuestro sistema puede estar caracterizado por un vector x_i , y su dinámica la gobierna alguna función $f_i(x_k, u_j)$, donde su forma depende del sistema que se esté analizando y del estímulo de entrada. Hemos usado un sistema de ecuaciones de primer orden, ya que como es bien sabido, si la dinámica está gobernada por ecuaciones diferenciales de orden superior, estas siempre pueden ser reducidas a un conjunto de ecuaciones diferenciales de primer orden. Las mediciones que aquí pusimos son de un índice diferente, ya que en determinados sistemas no se podrán determinar todas las variables, o simplemente se encuentra una función de estas.

Ecuaciones en diferencias

En algunos sistemas es más natural que el sistema evolucione a tiempos discretos, en múltiplos de alguna cantidad, sus múltiplos serán caracterizados por una variable k , por lo que el marco natural para estas son las ecuaciones en diferencias, en base a la imagen que esta al principio de esta sección, estas ecuaciones serán las siguientes:

$$\begin{aligned}x_i(k+1) &= F_i(x_i(k), u_j(k)), \\ y_l(k) &= h_l(x_i(k), u_j(k)).\end{aligned}\tag{1.2}$$

A pesar de ser una ecuación en diferencias, está todavía tiene algunas de las características que dimos en el sistema continuo. Algunas ecuaciones en diferencias pueden ser derivadas de las ecuaciones diferenciales, o en el sentido inverso, sin embargo, algunas veces no se llega a capturar todo lo relevante del sistema [1].

Ley de acción de masas

Una de las ideas más usados para realizar el análisis de un proceso en un conjunto de módulos o componentes que interaccionan entre sí, es la ley de acción de masas. Aunque esta idea solo aplica a procesos que están cerca del equilibrio, esta ha resultado ser muy útil. Comúnmente se utiliza en química, pero estas ideas pueden ser aplicadas en otros campos que tengan una estructura y/o interacción similar a la que esta en los procesos químicos.

Para esto imagine que hay un proceso a través en el cual hay N especies S_j ($j \in \{1, 2, \dots, N\}$), y M reacciones \mathcal{R}_i ($i \in \{1, 2, \dots, M\}$) a través de la cual se transforman estas especies, es decir



Los coeficientes α_{ij} y β_{ij} son enteros positivos, estos son los coeficientes estequiométricos. En química los elementos que están del lado izquierdo de (1.3) se les conoce como reactantes y a los del lado derecho como productos. Ahora definimos la matriz estequiométrica

$$\Gamma_{ji} = \beta_{ij} - \alpha_{ij}. \quad (1.4)$$

Denotamos por x_j a la concentración de la especie S_j involucrada en la reacción, R_i a la forma algebraica de la i -enésima tasa de reacción en función de las concentraciones de las especies. La suposición más común es la acción de masas, en el caso determinista la tasa de reacción toma la siguiente forma

$$R^i = k_i \prod_{j=1}^N x_j^{\alpha_{ij}}, \quad (1.5)$$

que dice simplemente que la velocidad de reacción es proporcional al producto de las concentraciones de los reactivos, con exponentes más altos cuando se necesita más de una molécula. Los coeficientes k_i se denominan "constantes de reacción"; normalmente etiquetan las flechas en los diagramas (1.3) [8]. En el caso determinista uno supone que la razón de cambio de la concentración de las especies está dada por un sistema de ecuaciones diferenciales,

$$\frac{dx_j}{dt} = \Gamma_{ji} R^i. \quad (1.6)$$

Como más adelante veremos, la ley de acción de masas puede ser aplicada al caso estocástico. Como estamos enfocados en sistemas con retroalimentación, nuestra ecuación del caso clásico puede tener alguna entrada (retroalimentación) que modifique su estado, además de presentar otros fenómenos como la difusión, la modificación del estado estará dado por alguna función $f_j(x_j, u_l)$ por lo que ahora la dinámica de nuestro sistema será

$$\frac{dx_j}{dt} = \Gamma_{ji} R^i + f_j(x_j, u_l). \quad (1.7)$$

La ley de acción de masas es un muy buen enfoque para analizar sistemas con retroalimentación, principalmente en procesos de transcripción, por lo que (1.7) será una ecuación relevante para nuestro trabajo.

1.4.2. Ejemplos

A continuación daremos un ejemplo para sistemas continuos y discretos respectivamente. Nos enfocaremos principalmente en sistemas biológicos. Los dos ejemplos son para sistemas con retroalimentación.

Caso continuo

En este ejemplo analizaremos la regulación de la transcripción. La transcripción es el proceso por el cual el ARN mensajero (ARNm) se transcribe de algún segmento de ADN. La región promotora de un gen permite que la transcripción sea controlada por la presencia de otras proteínas, que se unen a la región promotora y reprimen o activan la ARN polimerasa, la enzima que produce una transcripción de ARNm a partir del ADN. Luego, el ARNm se traduce en una proteína de acuerdo con su secuencia de nucleótidos. Como se observa en la figura 1.4.

Un modelo simple del proceso de regulación transcripcional es mediante el uso de una función de Hill [7]. Considere la regulación de una proteína A con una concentración dada por p_a y una concentración de ARNm correspondiente m_a . Sea B una segunda proteína con concentración p_b que

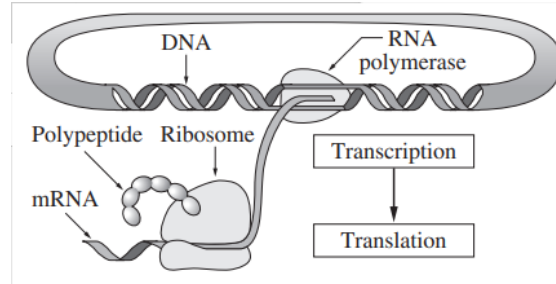


Figura 1.4: Esquemización del proceso de transcripción y traducción [6].

reprime la producción de proteína A mediante regulación transcripcional. La dinámica resultante de p_a y m_a puede ser

$$\begin{aligned} \frac{dp_a}{dt} &= \beta_a m_a - \delta_a p_a, \\ \frac{dm_a}{dt} &= \frac{\alpha_{ab}}{1 + k_{ab} p_a^n} + \alpha_{a0} - \gamma_a m_a \end{aligned} \quad (1.8)$$

El parámetro β_a se refiere al factor de traducción de m_a , tanto el parámetro δ_a y γ_a representan el rango de degradación, la ecuación para m_a tiene asociada la función de Hill $\frac{\alpha_{ab}}{1 + k_{ab} p_a^n}$ para la supresión de la producción de la proteína A , en el capítulo 6 damos algunos detalles de esta función.

Caso discreto

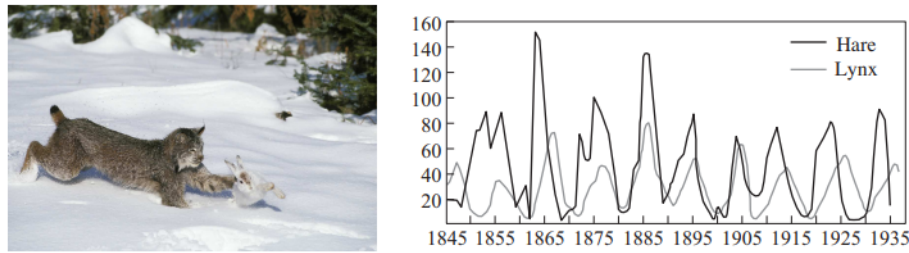


Figura 1.5: Depredador-presa [6]. En la figura de la izquierda se observa como un depredador atrapa a su presa, en la izquierda la dinámica discreta del modelo de depredador-presa.

Aquí veremos el modelo discreto de depredador-presa. Imagine que en un espacio solo viven dos especies, donde una se come a la otra, por ejemplo, Lobos L y conejos C , los conejos nacen y mueren, la causa de su muerte es debida a los lobos, de igual manera los lobos perecen a un ritmo constante. Con esto en mente proponemos el siguiente modelo:

$$\begin{aligned} C(k+1) &= C(k) + b_r(u)C(k) - aL(k)C(k), \\ L(k+1) &= L(k) - d_f L(k) + cL(k)C(k), \end{aligned} \quad (1.9)$$

la función $b_r(u)$ nos habla sobre el nacimiento de los conejos y como estos dependen del entorno, como la disposición de comida, la variable a el ritmo en el que son cazadas por los lobos o un término de interacción, esta interpretación es similar para la variable c , d_f es el coeficiente de defunción de los lobos.

Un buen ejemplo de lo que se obtiene en este sistema está en la figura 1.5, ahí podemos observar cómo la población de ambos va oscilando, es decir, uno puede pensar que se regulan el uno al otro, aunque este resultado depende del valor de las constantes del modelo, ya que para determinados valores ocurre bifurcación, al igual que sucedería en el caso continuo.

1.5. Retroalimentación del estado

En este capítulo expondremos como se puede estabilizar el estado de un sistema usando retroalimentación, como es necesario que el sistema este cerca de un punto estable, además el valor de la entrada u es prácticamente estacionario, linealizaremos las ecuaciones que describen nuestro sistema, posteriormente expondremos como se logra esta estabilización usando un modelo sencillo.

1.5.1. Linealizar sistemas con retroalimentación

Como hemos mencionado en la sección anterior, algunos sistemas pueden estar descritos por ecuaciones diferenciales o ecuaciones en diferencias, en los sistemas con control o retroalimentación, siempre es deseado que se encuentren cerca de algún punto de equilibrio (aunque podemos elegir otros puntos, pero sería ligeramente más complicado), por lo que es bastante útil linealizar nuestro sistema, para ello lo vemos como se realiza la linealización en los dos casos.

Ecuaciones diferenciales

Como ya hemos mencionado, nuestro sistema es descrito por un sistema de ecuaciones diferenciales de la forma

$$\begin{aligned}\frac{dx_i}{dt} &= f_i(x_i, u_j), \\ y_l &= h_l(x_i, u_j).\end{aligned}\tag{1.10}$$

El punto de equilibrio de este sistema de ecuaciones diferenciales se encuentra al resolver la siguiente ecuación, considerando que no haya retroalimentación e igualar la relación anterior a cero,

$$\frac{dx_{i,eq}}{dt} = 0 = f_i(x_{i,eq}, 0),\tag{1.11}$$

al ya encontrar el punto de equilibrio, ahora nos interesa describir la dinámica alrededor de este punto, en general un sistema no tiene un solo punto de equilibrio, y estos pueden ser estables o inestables, para más detalles ver [1]. Suponemos que nos encontramos cerca del punto de equilibrio $(x_{i,eq}, 0)$ (el cero que aparece quiere decir que no hay ningún estímulo externo u_j), por lo que definimos una variable $z_i = x_i - x_{i,eq}$, despejando, $x_i = z_i + x_{i,eq}$, sustituyendo en la primera expresión de (1.10), obtenemos

$$\frac{dz_i}{dt} = f_i(z_i + x_{i,eq}, u_j).\tag{1.12}$$

Realizando una expansión a primer orden de la función que está en el lado izquierdo

$$f_i(z_i + x_{i,eq}, u_j) \approx f_i(x_{i,eq}, 0) + z_j \left. \frac{df_i(x_i, u_j)}{dx_j} \right|_{x_i=x_{i,eq}, u_j=0} + u_j \left. \frac{df_i(x_i, u_j)}{du_j} \right|_{x_i=x_{i,eq}, u_j=0},\tag{1.13}$$

Los términos que involucran derivas son de hecho Jacobianos, procediendo de manera similar con la otra expresión de (1.10), definimos las siguientes matrices

$$\begin{aligned}A_{ij} &= \left. \frac{df_i(x_i, u_j)}{dx_j} \right|_{x_i=x_{i,eq}, u_j=0}, & B_{ij} &= \left. \frac{df_i(x_i, u_j)}{du_j} \right|_{x_i=x_{i,eq}, u_j=0}, \\ C_{lj} &= \left. \frac{dh_l(x_i, u_j)}{dx_j} \right|_{x_i=x_{i,eq}, u_j=0}, & D_{lj} &= \left. \frac{dh_l(x_i, u_j)}{du_j} \right|_{x_i=x_{i,eq}, u_j=0}.\end{aligned}\tag{1.14}$$

Sustituyendo estas expresiones en (1.12), nuestro sistema estará descrito por el siguiente conjunto de ecuaciones diferenciales linealizadas

$$\begin{aligned}\frac{dz_i}{dt} &= A_{ij}z_j + B_{ij}u_j, \\ y_l &= C_{lj}z_j + D_{lj}u_j.\end{aligned}\tag{1.15}$$

Este conjunto de ecuaciones diferenciales es el componente principal de los sistemas con retroalimentación, a partir de los elementos que la componen se pueden determinar una gran cantidad de información respecto a la dinámica del sistema.

Ecuaciones en diferencias

Para el caso discreto nuestro sistema es descrito por las siguientes ecuaciones,

$$\begin{aligned}x_i(k+1) &= F_i(x_i(k), u_j(k)), \\ y_l(k) &= h_l(x_i(k), u_j(k)).\end{aligned}\tag{1.16}$$

Como nos interesa estar cerca del punto de equilibrio $(x_{i,eq}(k), 0)$ (el cero que aparece quiere decir que no hay ningún estímulo externo u_j), lo primero es conocer el punto de equilibrio, para ello uno tiene que resolver la siguiente ecuación, suponiendo que no haya retroalimentación

$$x_{i,eq}(k+1) = F_i(x_{i,eq}(k), 0),\tag{1.17}$$

con la característica que $x_{i,eq}(k+1) = x_{i,eq}(k)$, después de encontrar el punto de equilibrio ahora nos interesa describir la dinámica alrededor de este punto, en general un sistema no tiene solo un punto de equilibrio, y estos pueden ser estables o inestables, para más detalles ver [1], por ahora nosotros pensaremos que son puntos de equilibrio estables. Por lo anterior suponemos $\Delta x_i(k+1) = x_i(k+1) - x_{i,eq}$, despejando, $x_i(k+1) = x_{i,eq} + \Delta x_i(k+1)$, sustituyendo en la primera expresión de (1.16)

$$x_{i,eq} + \Delta x_i(k+1) = F_i(x_{i,eq} + \Delta x_i(k), u_j(k)).\tag{1.18}$$

Ahora hacemos la expansión a primer orden de la parte izquierda

$$F_i(x_{i,eq} + \Delta x_i(k), u_j(k)) \approx F_i(x_{i,eq}, 0) + \Delta x_i(k) \left. \frac{dF_i(x_i, u_j)}{dx_i} \right|_{x_i=x_{i,eq}, u_j=0} + u_i(k) \left. \frac{dF_i(x_i, u_j)}{du_i} \right|_{x_i=x_{i,eq}, u_j=0}.\tag{1.19}$$

los términos que aparecen son matrices Jacobianas, con respecto a cada una de las variables que depende la función, realizando un procedimiento similar con la segunda expresión de (1.16), definimos a cada una de estas como sigue

$$\begin{aligned}A_{ij} &= \left. \frac{dF_i(x_i, u_j)}{dx_j} \right|_{x_i=x_{i,eq}, u_j=0}, & B_{ij} &= \left. \frac{dF_i(x_i, u_j)}{du_j} \right|_{x_i=x_{i,eq}, u_j=0}, \\ C_{lj} &= \left. \frac{dh_l(x_i, u_j)}{dx_j} \right|_{x_i=x_{i,eq}, u_j=0}, & D_{lj} &= \left. \frac{dh_l(x_i, u_j)}{du_j} \right|_{x_i=x_{i,eq}, u_j=0}.\end{aligned}\tag{1.20}$$

Sustituyendo cada una de estas expresiones en (1.18), y considerando (1.17), por lo que ahora nuestro sistema estará descrito por las siguientes ecuaciones linealizadas

$$\begin{aligned}\Delta x_i(k+1) &= A_{ij}\Delta x_j(k) + B_{ij}u_j(k), \\ y_l(k) &= C_{lj}\Delta x_j(k) + D_{lj}u_j(k).\end{aligned}\tag{1.21}$$

Como es bien sabido, los eigenvalores asociadas a las matrices que definimos, tanto en el caso continuo como discreto nos dicen como se comportara nuestro sistema, es decir la dinámica cerca de los puntos de equilibrio, además del tipo de punto estacionario.

1.5.2. Estabilización del estado por retroalimentación

Como ya hemos linealizado nuestro sistema de ecuaciones, ahora veremos como estabilizar nuestro sistema usando retroalimentación, para ello nos enfocaremos en un sistema continuo, para simplificar un poco los cálculos nos fijaremos en un sistema parecido a la figura 1.6. Tal y como está en la figura nuestro sistema esta descrito por un sistema de ecuaciones dado por

$$\begin{aligned}\frac{dz_i}{dt} &= A_{ij}z_j + B_{ij}u_j, \\ y_l &= C_{lj}z_j + D_{lj}u_j.\end{aligned}\tag{1.22}$$

Este sistema se toma linealizado, ya que la función del módulo de control es mantener al sistema lo más cercano al punto de equilibrio.

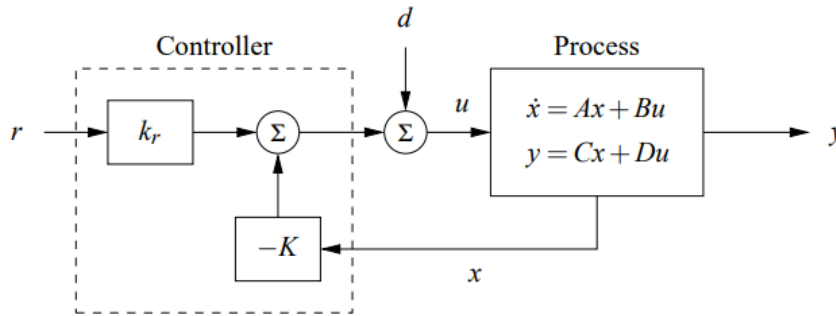


Figura 1.6: Control por retroalimentación [6].

Los símbolos que aparecen en nuestro sistema quieren decir lo siguiente:

1. Los cuadros amplifican k_r o $-K$ veces la señal de entrada ($k_r, K \in \mathbb{R}$).
2. El círculo que encierra a la variable Σ , significa la suma lineal de todas las señales de entrada.

Con esto en mente ahora se encuentra que u tiene la siguiente forma

$$u_j = -Kx_j + k_r r_j + d_j,\tag{1.23}$$

donde x_j representa a las variables del cuadro proceso, r_j es un valor de referencia en el cual debería estar y , y d_j representa al ruido que entra al sistema. Se supone que nosotros controlamos la retroalimentación, tal como en la figura 1.6, por lo que si damos algún valor de la K , entonces podremos determinar el valor de la k_r , de esta manera el sistema se mantiene en el estado estacionario aún cuando existan perturbaciones o interacciones con el entorno, ahora intentaremos determinar el valor de k_r , para ello sustituimos esta última expresión que hemos obtenido en (1.22), y considerando que ya estamos en el punto de equilibrio, por lo que ahora tendríamos lo siguiente

$$0 = A_{ij}x_{j,eq} + B_{ij}(-Kx_{j,eq} + k_r r_j + d_j),$$

despejando se obtiene

$$x_{j,eq} = -(A_{ij} - KB_{ij})^{-1}B_{ij}(k_r r_j + d_j),\tag{1.24}$$

si suponemos que $y_j = r_j$ (el valor deseado de la salida), entonces tendremos la siguiente relación

$$r_l = C_{lj}x_{j,eq} + D_{lj}u_j = C_{lj}x_{j,eq} + D_{lj}(-Kx_{j,eq} + k_r r_j + d_j),$$

sustituyendo el valor de $x_{j,eq}$

$$r_l = (-C_{lj}(A_{ij} - KB_{ij})^{-1}B_{ij} + D_{lj}(K(A_{ij} - KB_{ij})^{-1}B_{ij} + \delta_{jj}))(k_r r_j + d_j), \quad (1.25)$$

(δ_{jj} es la delta de Kronecker) definimos lo siguiente

$$H_{lj} = (-C_{lj}(A_{ij} - KB_{ij})^{-1}B_{ij} + D_{lj}(K(A_{ij} - KB_{ij})^{-1}B_{ij} + \delta_{jj})),$$

sustituyendo esta última expresión en (1.25) y despejando,

$$k_r = (H_{lj})^{-1}\delta_l^j - d_i r^i = (H_{ii})^{-1} - d_i r^i. \quad (1.26)$$

Por lo que el problema se reduce a encontrar una solución de la ecuación anterior, suponiendo que $\delta_i^j = r_i r^j$. Cuando $d_j = 0$ nuestro sistema se simplifica aún más,

$$k_r = (H_{ii})^{-1}. \quad (1.27)$$

Recordar que para que este procedimiento este bien definido es indispensable que los inversos de las matrices existan. Además, observamos que en realidad este valor solo depende de la traza, la cual es una cantidad invariante ante cambios de coordenadas. Con la notación y el desarrollo que hemos expuesto aquí, fácilmente se puede generalizar si suponemos que ahora el coeficiente k_r es una matriz.

Existe otra forma de determinar el valor de u , lo que permitiría construir un módulo de control diferente al que hemos presentado aquí, para ello se requiere que se minimice una función de coste [13], esta función nos dice como mantener al estado de nuestro sistema en algún punto (o al menos muy cercano), esta función está dada por

$$P[u(x, t)] = \Phi_1(x, u, t_f) + \int_0^t \Phi_2(x, u, t') dt', \quad (1.28)$$

ya que esta expresión es dada en forma integral es fácil insertar las restricciones que el sistema podría tener, existen diversas formas de la función que esta dentro del integrando y de la otra que esta fuera, y estos dependen del sistema, puede ver algunos ejemplos en [13], se pueden aplicar diversas técnicas para resolver esta ecuación, entre las más usadas se encuentra el método variacional, una de las funciones de coste más conocidas y utilizadas, es el regulador lineal cuadrático [14],

$$P[u] = \int_0^{t_f} \frac{1}{2}(x^i Q_{ij} x^j + u^l R_{ls} u^s) d\tau + \frac{1}{2} x(t_f)^i Q_{ij}(t_f) x^j(t_f), \quad (1.29)$$

(t_f es el tiempo en el que se esta la dinámica del sistema) de hecho, se puede encontrar que en un límite, la solución a este problema es $u_i = -K_{ij} x_j$, para más detalles ver [14], la solución que se encuentra para este problema de optimización es muy similar a la elección que nosotros usamos para u en (1.23).

1.6. Funciones de transferencia

En esta sección se presenta el concepto de función de transferencia, que es una descripción compacta de la relación entrada/salida para un sistema lineal. La combinación de funciones de transferencia con diagramas de bloques proporciona un método poderoso para tratar con sistemas lineales complejos. También se analiza la relación entre las funciones de transferencia y otras descripciones de la dinámica del sistema.

El modelado de un sistema a través de su respuesta a señales sinusoidales y exponenciales se conoce como modelado en el dominio de la frecuencia. Esta terminología se deriva del hecho de

que representamos la dinámica del sistema en términos de la frecuencia generalizada s en lugar de la variable del dominio del tiempo t . La función de transferencia proporciona una representación completa de un sistema lineal en el dominio de la frecuencia.

Este tipo de representaciones nos ayudarán a estudiar y entender mejor como se analizan sistemas modulares y además nos permiten introducir diferentes tipos de controladores al sistema.

1.6.1. Derivación de la función de transferencia

Como ya hemos mencionado será necesario usar señales sinusoidales y exponenciales o alguna combinación lineal de estas (una función puede llegar a descomponerse en senos y cosenos usando la transformada de Fourier), por lo que haremos uso de una señal exponencial, la cual sería proporcional a e^{st} , donde $s = \sigma + i\omega$ un número complejo. Este tipo de señales aparece de manera recurrente al solucionar diversos problemas lineales, muchas soluciones pueden ser presentadas como combinación lineal de estas. Podemos expandir esta función

$$e^{(\sigma+i\omega)t} = e^{\sigma t}(\cos(\omega t) + i\sin(\omega t)),$$

diferentes valores para la σ y ω dan a diferentes tipos de señales. Derivaremos nuestra función de transferencia para el caso continuo, para ello recordamos que nuestro sistema esta descrito por el siguiente conjunto de ecuaciones

$$\begin{aligned} \frac{dx_i}{dt} &= A_{ij}x_j + B_{ij}u_j, \\ y_l &= C_{lj}x_j + D_{lj}u_j. \end{aligned} \tag{1.30}$$

La solución a este sistema de ecuaciones (si A_{ij} , B_{ij} , C_{ij} y D_{ij} no depende del tiempo) será (simplificando un poco la notación quitando los índices, pero recordar que aún tenemos vectores y matrices),

$$x(t) = x(0)e^{At} + e^{At} \int_0^t e^{-A\tau} \mathbf{B}u d\tau, \tag{1.31}$$

si consideramos que las señales de entrada están dadas por $u_i = z_i e^{st}$, por lo que ahora nuestra solución será

$$x(t) = x(0)e^{At} + e^{At} \int_0^t e^{(sI-A)\tau} \mathbf{B}z d\tau, \tag{1.32}$$

es posible que nuestra solución se simplifique aún más si $sI - A = 0$, es decir que s sea algún eigenvalor de la matriz A , pero este caso no nos interesa, por lo que supondremos que s no es ningún eigenvalor de A , por lo que ahora la solución de (1.32) es

$$\begin{aligned} x(t) &= x(0)e^{At} + e^{At}(sI - A)^{-1}(e^{(sI-A)t} - I)Bz \\ &= e^{At}(x(0) - (sI - A)^{-1}Bz) + (sI - A)^{-1}Bze^{st}, \end{aligned} \tag{1.33}$$

por lo que ahora la salida en nuestro sistema será

$$\begin{aligned} y(t) &= Cx(t) + Du(t) \\ &= Ce^{At}(x(0) - (sI - A)^{-1}Bz) + (C(sI - A)^{-1}Bz + Dz)e^{st}, \end{aligned} \tag{1.34}$$

el primer término se debe completamente a la respuesta del sistema, mientras que el segundo se debe a la función de entrada, dependiendo de los eigenvalores de A nos diría si el sistema es estable o no, para deshacernos de esta incertidumbre uno siempre puede elegir al estado inicial

como $x(0) = (sI - A)^{-1}Bz$, por lo que ahora el sistema solo depende de una cantidad proporcional al impulso de entrada, entonces tendríamos lo siguiente

$$\begin{aligned} x(t) &= (sI - A)^{-1}Bze^{st} = (sI - A)^{-1}Bu(t), \\ y(t) &= (C(sI - A)^{-1}Bz + Dz)e^{st} = (C(sI - A)^{-1}B + D)u(t), \end{aligned} \quad (1.35)$$

con este resultado uno observa que conociendo el estímulo de entrada (que es en forma exponencial o sinusoidal) uno puede llegar a conocer la salida a través de una función de transferencia

$$G_{yu}(s) = C(sI - A)^{-1}B + D, \quad (1.36)$$

entonces para encontrar la salida solo tenemos que computar de la manera siguiente

$$y(t) = G_{yu}(s)u(t), \quad (1.37)$$

una propiedad importante de la función de transferencia es que esta es invariante bajo cambios de coordenada de nuestro sistema, además se puede obtener la función de transferencia usando la transformada de Laplace [6], para ello se aplica esta transformada a (1.30) y se considera una condición inicial $x(0) = 0$, con las mismas restricciones que hemos considerado se llega a una relación similar,

$$Y(s) = G_{yu}(s)U(s). \quad (1.38)$$

1.6.2. Diagramas de bloque y funciones de transferencia

A continuación, analizaremos y explicaremos como usar y representar las funciones de transferencia en los bloques, nos centraremos en explicar la estructura de los que se encuentran la figura 1.7.

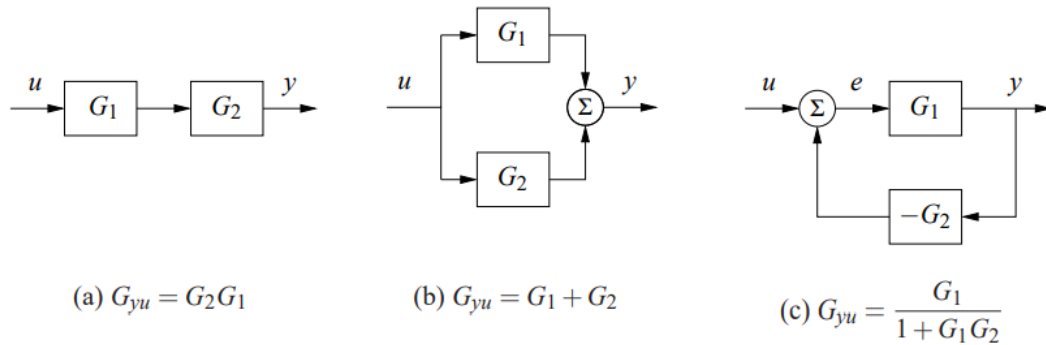


Figura 1.7: Bloques y funciones de transferencia [6].

Para el caso de la figura 1.7(a), el que aparezca un cuadro significa que amplifica la entrada un número veces igual al que está representado dentro de esta, por lo que tendremos

$$y = G_2G_1u, \quad (1.39)$$

es decir, nuestra nueva función de transferencia será $G_{yu} = G_2G_1$. En el caso de la figura 1.7(b), recordar que el símbolo sumatorio significa la suma lineal de todas las entradas por lo que tendremos

$$y = (G_2 + G_1)u, \quad (1.40)$$

con lo que nuestra función de transferencia será $G_{yu} = G_2 + G_1$. Para la figura 1.7(c), tendremos las siguientes dos relaciones

$$\begin{aligned} e &= u - G_2 y, \\ y &= G_1 e, \end{aligned}$$

sustituyendo una dentro de la otra y despejando se encuentra

$$y = \frac{G_1}{1 + G_1 G_2} u, \tag{1.41}$$

por lo que nuestra función de transferencia quedara $G_{yu} = \frac{G_1}{1 + G_1 G_2}$. En base a las ideas presentadas aquí uno puede construir algún sistema compuesto de muchos módulos y encontrar su respectiva función de transferencia. No olvidar que el formalismo que desarrollamos en esta sección tiene tres grandes restricciones

1. Se cumple para puntos cerca del equilibrio.
2. Es indispensable que la entrada tenga una estructura exponencial o una combinación lineal de estas.
3. La condición inicial que encontramos es diferente según el método que se use.

A pesar de estas restricciones existe un gran número de sistemas que las cumplen o se pueden forzar a cumplirlas.

1.7. Control PID

En esta sección se tratarán las propiedades básicas del control proporcional-integral-derivativo (PID). El control PID, es la forma más común de control por retroalimentación. Aparece en dispositivos simples y hasta en grandes fábricas con miles de controladores. La mayoría de los controladores PID no utilizan acciones derivadas, por lo que, estrictamente hablando, deberían llamarse controladores PI. También existe una creciente evidencia de que el control PID aparece en sistemas biológicos.

1.7.1. Funciones básicas de control

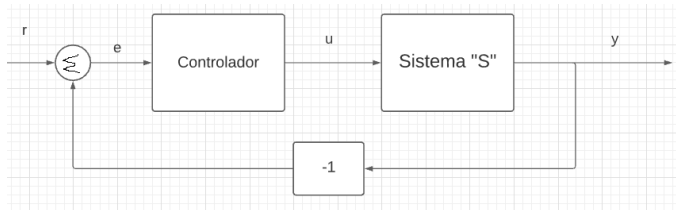


Figura 1.8: Módulo de control.

Para plantear un sistema con retroalimentación PID, primero pensemos en un sistema parecido el de la figura 1.8, la r es un parámetro de referencia estos pueden llegar a ser los datos que se obtienen de manera experimental, o simplemente algún valor en el cual queremos mantener nuestro sistema, ahora el objetivo es construir el módulo del controlador de tal manera que y sea lo más cercano a r , la diferencia entre estos valores será caracterizado por una función error definida como $e = r - y$, ahora analizaremos con más detalle el controlador PID, para ello se intenta eliminar el error cuando es independiente o dependiente del tiempo.

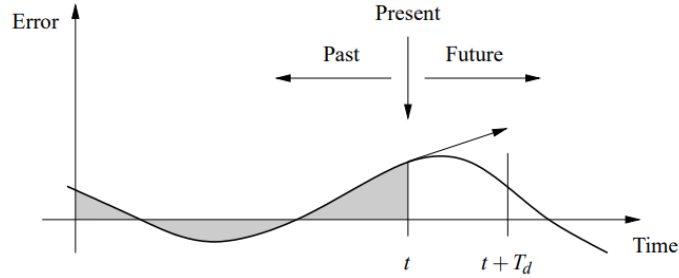


Figura 1.9: Control PID [6].

1. El primer caso es cuando el controlador tiene una constante k_p (la constante k_p se puede pensar como el valor de la ganancia del controlador), es decir que es lineal la supresión del error, y está dada por la siguiente formula

$$u = k_p e.$$

Esta fórmula resulta ser bastante útil cuando el error es independiente del tiempo, o el error en un determinado tiempo, como en la figura 1.9, el error en el presente.

2. Una de las maneras para analizar los sistemas en el que el error depende del tiempo es a través del control integral, piense en la figura 1.9, imagine que el error depende del tiempo y se ha ido acumulando desde un tiempo inicial $t_0 = 0$ hasta el tiempo actual t , por lo que ahora para saber cuánto error se ha ido acumulando basta con integrar esta cantidad, ahora u será una constante por ese error acumulado,

$$u = \frac{k_p}{T_i} \int_0^t e(t') dt'.$$

La constante T_i está estrechamente relacionada con la velocidad con la que se repite el promedio.

3. Cuando sabemos que el error depende del tiempo, posiblemente conocemos la expresión o algún conjunto de valores hasta un tiempo t , queremos prever el error que va a ocurrir, como en la figura 1.9, para ello realizamos una expansión a primer orden de la función error, por lo que para quitar este error futuro debemos de insertar una función de entrada similar a

$$u = k_p T_d \frac{de(t)}{dt}.$$

Esta constante T_d está relacionada con el tiempo de la acción predicativa.

Juntando las tres relaciones; el error acumulado, el actual y el futuro construimos nuestro controlador PID

$$u(t) = k_p e(t) + k_i \int_0^t e(t') dt' + k_d \frac{de(t)}{dt} = k_p \left(e(t) + \frac{1}{T_i} \int_0^t e(t') dt' + T_d \frac{de(t)}{dt} \right). \quad (1.42)$$

Para tener una ecuación similar a la función de transferencia, aplicaremos la transformada de Laplace a esta última relación (como hemos mencionado se puede obtener la función de transferencia con la transformada de Laplace), obteniendo

$$U(s) = k_p \left(1 + \frac{1}{T_i s} + T_d s \right) E(s) = G_{ue}(s) E(s) = C(S) E(s). \quad (1.43)$$

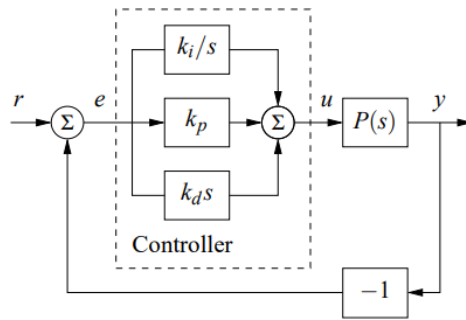


Figura 1.10: Diagramas de bloque de control PDI [6].

Una buena representación de un sistema con estas características estaría dada por la figura 1.10, ahí podemos observar cómo se va corrigiendo el error a través del controlador.

Para conocer la función de transferencia de todo nuestro sistema gobernado por una dinámica $P(s)$, decimos que nuestro controlador es $C(s)$, por lo que la función de transferencia será

$$G_{yr} = \frac{PC}{1 + PC} = \frac{Pk_p \left(1 + \frac{1}{T_i s} + T_d s\right)}{1 + Pk_p \left(1 + \frac{1}{T_i s} + T_d s\right)}. \quad (1.44)$$

En estas ecuaciones podemos eliminar las contribuciones integral y derivativa eligiendo algunas de las variables de manera adecuada. Podemos ver como son los efectos de este módulo de control en un sistema que esta caracterizado por la dinámica $P(s) = \frac{1}{(s+1)^3}$ como en la figura 1.11, en la figura 1.11(a) solo tenemos al controlador proporcional, en la siguiente se agrega el integral, en el último el derivativo, la flecha señala la dirección en la que crece las constantes. Podemos observar que dependiendo de valor de las constantes nuestro sistema oscila hasta un determinado tiempo y luego se mantiene estable. Existen dos maneras de determinar las constantes que aparecen, una de ellas es realizando simulaciones hasta encontrar algún parámetro óptimo, el otro es a través del método de Zieger-Nichols [6].

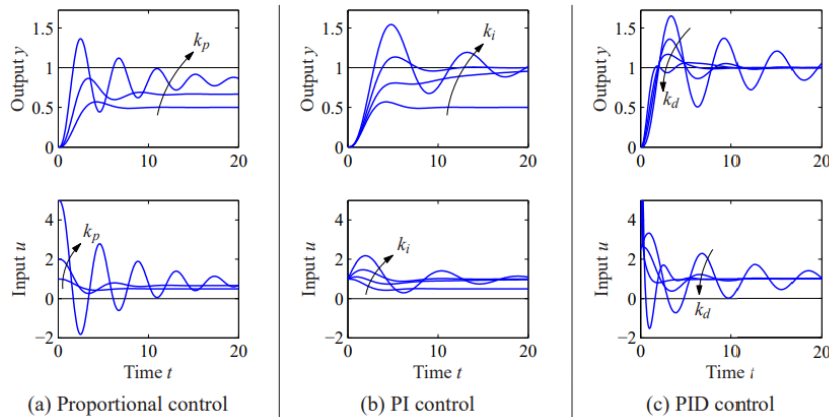


Figura 1.11: Respuesta de control PDI [6].

Capítulo 2

Biología sintética

A continuación, daremos algunos conceptos y métodos desarrollados en biología sintética (SynBio), especialmente en circuitos genéticos, se hará un estudio de sistemas biológicos desde una perspectiva de sistemas con retroalimentación (enfoque determinista). Nos enfocaremos principalmente en especial en el tema de redes de regulación genética; como se construyen y que ecuaciones gobiernan su dinámica, y por otro lado se describe un caso en el que los efectos de la interconexión entre módulos "retroactividad" son débiles, esta cantidad se cuantifica a través de una función, esto último se aplica al proceso de fosforilación (que está muy relacionado con las redes de regulación genética), la mayoría de los desarrollos, ideas e imágenes fueron tomados directamente de [8] y [9]. Todos los análisis que se realizan en este capítulo serán deterministas (principalmente describiremos a nuestros sistemas con ecuaciones diferenciales), hacemos todos estos análisis con el objetivo de realizar una contra parte estocástica en los últimos capítulos.

2.1. ¿Qué es la biología sintética?

Para empezar, nos preguntamos qué es un sistema biológico, esta definición es complicada de dar, aunque los sistemas biológicos entran en la categoría de sistemas complejos (ver capítulo 1), sin embargo, podemos quedarnos con alguna de las dos siguientes definiciones de sistema biológico:

- i) Una red compleja de interacciones entre entes biológicos, como biomoléculas, células, tejidos, organismos, poblaciones y comunidades [2].
- ii) El conjunto de partes similares que trabajan en armonía para cumplir alguna función fisiológica determinada. La interacción de sistemas biológicos da lugar a una red de sistemas que, a su vez, pueden organizarse para dar lugar a procesos de mayor complejidad [3].

Con esto ahora definimos a la biología sintética como:

Una disciplina que sirve para diseñar o rediseñar sistemas biológicos y otorgarles cualidades mejoradas o nuevas cualidades [3].

Dentro de los acontecimientos científicos más relevantes que le han dado a la SynBio una presencia central en esta revolución bio-industrial se encuentran los siguientes: el descubrimiento y los avances logrados en la regulación génica (e.g. el operón LacZ); el perfeccionamiento de las técnicas de ADN recombinante y la ingeniería genética; el lanzamiento del proyecto genoma; la secuenciación Sanger; el perfeccionamiento de la técnica de la Reacción en Cadena de la Polimerasa (PCR por sus siglas en inglés); la generación de bases de datos mundiales tales como NCBI, GeneBank y EMBL-net; la optimización de nuevas tecnologías de Secuenciación de Nueva Generación (NGS's) para la decodificación del ADN; el ensamble Gibson; el uso de la herramienta de edición de genomas

CRISPR-Cas9, los modelos computacionales de célula completa; así como la reciente creación de la bacteria sintética *Micoplasma mycoides* JCVI-syn 3.0 [3].

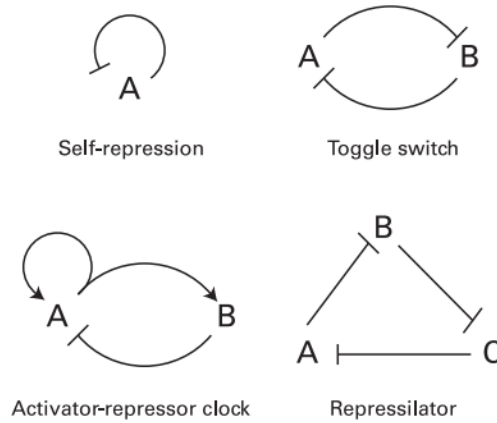


Figura 2.1: Circuitos de genéticos [8].

Una de las direcciones en la que se está desarrollando la SynBio es en el diseño de circuitos genéticos, algunos ejemplos de esto se pueden ver en la figura 2.1, más adelante hablaremos un poco sobre los ejemplos, por lo que podemos decir que la SynBio puede estar avanzando hacia un desarrollo similar al de diseño de circuitos integrados y modulares, muy similar a lo que ocurre en los sistemas electrónicos tales como las computadoras.

2.2. Módulo de transcripción/traducción

En lo que sigue del texto usaremos mucho a los módulos de transcripción/traducción, entonces explicaremos brevemente que procesos bioquímicos ocurren, así como la aparición de la función de Hill.

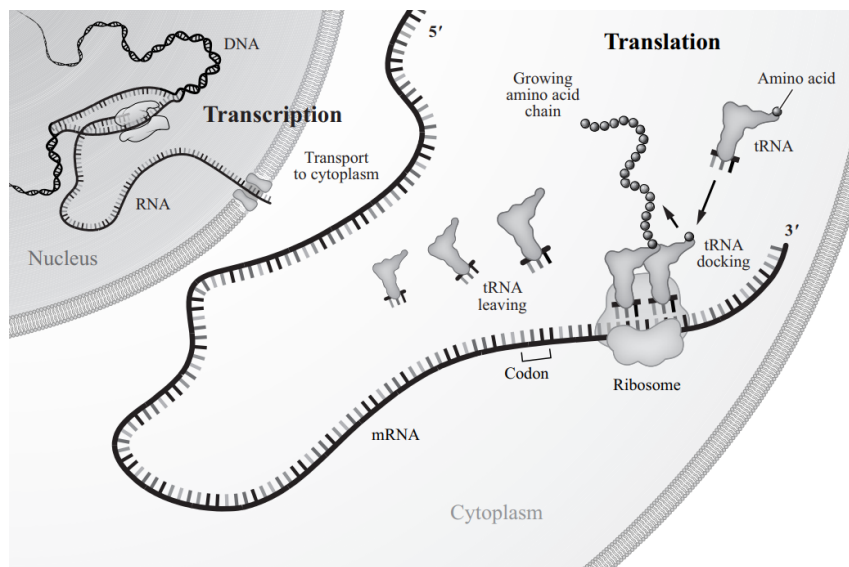


Figura 2.2: Proceso de transcripción y traducción [15].

Transcripción

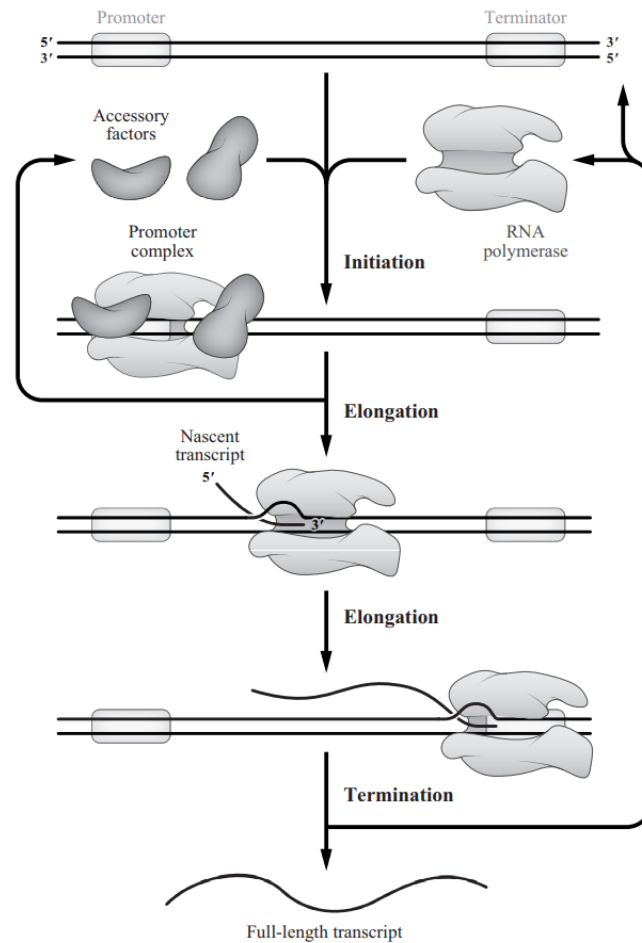


Figura 2.3: Proceso de transcripción [15].

El primer proceso que ocurre es la transcripción, por la figura 2.3, podemos notar que lo que esencialmente hace este proceso es formar un ARNm (ARN mensajero) a partir del ADN. La RNA polimerasa, junto con otros factores, se une a la región promotora del ADN y luego “abre” el ADN para comenzar la transcripción (iniciación). A medida que la RNA polimerasa avanza por el ADN en el complejo de elongación de la transcripción (TEC), produce una transcripción de ARN (elongación), que luego se traduce en una proteína. El proceso termina cuando la RNA polimerasa alcanza el terminador (terminación) [15].

Traducción

Nos apoyaremos de la figura 2.2 para explicar un poco cómo están conectados los procesos de transcripción y traducción, el primer proceso es la transcripción, ahora lo que sigue es explicar un poco el segundo proceso. La traducción es el proceso de traducir la secuencia de una molécula de ARNm (ARN mensajero) a una secuencia de aminoácidos durante la síntesis de proteínas. El código genético describe la relación entre la secuencia de pares de bases en un gen y la correspondiente secuencia de aminoácidos que codifica. En el citoplasma celular, el ribosoma lee la secuencia del ARNm en grupos de tres bases para ensamblar la proteína [15].

Regulación transcripcional

Lo que mencionamos en el proceso de transcripción no considera que la producción del ARm este regulado por otras proteínas, las formas más simples de esto es la activación y la inhibición, comúnmente llamados factores de transcripción. Podemos pensar que una proteína activa o inhibe la actividad de la promotora, además hay una gran escala de tiempo en lo que suceden la actividad promotora y la transcripción/traducción, el más corto es el efecto de la región de la promotora, la cual dura en el orden de segundos, mientras que el de la transcripción/traducción puede llegar a durar de minutos a horas. Entonces, uno puede pensar que el efecto de la promotora se encuentra prácticamente en equilibrio. Por lo que se puede capturar la actividad de la promotora con alguna función.

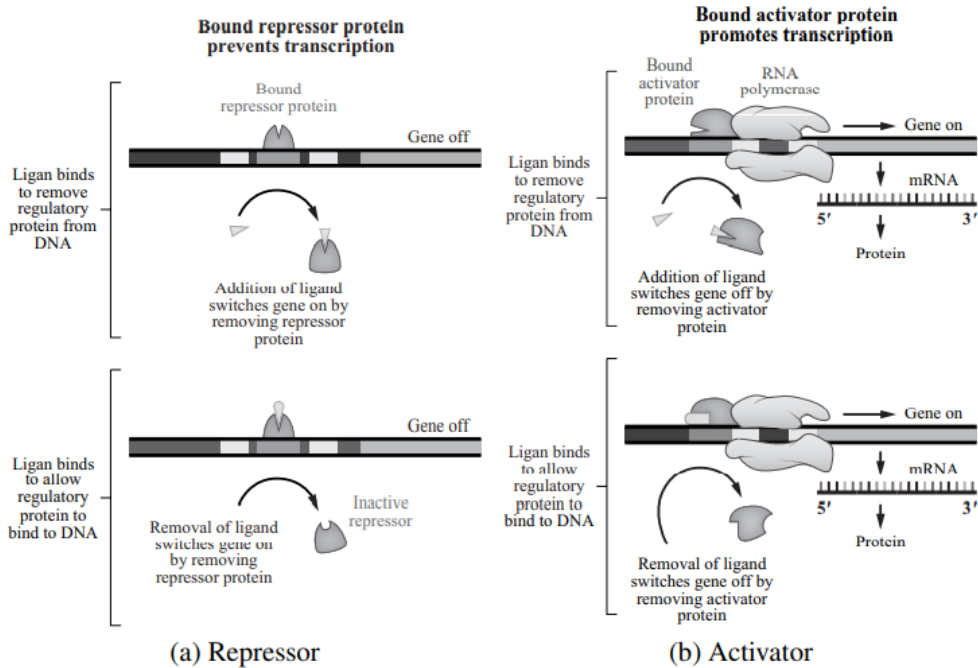


Figura 2.4: Regulación genética [15].

De hecho, algunas proteínas pueden actuar como factores de transcripción en el MTT, esto quiere decir que las proteínas pueden actuar como activadora; al acelerar la actividad de la promotora, de tal manera que aumenta la tasa de transcripción del gen. También como supresor; cuando disminuye la tasa de transcripción del gen. Suponiendo que estos procesos ocurren muy rápido y con las ideas que están detrás de la función de Hill, podemos sustituir la contribución de la promotora por la función de Hill, para ser más precisos la función de Hill es proporcional al número de promotoras activas (más adelante veremos a detalle esta función).

Para el caso del activador supondremos que la función de Hill toma esta forma

$$H = \frac{\beta p^n}{K^n + p^n}. \tag{2.1}$$

A K se le conoce como el parámetro de activación y en muchos sistemas esta es del orden 10^{-4} , β es la máxima actividad promotora, n determina la inclinación de la función de entrada, mientras más grande sea n , la función de Hill se parecerá más a una función escalón. Cuando tenemos un

supresor la función de Hill toma la forma siguiente:

$$D = \frac{\beta K^n}{K^n + p^n}. \quad (2.2)$$

Los parámetros tienen el mismo significado que en el activador, sin embargo, muchos genes tienen un nivel de expresión mínimo distinto de cero, llamado nivel de expresión basal del gen. Un nivel basal se puede describir agregando a la función de entrada un término β_0 [4].

Supuesto de modularidad

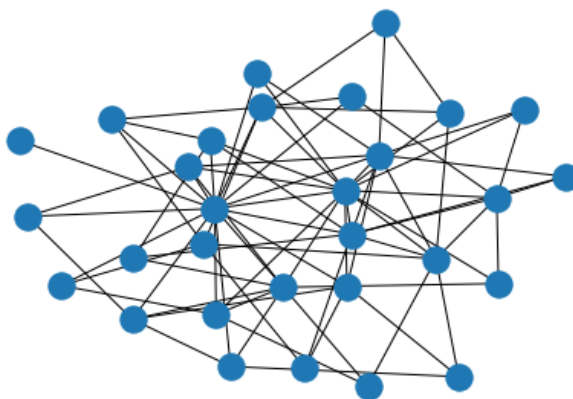


Figura 2.5: Grafo.

Como hemos venido mencionando al principio de esta sección, que una de las vertientes de la SynBio es en el desarrollo de circuitos genéticos, para ello es indispensable pensar que se pueden encontrar y/o diseñar sistemas biológicos modulares, es decir sistemas que se pueden entender como si estos estuviesen formados por módulos y cada uno de ellos cumplen alguna determinada función, esto último va de la mano con los conceptos de retroalimentación y control que están al principio de este escrito.

Nosotros nos interesaremos principalmente en el estudio de circuitos de transcripción, podemos pensar en estos como una computadora o aún mejor en un grafo (ver figura 2.5), donde cada nodo de este grafo representaría a un módulo de transcripción/traducción (MTT) de algún gen como en la figura 2.6, si pensamos que uno de estos nodos pertenece a un gen \mathcal{Y} , entonces este puede modelarse como un módulo de entrada-salida que toma como entrada las concentraciones de factores de transcripción que regulan el gen \mathcal{Y} y da como salida la concentración de proteína expresada por el gen \mathcal{Y} , denominada Y .

2.3. Dinámica de MTT con regulación

En concordancia con los desarrollos que hemos hecho en retroalimentación, pensaremos en la figura 2.6 como un sistema con retroalimentación, esta figura nos explica una versión simplificada del proceso de transcripción, aquí la dinámica interna del MTT está determinada por los procesos de transcripción y traducción, que son mucho más lentos que la dinámica de la unión del factor de transcripción. La unión de los factores de transcripción al sitio del promotor alcanza el equilibrio

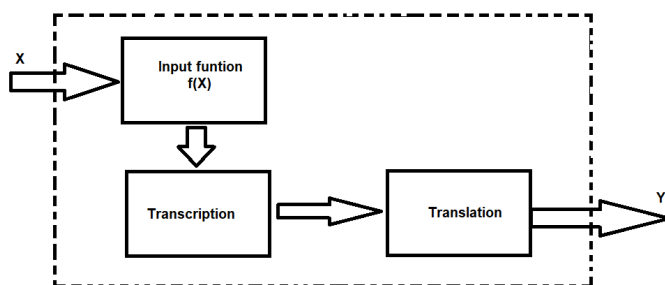


Figura 2.6: Módulo de transcripción/traducción.

en segundos, mientras que la transcripción y traducción del gen le toma de minutos a horas. Esta separación de escala de tiempo, una característica clave de los circuitos de transcripción, conduce a la siguiente simplificación del modelo; de acuerdo con el supuesto de modularidad, la dinámica de la unión del factor de transcripción/ADN se considera en equilibrio, y cada concentración de factor de transcripción ingresa al módulo de transcripción de entrada-salida a través de una función de entrada estática que impulsa la dinámica del sistema.

En el caso más simple de una función de entrada que actúa como represor o activador toman la forma de la función Hill (para más detalles ver el capítulo 6). Al analizar la dinámica interna del MTT generalmente se modela la dinámica de ARNm y proteínas a través de los procesos de transcripción y traducción. La producción de proteínas se equilibra tanto por la descomposición, a través de la degradación, que ocurre cuando la proteína es destruida por proteínas especializadas en la célula, por ejemplo, reconocen una parte específica de la proteína y la destruyen, así como por la dilución, que es causada por la reducción en la concentración de la proteína debido al aumento del volumen celular durante el crecimiento. De manera similar, la producción de ARNm también se equilibra mediante degradación y dilución. Por lo tanto, la dinámica de un módulo de transcripción a menudo está bien capturada por las siguientes ecuaciones diferenciales ordinarias, que pueden obtenerse a partir de la ley de acción de masas que presentamos en una de las secciones anteriores

$$\frac{dm_y}{dt} = f(X) - \alpha_{my}m_y, \quad \frac{dY}{dt} = \delta_{my}m_y - \alpha_Y Y. \quad (2.3)$$

Denotaremos por x la proteína, por X la concentración media de proteína y por \mathcal{X} el gen que expresa la proteína x , m_y denota la concentración del ARNm de traducción del gen \mathcal{Y} , mientras que Y la concentración del gen \mathcal{Y} , $f(X)$ es la función de Hill y actuará como represor o activador según sea el caso, las constantes α nos hablan sobre la razón de degradación y/o disolución del gen así como del ARNm, mientras que la δ sobre la razón de traducción del ARNm.

2.4. Diseño de circuitos genéticos

Con el supuesto de modularidad ahora podemos diseñar y/o analizar algunos circuitos genéticos, en particular algunos de los que se encuentran en la figura 2.7. Con este procedimiento, se busca predecir el comportamiento de un circuito a partir del de sus componentes, una vez que estas han sido bien caracterizadas de forma aislada. Se ha supuesto que una red de regulación genética esta formada por la interacción (regulación) de varios módulos. Existe un gran interés en diseñar circuitos genéticos con oscilaciones, ya que a partir de estos es posible regular otros procesos, donde estas oscilaciones actuarían como una especie de reloj, tal como sucede en el cuerpo humano, como el ciclo circadiano, el diseño de estos circuitos también está dando pie a la computación biológica [4].

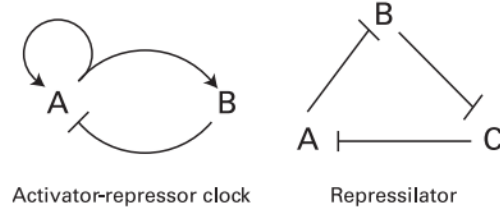


Figura 2.7: Circuitos de genéticos [8].

2.4.1. El Represilador

El primer sistema que analizaremos será el Represilador, aparece en la figura 2.7, aquí tenemos tres genes donde todos se suprimen en la dirección $A \rightarrow B \rightarrow C \rightarrow A$, pensemos que el proceso inicia en el gen A (puede empezar desde cualquier gen ya que no hay causalidad en los sistemas con retroalimentación cerrada) que va suprimiendo al gen B , por lo que aumenta la producción del gen C , al aumentar este gen suprime al gen A , por lo que ahora aumentaría el gen B , se seguiría un proceso de este tipo de tal manera que hay un aumento y disminución cíclico de los genes, este sistema fue uno de los primeros circuitos genéticos que se diseñaron, para esto se usó la bacteria *E. coli* [5].

Para conocer la dinámica de nuestro sistema suponemos que cada uno de los genes son un MTT, por lo que su dinámica estaría descrita por las ecuaciones modulares (2.3), de esta manera nuestro sistema tendrá una dinámica dada por las siguientes ecuaciones diferenciales

$$\begin{aligned}
 \frac{dm_A}{dt} &= f_1(C) - \alpha_{m_A} m_A, & \frac{dA}{dt} &= \delta_{m_A} m_A - \alpha_A A, \\
 \frac{dm_B}{dt} &= f_2(A) - \alpha_{m_B} m_B, & \frac{dB}{dt} &= \delta_{m_B} m_B - \alpha_B B, \\
 \frac{dm_C}{dt} &= f_3(B) - \alpha_{m_C} m_C, & \frac{dC}{dt} &= \delta_{m_C} m_C - \alpha_C C.
 \end{aligned} \tag{2.4}$$

Nos enfocaremos en dos casos para analizar este sistema, el primero es el caso simétrico, es cuando todos los genes actúan como supresores, el segundo caso es cuando uno de los genes es un activador, cuando estamos en el caso simétrico todas las funciones de Hill son represores, es decir

$$f_1(p) = f_2(p) = f_3(p) = \frac{\gamma_{1p}}{1 + \gamma_{2p} p^n}, \tag{2.5}$$

para el Represilador que se diseñó en [5], se usaron el siguiente conjunto de parámetros $\delta_{mp} = \gamma_{2p} = 1$, $\alpha_p = \alpha_{mp} = \alpha$ y $\gamma_{1p} = \alpha^2$, con el valor de estos parámetros se calculara ahora la sección donde existirían oscilaciones en nuestro sistema. Para encontrar esta región nos fijamos en un método similar al desarrollado en [5], de donde se obtiene que el sistema tiene soluciones periódicas si cumple la siguiente condición

$$\frac{\alpha^2}{\delta^2} > \sqrt[n]{\frac{\frac{4}{3}}{n - \frac{4}{3}}} \left(1 + \frac{\frac{4}{3}}{n - \frac{4}{3}} \right) \tag{2.6}$$

Esta relación se ve en la figura 2.8(a), podemos ver que cuando n incrementa las soluciones periódicas del sistema están prácticamente garantizadas, de igual manera cuando el parámetro δ aumenta se nota que se van perdiendo las oscilaciones, de hecho, este parámetro está estrechamente relacionado con el periodo de las oscilaciones.

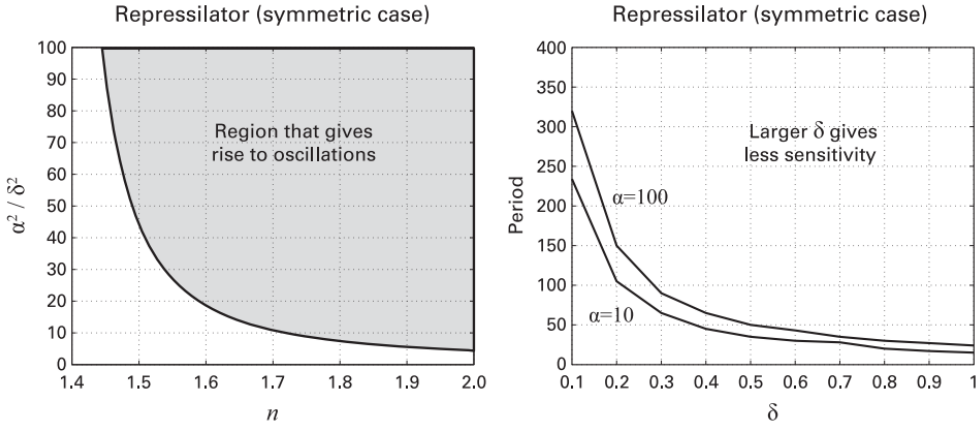


Figura 2.8: Regiones de oscilación del represilador [8].

El segundo caso es cuando uno de los genes es un activador, es decir nuestras funciones de Hill serán:

$$f_1(p) = f_2(p) = \frac{\gamma_{1p}}{1 + \gamma_{2p}p^n}, \quad f_3(p) = \frac{\gamma_{3p}p^n}{\gamma_{4p} + p^n}. \quad (2.7)$$

Para que nuestro modelo sea similar al Represilador que se diseñó en [5], se usarán el siguiente conjunto de parámetros $\delta_{mp} = \gamma_{2p} = \gamma_{4p} = 1$, $\alpha_p = \alpha_{mp} = \alpha$ y $\gamma_{1p} = \gamma_{3p} = \gamma_1^2$, con el valor de estos parámetros se calculará ahora la sección donde existirían oscilaciones en nuestro sistema. Para ello se emplea métodos similares al caso simétrico. Con los resultados que se obtuvieron tanto para el caso simétrico como para el caso antisimétrico es posible diseñar un circuito genético que contenga oscilaciones, para ello lo único que se tendría que hacer es ajustar los parámetros, en la figura 2.8 están las gráficas que nos muestran las regiones de oscilación como función de n y δ .

2.4.2. El reloj activador-represor

El otro sistema que analizaremos será el reloj activador-represor, tal como aparece en la figura 2.7, en esta imagen se ve como el gen \mathcal{A} suprime al gen \mathcal{B} y como también este se activa a sí mismo, mientras que el gen \mathcal{B} solo suprime al gen \mathcal{A} , en este sistema también se podrían encontrar oscilaciones en las densidades de los genes. Para analizar este sistema suponemos que se cumple el supuesto de modularidad, las ecuaciones que describen la dinámica de nuestro sistema estarán dadas por las ecuaciones para el MTT,

$$\begin{aligned} \frac{dm_A}{dt} &= f_1(A, B) - \alpha_{mA}m_A, & \frac{dA}{dt} &= \delta_{mA}m_A - \alpha_A A, \\ \frac{dm_B}{dt} &= f_2(A) - \alpha_{mB}m_B, & \frac{dB}{dt} &= \delta_{mB}m_B - \alpha_B B, \end{aligned} \quad (2.8)$$

las funciones de entrada son las siguientes:

$$f_1(A, B) = \frac{\gamma_{1A}A^n + \gamma_{0A}}{1 + \gamma_{2A}A^n + \gamma_{3A}B^n}, \quad f_2(A) = \frac{\gamma_{1B}A^n + \gamma_{0B}}{1 + \gamma_{2B}A^n}. \quad (2.9)$$

Ya hemos visto el significado de algunas de estos parámetros, ahora lo único que faltaría explicar es lo que significan los parámetros γ_{0p} el cual está relacionado con la transcripción basal. Para simplificar el modelo se elige el siguiente conjunto de parámetros $\alpha_{mp} = \frac{\delta_p}{\epsilon}$, $\delta_{mp} = \frac{k_p}{\epsilon}$, $\alpha_A = \delta_A \xi$, el parámetro ϵ y ξ están estrechamente relacionados con la escala de tiempos entre la producción

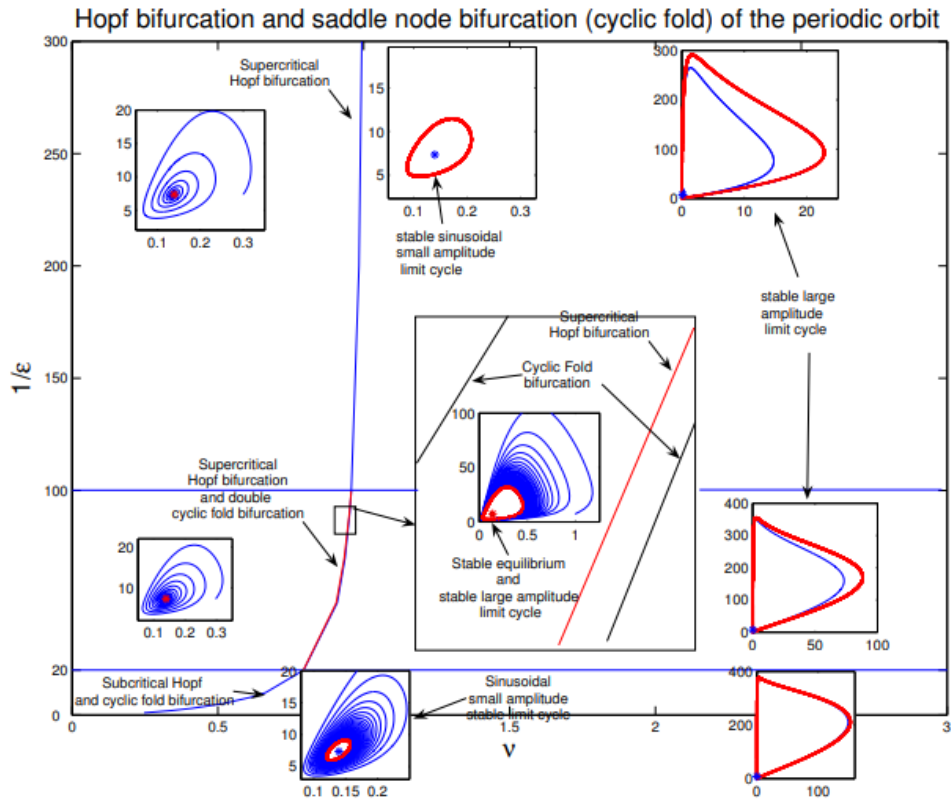


Figura 2.9: Regiones de oscilación del reloj activador-represor [16].

de la proteína y ARNm, y la diferencia de escalas entre el activador y el represor. Podemos ver un análisis de este sistema en términos de estas dos variables, tal como se ve en la figura 2.9, esta figura como los detalles del análisis se pueden llegar a ver en [16].

En términos de los parámetros ϵ , ξ y n , es posible diseñar un sistema con un comportamiento oscilatorio: a medida que la escala de tiempo de la dinámica del activador aumenta con respecto a la dinámica del represor, los parámetros del sistema se mueven a través de una bifurcación de Hopf y las oscilaciones estables se moverán. Desde el punto de vista de la fabricación, esto se puede lograr añadiendo etiquetas de degradación adecuadas a la proteína activadora. La región del espacio de parámetros en la que el sistema exhibe oscilaciones amortiguadas casi sinusoidales está en el lado izquierdo de la curva correspondiente a la bifurcación de Hopf. Si esto fuera el caso, el aumento de la separación de escalas de tiempo entre el activador y el represor, ξ , puede conducir a un ciclo límite estable.

2.5. Mas allá de la asunción de modularidad

Hasta ahora hemos diseñado circuitos genéticos basados en el supuesto de modularidad, por ejemplo, la monotonicidad de las funciones de entrada en los módulos transcripcionales que componen el represor fue clave para mostrar formalmente la existencia de soluciones periódicas, de igual manera en el diseño del reloj activador-represor permitió predecir fácilmente la ubicación y el número de puntos de equilibrio a medida que se cambiaban los parámetros. El supuesto de modularidad implica que, cuando dos módulos están conectados entre sí, su comportamiento no

cambia en la interconexión. Sin embargo, un problema fundamental de ingeniería de sistemas surge cuando se interconectan subsistemas; cuando el proceso de transmisión de una señal a un componente descendente afecta el estado dinámico del componente emisor.

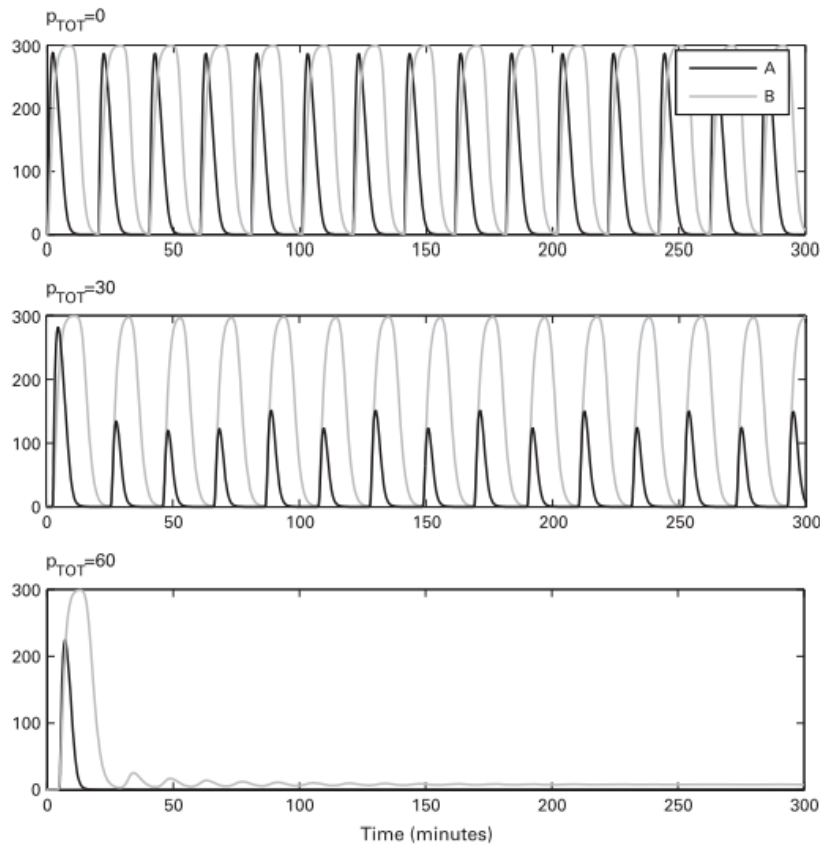


Figura 2.10: Perdida del reloj activador-represor [16].

De hecho, después de diseñar, probar y caracterizar el comportamiento de entrada-salida de un componente individual de forma aislada, es ciertamente deseable que sus características no cambien cuando otro componente está conectado a su canal de salida. Este problema, el efecto de las cargas en la salida de un sistema, se comprende bien en muchos campos de la ingeniería, por ejemplo, en el diseño de circuitos eléctricos. A menudo se ha señalado que surgen problemas similares para los sistemas biológicos.

Los módulos deben tener características especiales que les permitan integrarse fácilmente en cualquier sistema, como impedancia de salida cero e impedancia de entrada infinita. Una revisión extensa sobre problemas de cargas y modularidad en redes de señalización se puede encontrar en [8], donde los autores proponen analogías concretas con problemas similares que surgen en los circuitos eléctricos.

Estas cuestiones son especialmente delicadas en biología sintética. Por ejemplo, considere el reloj activador-represor de la figura 2.7. Supongamos que queremos emplear este reloj (sistema de flujo ascendente (UT)) para impulsar uno o más componentes (sistemas de flujo descendente (DT)), utilizando como señal de salida la concentración oscilante $A(t)$ del activador. Desde el punto de vista de los sistemas/señales, $A(t)$ se convierte en una entrada para el segundo sistema. Los términos UT

y DT reflejan la dirección en la que pensamos que las señales viajan, desde el reloj hasta los sistemas que se sincronizan. Sin embargo, esto es solo una idealización, porque la unión y desvinculación de A a los sitios promotores en un sistema DT compite con las interacciones bioquímicas que constituyen el bloque UT (retroactividad) y, por lo tanto, pueden interrumpir el funcionamiento del reloj en sí (figura 2.10). Un posible enfoque para evitar interrumpir el comportamiento del reloj, motivado por el enfoque utilizado en GFP (green fluorescent protein), es introducir un gen que codifique una nueva proteína X , colocada bajo el control del mismo promotor que el gen de A , y utilizar la concentración de X , que presumiblemente refleja la de A (concentración de la proteína), para impulsar el sistema DT. Este enfoque, sin embargo, todavía tiene el problema de que el comportamiento de la concentración de X en el tiempo puede ser alterado e incluso interrumpido por la adición de sistemas DT que drenan X . El resultado neto sigue siendo que los sistemas DT no están sincronizados adecuadamente.

2.5.1. Modelando la retroactividad

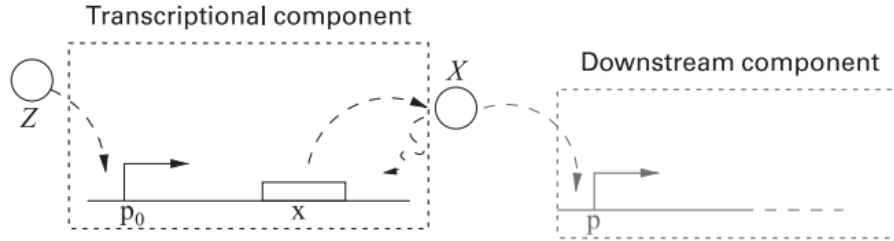


Figura 2.11: Sistema DT [9].

En términos generales, llamamos retroactividad al fenómeno por el cual el comportamiento de un sistema UT cambia al interconectarse con un sistema DT. Como ejemplo uno puede pensar en cómo un componente transcripcional es conectado a un sistema DT tal como se ve en la figura 2.11, el gen Z influye directamente en el componente transcripcional de la proteína x , con concentración X , para los propósitos que siguen no nos centraremos en el ARNm por lo que englobaremos todos estos efectos en una sola función $k(t)$, por lo que la concentración de la proteína x estará dada por la siguiente ecuación

$$\frac{dX(t)}{dt} = k(t) - \delta X(t), \quad (2.10)$$

donde δ es la tasa de degradación de la proteína, a esta ecuación se le conocerá como sin retroactividad o aislado. Cuando nuestro sistema se ha conectado a un módulo DT, de tal manera que la proteína X ahora es la que activa el módulo de transcripción al unirse a una promotora P figura 2.11, con concentración p ,

$$X + p \xrightleftharpoons[k_{off}]{k_{on}} C, \quad (2.11)$$

aquí C es la proteína-promotora (una molécula compuesta por la unión de la proteína X y la promotora p , C no se degrada ya que son sus componentes los que lo harían, en este caso X y/o p), ya que C no está sujeta a degradarse, entonces se conserva la concentración total $p_{Tot} = p + C$, por lo que ahora la dinámica del sistema de será:

$$\begin{aligned} \frac{dX(t)}{dt} &= k(t) - \delta X(t) + k_{off}C(t) - k_{on}(p_{Tot} - C(t))X(t), \\ \frac{dC(t)}{dt} &= -k_{off}C(t) + k_{on}(p_{Tot} - C(t))X(t), \end{aligned} \quad (2.12)$$

podemos definir el siguiente termino, el cual aparece al final de (2.12)

$$s(t) = k_{off}C(t) - k_{on}(p_{Tot} - C(t))X(t), \quad (2.13)$$

el cual representa una señal, específicamente la retroactividad. Cuando $s = 0$ se pierde todos los efectos de la retroactividad, y regresamos al caso aislado. El efecto de la retroactividad s sobre el comportamiento de X puede ser muy grande (figura 2.12). Esto no es deseable en una serie de situaciones en las que nos gustaría que un sistema UT "impulsara" a uno DT, por ejemplo, cuando un oscilador biológico tiene que cronometrar varios procesos DT. La retroactividad puede hacer que la señal de salida del proceso ascendente se vuelve demasiado baja o desfasada con respecto a la señal de salida del sistema aislado (como en la figura 2.12), entonces se perderá la coordinación entre el oscilador y los procesos descendentes.

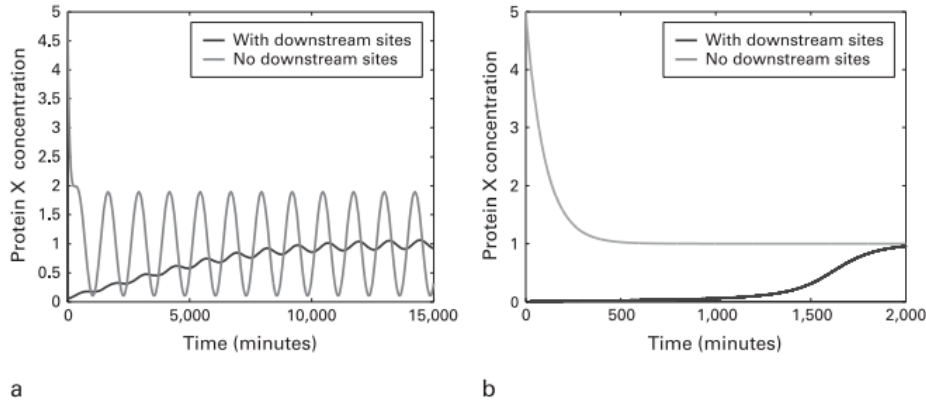


Figura 2.12: Desfase por retroactividad [9].

Podemos pensar en un sistema más general, como en la figura 2.13, donde ya se considera los efectos de la retroactividad, así como el de las funciones de entrada, muy similar a lo que hemos hecho en el capítulo anterior, las ecuaciones que describen la dinámica de este sistema son

$$\begin{aligned} \frac{dx(t)}{dt} &= f(x(t), u(t), s(t)), \\ y(t) &= Y(x(t), u(t), s(t)), \\ r(t) &= R(x(t), u(t), s(t)), \end{aligned} \quad (2.14)$$

aquí el sistema podría estar conectado a otro, por lo que se considera dos tipos de retroactividad, la retroactividad de salida s y la retroactividad de entrada r . Las funciones que aparecen de lado derecho de la ecuación (2.14), pueden tomar cualquier forma, dependiendo del sistema que se esté analizando, al igual que la retroactividad, estas pueden ser funciones escalares o vectoriales, también se podría analizar un sistema que está compuesto de varios de estos módulos. Cuando la retroactividad ya no se encuentra en el sistema se elige $r = 0$ y $s = 0$.

2.5.2. Cuantificando la retroactividad

Se puede obtener una cuantificación de la retroactividad de la salida explotando la diferencia de escalas de tiempo entre la dinámica de la salida del módulo UT y la dinámica de la etapa de entrada del módulo DT. Esta separación de escalas de tiempo siempre se encuentra en circuitos transcripcionales como se discutió anteriormente. Cuantificamos la diferencia entre la dinámica de X en el sistema aislado (2.10) y la dinámica de X en el sistema conectado (2.12) estableciendo



Figura 2.13: Sistema con retroactividad [9].

condiciones sobre los parámetros biológicos bajo las cuales los dos sistemas exhiben un comportamiento similar. Esto se logra aprovechando la diferencia de escalas de tiempo entre los procesos de producción y desintegración de proteínas y el proceso de unión y desvinculación en el promotor p .

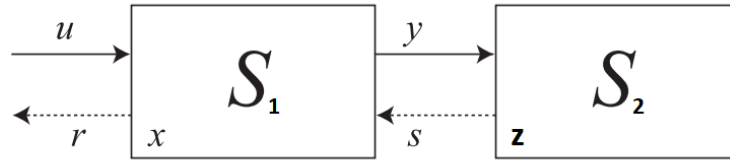


Figura 2.14: Sistemas con retroactividad.

Podemos pensar en que un sistema similar al de la figura 2.14, el primer sistema está descrito por las ecuaciones que se encuentran en (2.14), ahí suponemos que nuestro sistema se encuentra cerca de algún punto de equilibrio por lo que podremos realizar una aproximación lineal, de manera similar a como hemos hablado en las secciones anteriores, con lo que obtenemos

$$\begin{aligned} \frac{dx(t)}{dt} &= A_1x(t) + B_1u(t) + C_1s(t), \\ y(t) &= D_1x(t) + E_1u(t) + F_1s(t), \\ r(t) &= G_1x(t) + H_1u(t) + J_1s(t), \end{aligned} \tag{2.15}$$

los términos $A_1 - J_1$ en general son matrices. Como mencionamos al principio, solo nos interesa la retroactividad s entre dos sistemas, por lo que por ahora no nos fijamos en $r(t)$, además suponemos que prácticamente toda la información que sale del sistema es debido a $x(t)$, entonces las ecuaciones anteriores se simplifican

$$\begin{aligned} \frac{dx(t)}{dt} &= A_1x(t) + B_1u(t) + C_1s(t), \\ y(t) &= D_1x(t), \end{aligned} \tag{2.16}$$

estas ecuaciones corresponderán al sistema S_1 , siguiendo un análisis similar para el sistema S_2 , obtenemos

$$\frac{dz(t)}{dt} = -s(t), \tag{2.17}$$

hemos supuesto que toda la dinámica de nuestro sistema S_2 es debido a la interacción química entre x y z , es decir, se supone que la retroactividad está completamente dada por la interacción de los procesos químicos de S_1 y S_2 . Debido a que hay una diferencia muy grande en la escala

de tiempo de las dinámicas de los procesos de S_1 con respecto de S_2 , para considerarlo vamos a multiplicar a la segunda ecuación de (2.17) por un parámetro ϵ , fijándonos solo en (2.16) y (2.17):

$$\begin{aligned}\frac{dx(t)}{dt} &= A_1x(t) + B_1u(t) + C_1s(t), \\ \epsilon \frac{dz(t)}{dt} &= -\epsilon s(t),\end{aligned}\tag{2.18}$$

introducimos un cambio de variable, $\Omega = x + C_1z$, esto se agrega como una combinación lineal de las z 's debido a que es posible que más de una variable participe en una reacción

$$\begin{aligned}\frac{d\Omega(t)}{dt} &= A_1(\Omega(t) - C_1z(t)) + B_1u(t), \\ \epsilon \frac{dz(t)}{dt} &= -\epsilon s(t),\end{aligned}\tag{2.19}$$

estas ecuaciones aparecen muy a menudo en la perturbación singular, si pensamos en el sistema que está en la figura 2.11, la Ω jugaría un papel equivalente a la concentración total de la proteína, podemos obtener una aproximación de la dinámica de x , eligiendo al parámetro $\epsilon = 0$, eligiendo este parámetro de esta manera la segunda ecuación de (2.19) sería de inmediato cero, es decir $\epsilon s(t) = 0$, $s(t)$ es función de Ω , x y z , de aquí se solucionaría para encontrar una z que cumpliría lo anterior, diciéndolo de manera alternativa, suponemos que existe una z que definiremos por $w(t) = w(\Omega, t)$, tal que es el cero más pequeño de la función $g(\Omega, z, t) = -s(t)$, con $\epsilon = 0$ y al evaluar $w(t)$ (2.19) se reduciría

$$\frac{d\Omega_l(t)}{dt} = A_1(\Omega_l(t) - C_1w(t)) + B_1u(t),\tag{2.20}$$

suponemos que existe alguna solución de (2.20), con lo cual, regresando a la definición de la Ω , tendremos que $x = \Omega_l - w$, y además se supone que w es función de la omega, por lo que al tomar la derivada total con respecto del tiempo

$$\frac{dx(t)}{dt} = \frac{d\Omega_l(t)}{dt} - C_1 \frac{dw(t)}{d\Omega_l} \frac{d\Omega_l(t)}{dt} = \left(1 - C_1 \frac{dw(t)}{d\Omega_l}\right) \frac{d\Omega_l(t)}{dt},\tag{2.21}$$

con esto finalmente tendremos

$$\frac{dx(t)}{dt} = (A_1x(t) + B_1u(t)) \left(1 - C_1 \frac{dw(t)}{d\Omega_l}\right),\tag{2.22}$$

cuando $\frac{dw(t)}{d\Omega_l}$ valgan cero nos encontraremos ante el caso aislado,

$$\frac{dx(t)}{dt} = A_1x(t) + B_1u(t),\tag{2.23}$$

en base a los desarrollos que ya hemos realizado, podemos realizar una aproximación cuasi-estática del sistema eligiendo a $\epsilon \approx 0$, por lo que se podría pensar al factor $\frac{dw(t)}{d\Omega_l}$ como una variación porcentual de la dinámica del sistema conectado con respecto a la dinámica del sistema aislado en la aproximación cuasi-estática. Por lo que ahora definimos

$$\mathcal{R}(x) = C_1 \frac{dw(\Omega_l)}{d\Omega_l} = -C_1 \frac{\partial g}{\partial \Omega_l} \left(\frac{\partial g}{\partial w} \right)^{-1},\tag{2.24}$$

está es una cantidad que cuantifica la retroactividad, \mathcal{R} es una cantidad que se puede calcular empleando derivada implícita sobre $g(t)$. Cuando el sistema tiene retroactividad \mathcal{R} es una cantidad que siempre es menor que uno, de hecho, se puede calcular para el sistema (2.12), obteniendo,

$$\mathcal{R}(X) = \frac{1}{1 + \frac{(1 + \frac{X}{k_d})^2}{\frac{p_T O_T}{k_d}}}.\tag{2.25}$$

En este ejemplo, para que la retroactividad desaparezca por completo o su contribución sea muy pequeña, el termino $\frac{(1+\frac{X}{k_d})^2}{\frac{p_{Tot}}{k_d}}$ tiene que ser lo suficientemente grande, para ello se pueden elegir algunas variables adecuadas, por ejemplo, que p_{Tot} sea muy pequeño, también eligiendo a k_d y/o X muy grandes.

2.6. Dispositivos de aislamiento para reforzar la modularidad

No siempre es posible diseñar una interconexión tal que la retroactividad sea baja. Este es, por ejemplo, el caso de un oscilador que tiene que sincronizar una carga DT: en general, la carga no se puede incluir en el diseño ya que el oscilador debe funcionar bien frente a cargas desconocidas y posiblemente variables. Dicho esto, al igual que con los circuitos eléctricos, se puede diseñar un dispositivo que será colocado entre el oscilador y la carga, de modo que la salida del dispositivo no se vea afectada por la carga, y el dispositivo en sí no afecte a la carga. Específicamente, considere un sistema S como se muestra en la figura 2.15 que toma u como entrada y da y como salida, diseñado de tal manera que

1. La retroactividad de la entrada r sea pequeña.
2. Los efectos de la retroactividad de la salida s en la dinámica del sistema sean muy pequeños.
3. La relación entre la salida y entrada sean lineales (similar a lo que aparece en 2.15).

Un sistema que llegue a cumplir estas características se le nombra sistema "aislado". En un amplificador electrónico, el efecto de la retroactividad de la salida en el comportamiento del amplificador se reduce a casi cero en virtud de una gran ganancia de amplificación de entrada (teóricamente infinita) y una retroalimentación de salida negativa. Tal mecanismo se puede ilustrar en su forma más simple en la figura 2.15(a), que es muy conocida por los ingenieros de control. Para simplificar, hemos asumido un diagrama en el que la retroactividad s es solo una perturbación aditiva. El efecto de la retroactividad s es insignificante para grandes valores de G , esto puede verificarse

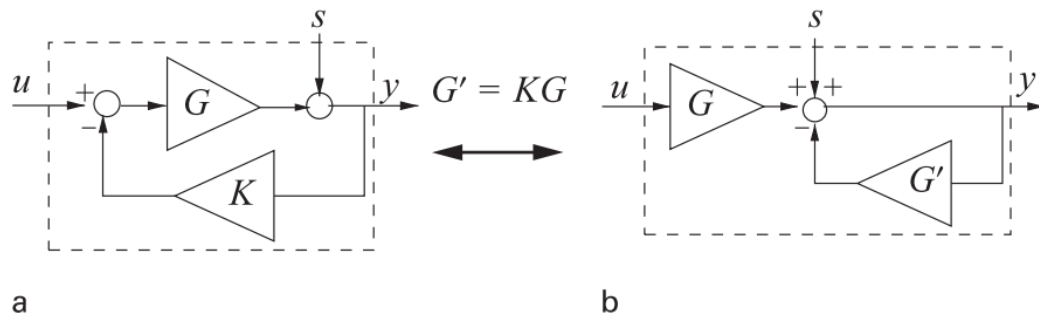


Figura 2.15: Sistema con retroalimentación [9].

mediante el siguiente cálculo, de acuerdo con las funciones de transferencia que ya hemos explicado $y = G(u - Ky) + s$, entonces

$$y = u \frac{G}{1 + KG} + \frac{s}{1 + KG}, \tag{2.26}$$

Cuando G crece mucho, nuestra variable de salida es $\frac{u}{K}$ y es independiente de s . Para atenuar el efecto de retroactividad en la salida de un componente, se pueden hacer alguna de las dos opciones:

1. amplifica la entrada del componente a través de una gran ganancia (G muy grande),

2. aplica una gran retroalimentación negativa de salida (figura 2.15b) (G' muy grande).

A continuación, ilustramos esta idea en el contexto del ejemplo transcripcional. Inspirándonos en el sistema de la figura 2.15(b) y con el propósito de ir atenuando la retroactividad. Considerando una dinámica aproximada de X . Supongamos que podemos aplicar una ganancia G a la entrada $k(t)$ y una ganancia de retroalimentación negativa G' a X con $G' = KG$. Esto conduce a la nueva ecuación diferencial para el sistema conectado dada por

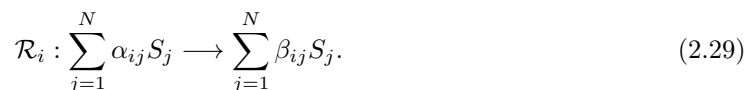
$$\frac{dX(t)}{dt} = (Gk(t) - (G' + \delta)X(t))(1 - \mathcal{R}), \quad (2.27)$$

Donde el parámetro \mathcal{R} es la que cuantifica la retroalimentación del sistema, cuando este parámetro es cero, regresamos al caso en el cual nuestro sistema se encuentra aislado.

$$\frac{dX(t)}{dt} = Gk(t) - (G' + \delta)X(t). \quad (2.28)$$

2.7. Retroactividad en procesos bioquímicos

Los desarrollos hechos por Del Vecchio son exclusivamente para sistemas cerca del equilibrio y lineales, por lo que es necesario linealizar el sistema para poder llegar al termino de retroactividad, sin embargo, nosotros mostraremos una alternativa para poder analizar redes bioquímicas, aquellas que están descritas por la ley de acción de masas, recordemos que en la ley de acción de masas se parte de un sistema en el cual hay N especies S_j ($j \in \{1, 2, \dots, N\}$), y M reacciones \mathcal{R}_i ($i \in \{1, 2, \dots, M\}$) a través de la cual se transforman estas especies, es decir



Los coeficientes α_{ij} y β_{ij} son enteros positivos, estos son los coeficientes estequiométricos. En química los elementos que están del lado izquierdo de (2.29) se les conoce como reactantes y a los del lado derecho como productos. Ahora definimos la matriz estequiométrica

$$\Gamma_{ji} = \beta_{ij} - \alpha_{ij}. \quad (2.30)$$

Denotamos por x_j a la concentración de la especie S_j involucrada en la reacción, R_i a la forma algebraica de la i -ésima tasa de reacción en función de las concentraciones de las especies. Ahora supondremos que tenemos algún sistema bioquímico que consiste en n módulos, la elección de módulo es un poco arbitraria por lo que uno puede llegar a elegir una sola variable como módulo, pero para los sistemas que analizaremos estos se eligen considerando que alguna reacción (o reacciones) se encuentra aislada, para esto pensar que se tiene un sistema similar al de la figura 2.16.

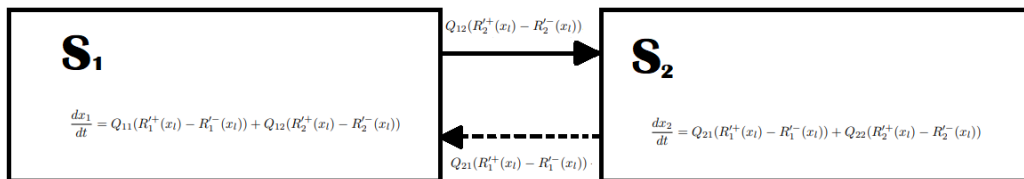


Figura 2.16: Sistema modular. Separación de un sistema en módulos y los términos de interacción.

(Este sistema solo tiene 2 módulos, pero ayudará a entender las ideas que explicaremos) si existen cantidades conservadas en el sistema, reducimos aquellas que pertenecen a un solo módulo

introduciendo las correspondientes cantidades conservadas para así tener una matriz estequiométrica (2.30) cuadrada, esta puede ser escrita de la siguiente manera

$$\Gamma_{ij} = \begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1n} \\ Q_{21} & Q_{22} & \dots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \dots & Q_{nn} \end{pmatrix}, \quad (2.31)$$

(los términos Q_{ij} pueden ser matrices, vectores o escalares) donde los términos Q_{nn} corresponden a la dinámica del n-ésimo módulo sin considerar la interacción con otros, por lo que los términos Q_{ni} ($i \neq n$ y $i > n$) corresponden a la interacción de los módulos tipo DT, mientras que los términos Q_{ni} ($i \neq n$ y $i < n$) son interacción del tipo UT, por lo general $Q_{in} \neq Q_{ni}$, como se puede observar en la figura 2.16. Siguiendo la misma idea podemos escribir a nuestras tasas de reacción (1.5) divididos con respecto a cada módulo del sistema,

$$R'^+(x_l) = \begin{pmatrix} R_1^+(x_l) \\ R_2^+(x_l) \\ \vdots \\ R_n^+(x_l) \end{pmatrix}, \quad R'^-(x_l) = \begin{pmatrix} R_1^-(x_l) \\ R_2^-(x_l) \\ \vdots \\ R_n^-(x_l) \end{pmatrix}, \quad (2.32)$$

(los términos $R_i^+(x_l)$ y $R_i^-(x_l)$ pueden ser vectores o escalares) recordar que la elección de modulo es un poco arbitraria por lo que uno puede llegar a elegir una sola variable como modulo. Las ecuaciones de movimiento de algún módulo serán de la forma

$$\frac{dx_i}{dt} = \sum_j Q_{ij}(R_j^+(x_l) - R_j^-(x_l)), \quad (2.33)$$

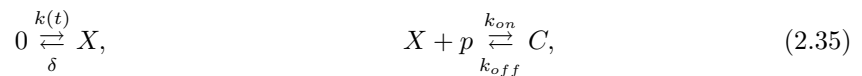
(esta ecuación diferencial es solo otra versión de la que se obtiene de la ley de acción de masas $\frac{dx_i}{dt} = \sum_j \Gamma_{ij}(R_j^+(x_l) - R_j^-(x_l))$) los coeficientes de interacción vuelven a aparecer en los módulos con los que interactúan, pero con una cantidad proporcional a ellos mismos, es decir, si tuviéramos un sistema similar a la figura 2.16 entonces κQ_{12} ($\kappa < 0$) aparecerá dentro de Q_{22} (en algunos casos es necesario introducir algún cambio de coordenadas), para saber que módulos están conectados entre sí usamos una matriz de conexiones A_{ij} , esta matriz de conexiones la construimos a partir de los valores absolutos de las entradas de Γ_{ij} pero quitando los términos en la diagonal, es decir

$$A_{ij} = \begin{pmatrix} 0 & |Q_{12}| & \dots & |Q_{1n}| \\ |Q_{21}| & 0 & \dots & |Q_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |Q_{n1}| & |Q_{n2}| & \dots & 0 \end{pmatrix}, \quad (2.34)$$

hay que recordar que el sentido de conexión va en un solo sentido por lo que $A_{ij} \neq A_{ji}$, antes de continuar expondremos un ejemplo de esto.

Ejemplo

Para esto pensemos en un sistema que esta descrito por los siguientes dos procesos



donde denotaremos a x , p , c las concentraciones de X , P y C respectivamente, hemos denotado a la reacción de la izquierda como el módulo 1 y a la de la derecha como el 2, a partir de estas reacciones calculamos la matriz de coeficientes estequiométricos y la estequiométrica del sistema

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Gamma_{ij} = \begin{pmatrix} 1 & -1 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}, \quad (2.36)$$

este sistema tiene una cantidad conservada $p_{Tot} = p + c$, considerándola obtendremos la siguiente matriz estequiométrica cuadrada

$$\Gamma_{ij}^r = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \quad (2.37)$$

de aquí vemos que $Q_{12} = -Q_{22}$, como ya hemos comentado que los términos de interacción aparecen en el otro modulo. Las tasas de reacción serán

$$R'^+(x_l) = \begin{pmatrix} R'_1+(x_l) \\ R'_2+(x_l) \end{pmatrix} = \begin{pmatrix} k(t) \\ k_{on}(p_{Tot} - c)x \end{pmatrix}, \quad R'^-(x_l) = \begin{pmatrix} R'_1-(x_l) \\ R'_2-(x_l) \end{pmatrix} = \begin{pmatrix} \delta x \\ k_{off}c \end{pmatrix}, \quad (2.38)$$

y nuestra matriz de conexiones será

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \quad (2.39)$$

En este ejemplo hemos mostrado como serán los elementos que ocuparemos para analizar los sistemas bioquímicos.

Siguiendo la misma idea que Del Vecchio [9] introducimos el siguiente cambio de variable $\Omega_i = x_i + \sum_j A_{ij}x_j$, de esta manera podemos construir un conjunto de ecuaciones diferenciales de tal manera que coincidan con el que se obtiene cuando el i -ésimo módulo este aislado, obteniendo

$$\frac{d\Omega_i}{dt} = \frac{dx_i}{dt} + \sum_j A_{ij} \frac{dx_j}{dt} = Q_{ii}(R'_i+(x_l) - R'_i-(x_l)), \quad (2.40)$$

si suponemos que el término de interacción entre los dos módulos es muy pequeño, podemos realizar un desarrollo similar al realizado por Del Vecchio de tal manera que ahora las ecuaciones de movimiento de nuestro sistema serán como a continuación

$$\begin{aligned} \frac{dx_i}{dt} &= Q_{ii}(R'_i+(x_l) - R'_i-(x_l)) \left(1 + \sum_j A_{ij} \mathcal{R}_j^i \right), \\ A_{ij} \mathcal{R}_j^i &= A_{ij} \left(\frac{dw_j}{d\Omega_i} \right) = -A_{ij} \frac{\left(Q_{ij} \frac{\partial(R'_j+(\Omega_l + \sum_j A_{lj}w_j) - R'_j-(\Omega_l + \sum_j A_{lj}w_j))}{\partial\Omega_i} \right)}{\left(Q_{ij} \frac{\partial(R'_j+(\Omega_l + \sum_j A_{lj}w_j) - R'_j-(\Omega_l + \sum_j A_{lj}w_j))}{\partial w_i} \right)}, \end{aligned} \quad (2.41)$$

donde w_l es el cero más pequeño del término $Q_{ij}(t_j^+(\Omega_l + \sum_j Q_{lj}^I w_j) - t_j^-(\Omega_l + \sum_j Q_{lj}^I w_j))$, el segundo término se puede interpretar como la retroactividad del módulo j sobre el módulo i , notar que esta evaluada sobre el cambio de variable que propusimos. De esta forma hemos construido

el termino de retroactividad para los sistemas químicos, el cual resultará ser de gran ayuda en el futuro. Ya que estamos considerando la interacción entre varios módulos es necesario que $|\sum_j A_{ij} \mathcal{R}_j^i| \ll 1$ para que la retroactividad tenga un efecto pequeño. Notar que la retroactividad es una cantidad que nos dice cuanto se modificará la dinámica del nuevo módulo que introduzcamos al sistema, pero también puede decirnos cuanto cambiará el sistema al agregarle un nuevo módulo.

Para el ejemplo que tratamos la retroactividad del sistema será

$$\mathcal{R} = \frac{1}{1 + \frac{(1 + \frac{X}{k_d})^2}{\frac{pT_{OT}}{k_d}}} \quad (2.42)$$

El mismo resultado que se obtuvo en una de las secciones anteriores.

2.8. Proceso de fosforilación/desfosforilación

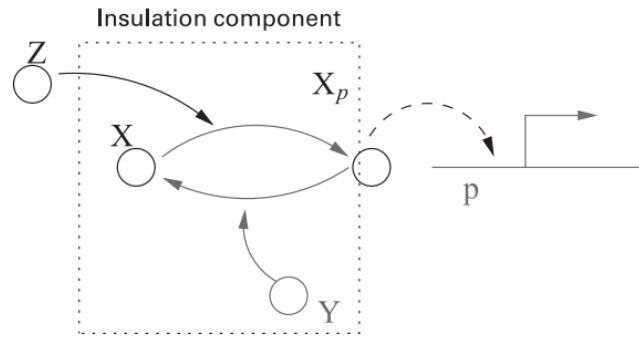
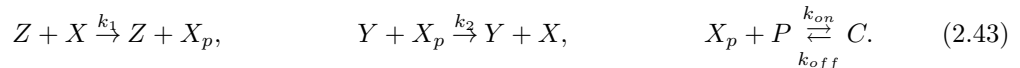


Figura 2.17: Componente aislado [9].

El proceso de fosforilación/desfosforilación es la unión (desunión) de algún grupo fosfato a una molécula. Este tipo de procesos es muy importante a nivel biológico ya que participa en muchos procesos celulares, como es el caso de fosforilación/desfosforilación del ADP/ATP respectivamente, estos dos procesos están estrechamente relacionados con la obtención y el transporte de energía. Este proceso también es usado en la fosforilación/desfosforilación de proteínas que es un mecanismo para modificar el comportamiento de una proteína, a menudo activando o desactivando una enzima. Los componentes del aparato de síntesis de proteínas también sufren fosforilación y desfosforilación y, por lo tanto, regulan las tasas de síntesis de proteínas [17]. Esto último es el que más nos interesa, ya que podría permitirnos construir módulos de control en alguna red de procesos bioquímicos al solo tener que controlar la fosforilación, particularmente algún MTT.

A continuación, aplicaremos el método de Del Vecchio que hemos descrito, para explicar el proceso de fosforilación/desfosforilación, para ello primero pensemos en una red bioquímica como en la figura 2.17. Vamos a considerar un modelo de una reacción de un paso para las reacciones de fosforilación analizadas por [8]



Las reacciones anteriores corresponden al proceso que ocurre de lado izquierdo, la primera reacción es la reacción a través de la cual la molécula X gana un fosfato (fosforilación) X_p , y la segunda cuando se desfosforara, en el lado derecho de la imagen una molécula fosforilada se une a una

proteína P formando un complejo C . La dinámica del sistema estará descrita por el siguiente conjunto de ecuaciones diferenciales (estas ecuaciones se obtienen a partir de las reacciones (2.43) con ayuda de la ley de acción de masas),

$$\begin{aligned}\frac{dX_p(t)}{dt} &= k_1 X_{Tot} Z(t) \left(1 - \frac{X_p(t)}{X_{Tot}} - \frac{C(t)}{X_{Tot}} \right) - k_2 Y(t) X_p(t) + k_{off} C(t) - k_{on} (p_{Tot} - C(t)) X(t), \\ \frac{dC(t)}{dt} &= -k_{off} C(t) + k_{on} (p_{Tot} - C(t)) X_p(t) = -s(t).\end{aligned}\quad (2.44)$$

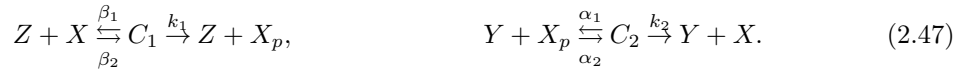
Consideramos que nos encontramos en un sistema donde $X_{Tot} \gg X_p$ y $X_{Tot} \gg C(t)$, por lo que nuestro sistema de ecuaciones se reduce,

$$\begin{aligned}\frac{dX_p(t)}{dt} &= k_1 X_{Tot} Z(t) - k_2 Y(t) X_p(t) + k_{off} C(t) - k_{on} (p_{Tot} - C(t)) X(t), \\ \frac{dC(t)}{dt} &= -k_{off} C(t) + k_{on} (p_{Tot} - C(t)) X_p(t) = -s(t).\end{aligned}\quad (2.45)$$

Ahora se considera que el sistema que se encuentra a la derecha de la figura 2.17 está casi en equilibrio, por lo que ahora usamos la aproximación con retroactividad, así como (2.27),

$$\frac{dX_p(t)}{dt} = (G(t)Z(t) - G'(t)X_p(t)) (1 - \mathcal{R}(X_p(t))), \quad (2.46)$$

se ha definido a $G(t) = k_1 X_{Tot}$ y $G'(t) = k_2 Y(t)$, para atenuar la retroactividad uno tendrá que hacer a $G(t)$ y $G'(t)$ muy grandes, o eligiendo a las variables dentro de \mathcal{R} de tal manera que se vuelva cero. Podemos pensar en un proceso más complejo para la fosforilación y desfosforilación, para ello pensamos en un modelo que implique dos pasos, este modelo que fue dado por [8], donde se supone que existe un proceso intermedio cuando se unen las diferentes especies en el módulo que se encuentra en la izquierda de la figura 2.17, como en la siguiente ecuación,



La dinámica de este sistema está dada por las siguientes ecuaciones diferenciales:

$$\begin{aligned}\frac{dZ(t)}{dt} &= k(t) - \delta Z(t) + (k_1 + \beta_2) C_1(t) - \beta_1 X_{Tot} Z(t) \left(1 - \frac{X_p(t)}{X_{Tot}} - \frac{C_1(t)}{X_{Tot}} - \frac{C_2(t)}{X_{Tot}} - \frac{C(t)}{X_{Tot}} \right), \\ \frac{dC_1(t)}{dt} &= -(k_1 + \beta_2) C_1(t) + \beta_1 X_{Tot} Z(t) \left(1 - \frac{X_p(t)}{X_{Tot}} - \frac{C_1(t)}{X_{Tot}} - \frac{C_2(t)}{X_{Tot}} - \frac{C(t)}{X_{Tot}} \right), \\ \frac{dC_2(t)}{dt} &= -(k_2 + \alpha_2) C_2(t) + \alpha_1 Y_{Tot} X_p(t) \left(1 - \frac{C_2(t)}{Y_{Tot}} \right), \\ \frac{dX_p(t)}{dt} &= k_1 C_1(t) + \alpha_2 C_2(t) - \alpha_1 Y_{Tot} X_p(t) \left(1 - \frac{C_2(t)}{Y_{Tot}} \right) + k_{off} C(t) - k_{on} X_p(t) (p_{Tot} - C(t)), \\ \frac{dC(t)}{dt} &= -k_{off} C(t) + k_{on} X_p(t) (p_{Tot} - C(t)),\end{aligned}\quad (2.48)$$

los términos que aparecen como fracciones se suponen lo suficientemente pequeños, por lo que nuestro sistema se reduce

$$\begin{aligned}
 \frac{dZ(t)}{dt} &= k(t) - \delta Z(t) + (k_1 + \beta_2)C_1(t) - \beta_1 X_{Tot}Z(t), \\
 \frac{dC_1(t)}{dt} &= -(k_1 + \beta_2)C_1(t) + \beta_1 X_{Tot}Z(t), \\
 \frac{dC_2(t)}{dt} &= -(k_2 + \alpha_2)C_2(t) + \alpha_1 Y_{Tot}X_p(t), \\
 \frac{dX_p(t)}{dt} &= k_1 C_1(t) + \alpha_1 C_2(t) - \alpha_1 Y_{Tot}X_p(t) + s(t), \\
 \frac{dC(t)}{dt} &= -k_{off}C(t) + k_{on}X_p(t)(p_{Tot} - C(t)) = -s(t),
 \end{aligned} \tag{2.49}$$

en la única expresión que aparece la retroactividad es en la tercera relación de la ecuación anterior, por lo que solo a esta expresión se le modifica de acuerdo con el modelo de Del Vecchio, $G(t)$ y $G'(t)$ se eligen de igual manera que en el proceso de un paso. Estos resultados se grafican en la figura 2.18.

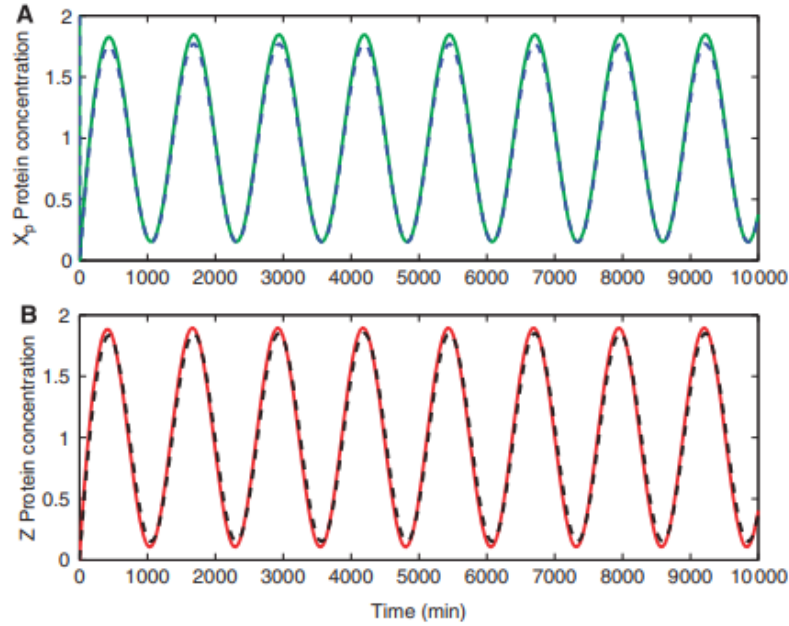


Figura 2.18: Proceso de fosforilación. Se modela la dinámica de las concentraciones de Z y X_p cuando hay retroactividad (líneas punteadas) y cuando no hay retroactividad (línea verde y roja respectivamente). Se han usado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$ [9].

En esta figura se puede observar cómo se contrasta la dinámica del sistema cuando hay y cuando no hay retroactividad, debido a los parámetros usados se puede observar que la retroactividad tiene un efecto muy pequeño en el sistema por lo que la dinámica es casi la misma que en el caso aislado, sin embargo, si se eligen otras variables de las constantes del sistema es posible que ya no haya retroactividad como se ve en la figura 2.19.

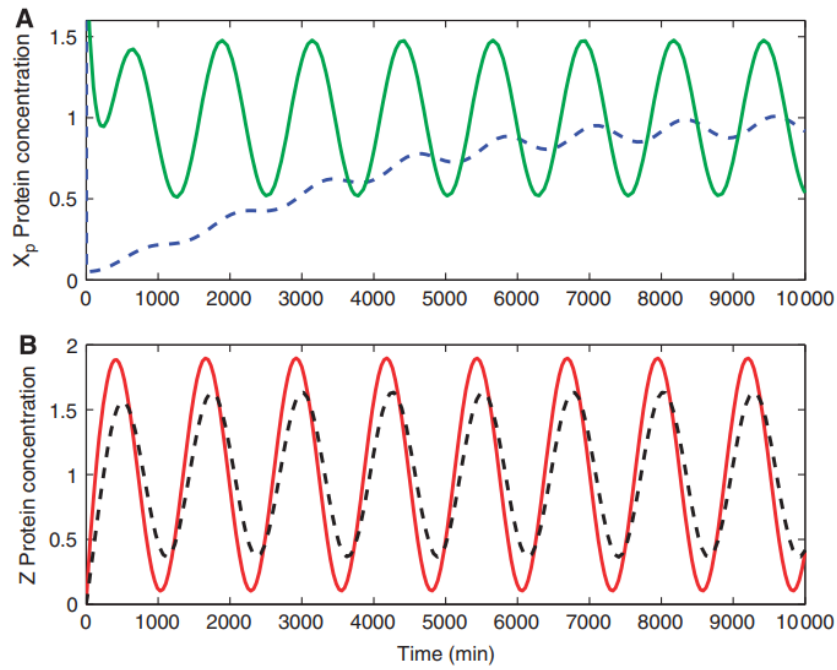


Figura 2.19: Proceso de fosforilación. Se modela la dinámica de las concentraciones de Z y X_p cuando hay retroactividad (líneas punteadas) y cuando no hay retroactividad (línea verde y roja respectivamente). Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 0,01$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$ [9].

Capítulo 3

Análisis estocástico

En esta sección hablaremos un poco sobre los modelos estocásticos, así como algunos conceptos y definiciones relacionados con esto, con el objetivo de llegar a la ecuación maestra, algunas de las ideas fueron tomadas de [10], estudiaremos los sistemas desde un enfoque estocástico debido a que estamos interesados en las redes bioquímicas que tengan un número pequeño de elementos y/o moléculas (particularmente en redes de regulación genética), por lo que el análisis determinista se vuelve insuficiente. Entre los temas que se tocarán serán los procesos de vida y muerte multivariable, el cual es una versión estocástica de la ley de acción de masas. Posteriormente mostramos una manera de cómo llegar al límite determinista, además describiremos algunos métodos de solución de la ecuación maestra, debido a que no siempre se puede encontrar una solución analítica también explicaremos el algoritmo de Gillespie, el cual es muy usado para modelar y/o simular sistemas estocásticos, en esta misma dirección mostramos un programa escrito en Python que nos describe este algoritmo y que usaremos en los próximos capítulos.

3.1. Introducción

Los modelos deterministas solo son válidos cuando el sistema se puede considerar macroscópico, e incluso en esta aproximación el sistema puede llegar a presentar desviaciones significativas, por lo que es necesario un enfoque estocástico.

Se conocen varias situaciones en las que asumir funciones continuas de concentración-tiempo resulta en una aproximación muy burda. A menudo, las fluctuaciones naturales (es decir, no causadas por influencias externas) son importantes en el sistema. Algunas de las posibles razones son las siguientes:

- i) **El tamaño del sistema químico es pequeño.** En este caso, el espacio de estados es discreto, la aproximación continua es muy mala. Esto es debido a que las fluctuaciones no se pueden descuidar ni siquiera en la aproximación “cero”, porque no se superponen al fenómeno, sino que representan el fenómeno en sí.
- ii) **El sistema opera cerca de un punto de inestabilidad de un modelo determinista.** En este caso, las pequeñas fluctuaciones pueden amplificarse y producir efectos observables, incluso macroscópicos.
- iii) **Las fluctuaciones pueden ser una fuente de información.** El teorema de fluctuación-disipación conecta las fluctuaciones espontáneas alrededor de un estado de equilibrio y el proceso disipativo que conduce al equilibrio. Usando este teorema aplicado a la cinética química, las constantes de velocidad se pueden calcular a partir de las fluctuaciones de equilibrio.

- iv) **Se encuentra una importancia específica para una concentración que cae a cero en un valor de tiempo finito.** Como se señaló anteriormente, esto no es posible dentro del marco matemático de la cinética determinista del tipo de acción de masas. Sin embargo, el enfoque estocástico puede asociar este evento con una probabilidad distinta de cero.

La alternativa a la cinética determinista que será utilizada se denomina modelo estocástico de estado discreto en tiempo continuo. Su relación con la cinética determinista es análoga a la conexión entre la termodinámica clásica y la estadística. La termodinámica estadística puede dar algún tipo de predicción teórica sobre las fluctuaciones de equilibrio, la cinética estocástica puede describir la evolución en el tiempo de los fenómenos de fluctuación química.

3.2. Procesos estocásticos

A continuación, daremos algunas definiciones y elementos necesarios para poder llegar a entender la ecuación maestra. Primero daremos la definición de proceso estocástico. Sea T algún conjunto no vacío (relacionado con el tiempo)

Un proceso estocástico es una sucesión de variables aleatorias $\xi(t)$ tal que $\{\xi(t) \mid t \in T\}$, donde estas variables aleatorias toman valores de algún espacio de estados S .

Por la forma en la que se ha definido un proceso estocástico, estas se dividen en cuatro categorías dependiendo de la naturaleza continua o discreta de la variable temporal y la variable aleatoria,

- i) **Tiempo discreto y estados discretos.** Aquí la variable temporal y aleatoria son ambas discretas, en el caso que sea un proceso de Markov, se le llama cadenas de Markov.
- ii) **Tiempo discreto y estados continuos.** La variable temporal es discreta y la aleatoria toma valores de un espacio continuo, si el proceso es Markoviano, se le llama proceso Markoviano de tiempo discreto.
- iii) **Tiempo continuo y estados discretos.** La variable temporal es continua y la aleatoria toma valores de un espacio discreto, si el proceso es Markoviano, se le llama cadena de Markov de tiempo continuo.
- iv) **Tiempo continuo y estados continuos.** La variable temporal es continua y la aleatoria toma valores de un espacio continuo, si el proceso es Markoviano, se le llama proceso Markoviano de tiempo continuo.

Para los procesos bioquímicos que nosotros estudiaremos en este texto nos enfocaremos principalmente en el tercero. Además, dentro de los procesos estocásticos nos interesara en particular los procesos estacionarios, este es un proceso estocástico cuyas distribuciones de probabilidad conjunta no cambian cuando se desplazan en el tiempo. De manera más precisa, suponemos que $F_t(\xi_t)$ es la función de distribución de probabilidad del proceso estocástico ξ_t en el instante t , se dice estacionario si

$$F_t(\xi_t) = F_{t+\tau}(\xi_{t+\tau}). \quad (3.1)$$

Dentro de los procesos estocásticos existe una gran familia muy conocida, los procesos Markovianos (en la figura 3.1 se pueden observar otros tipos), está definida de tal manera que la densidad de probabilidad condicional $f(x_n, t_n | x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1})$ ($x_i \in S$ y $t_i \in T$) solo depende del último estado, es decir

$$f(x_n, t_n | x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1}) = f(x_n, t_n | x_{n-1}; t_{n-1}). \quad (3.2)$$

En los procesos Markovianos podemos definir una transición de probabilidad entre dos diferentes estados

$$f_{ij}(s, t) = f(x_n = j, t_n = t | x_{n-1} = i; t_{n-1} = s), \quad (3.3)$$

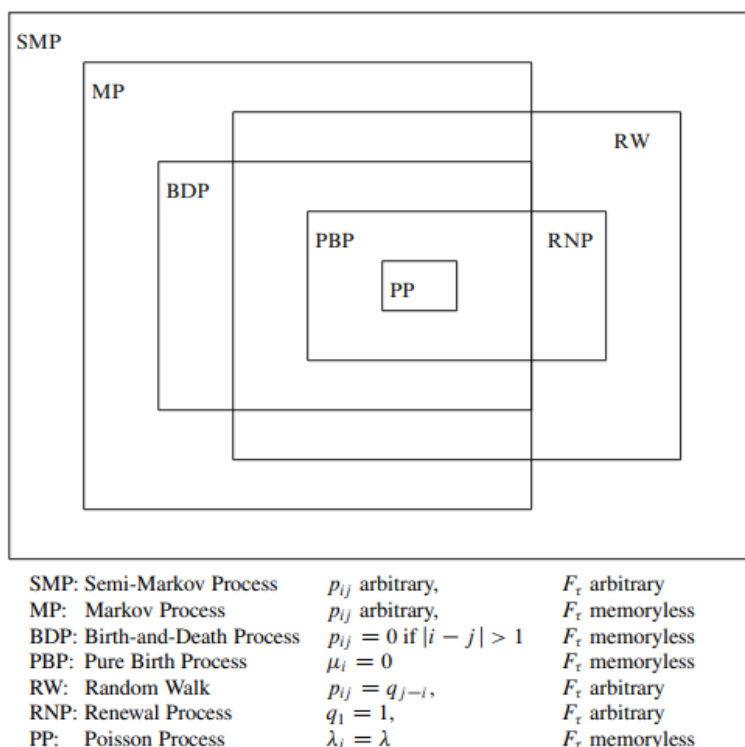


Figura 3.1: Clasificación de procesos estocásticos [10].

esta ecuación cumple una condición de normalización ($i, j \in S$ y $t, s \in T$)

$$\sum_j f_{ij}(s, t) = 1, \quad (3.4)$$

Si imaginamos que pasamos por algún estado intermedio para llegar a otro, uno puede llegar a deducir la ecuación de Chapman-Kolmogorov ($i, j \in S$ y $t, s \in T$)

$$f_{ij}(s, t) = \sum_k f_{ik}(s, t) f_{kj}(s, t), \quad (3.5)$$

es decir, sumamos sobre todos los posibles caminos para ir de un estado j a i al pasar por estados k . Su contra parte continua es

$$p(x_3|x_1, \tau + \tau') = \int_{-\infty}^{\infty} p(x_3|x_2, \tau') p(x_2|x_1, \tau) dx_2. \quad (3.6)$$

Esta ecuación en general es complicada de analizar, además, en muchos procesos físicos las transiciones de probabilidad ocurren en tiempos muy pequeños y las probabilidades solo dependen de la diferencia de tiempos en el que ocurren los procesos, s y t ($\tau = t - s$), por lo que la transición de probabilidad se puede aproximar de la manera siguiente

$$p(x|z, \tau') = (1 - a_0\tau')\delta(x - z) + \tau'w(x|z) + o(\tau'), \quad (3.7)$$

con $w(x|z)$ como la transición de probabilidad por unidad de tiempo y

$$a_0 = \int_{-\infty}^{\infty} w(x|z) dx, \quad (3.8)$$

al sustituir (3.7) en (3.6) se encuentra la siguiente ecuación diferencial

$$\frac{d}{d\tau}p(x_3|x_1, \tau) = \int_{-\infty}^{\infty} [w(x_3|x_2)p(x_2|x_1) - w(x_2|x_3)p(x_3|x_1, \tau)]dx_2. \quad (3.9)$$

la relación anterior usualmente es llamada la "ecuación maestra", de igual manera se tiene una versión discreta de la ecuación

$$\frac{d}{dt}p_n(t) = \sum_m [w_{nm}p_m(t) - w_{mn}p_n(t)]. \quad (3.10)$$

Esta ecuación resultará muy útil para analizar los sistemas bioquímicos, aunque esta ecuación podría llegar a representar algunas limitantes por la suposición de que las transiciones de probabilidad ocurren en tiempos muy pequeños, para los objetivos que tenemos no es necesario preocuparse de esto.

3.3. Proceso de vida y muerte multivariable

Debido a que estamos interesados en reacciones químicas, es necesario un formalismo estocástico que los describa, entonces usaremos a los procesos de vida y muerte multivariable. Existen varios procesos en donde están involucrados varios tipos de partículas (o diferentes especies), la dinámica detrás de estos se describen principalmente a través de saltos entre el número elementos, dado que existe la posibilidad que haya un número pequeño de estos o el sistema tienen pocos elementos, entonces el sistema presenta fluctuaciones, por lo que las ecuaciones deterministas se vuelven insuficientes, un buen enfoque para analizar tales sistemas es a través de los procesos de vida y muerte estocásticos, algunos sistemas donde esta descripción resulta ser bastante precisa son los siguientes:

1. Reacciones químicas: las moléculas se transforman por las colisiones entre ellas.
2. Poblaciones: en un espacio conviven diferentes especies y estas se depredan y/o compiten entre sí.
3. Epidemias: La enfermedad es transmitida de individuo a individuo. [18]

Estos procesos son un equivalente a la ley de acción de masas determinista, por lo que ahora construiremos su contra parte estocástica, una ecuación maestra, para esto imagine que hay un proceso a través del cual hay N especies S_j ($j \in 1, 2, ..N$), y hay M reacciones \mathcal{R}_i ($i \in 1, 2, ..M$) a través de la cual se transforman estas especies, es decir

$$\mathcal{R}_i : \sum_{j=1}^N \alpha_{ij} S_j \xrightleftharpoons[k_i^-]{k_i^+} \sum_{j=1}^N \beta_{ij} S_j. \quad (3.11)$$

Los coeficientes α_{ij} y β_{ij} son enteros positivos, estos son coeficientes estequiométricos. Una desventaja de este formalismo, es que es necesario conocer todas las especies y los coeficientes estequiométricos que involucran al proceso, lo cual a veces resultado complicado de saber (o determinar). En química los elementos que están del lado izquierdo de (3.11) se les conoce como reactivos y a los del lado derecho como productos. Ahora definimos la matriz estequiométrica

$$\Gamma_{ji} = \beta_{ij} - \alpha_{ij}. \quad (3.12)$$

A través de las colisiones (o interacciones) de los diferentes elementos estos se transforman siguiendo la ley de acción de masas, por lo que las tasas de propensión están dadas de la siguiente manera

$$t_i^+(S_j) = k_i^+ \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!}, \quad t_i^-(S_j) = k_i^- \prod_j \frac{S_j!}{\Omega^{\beta_{ij}} (S_j - \beta_{ij})!}. \quad (3.13)$$

(el índice i es el mismo que el de la reacción \mathcal{R}_i), estas tasas son las probabilidades de transición entre los diferentes estados del sistema. La Ω está estrechamente relacionada con el tamaño del sistema. Para poder escribir la ecuación maestra es indispensable conocer las reacciones que ocurren hacia "adelante", esto ocurre cuando



y en el sentido inverso



para escribir la ecuación maestra nos fijamos en las reacciones que llevan al sistema al punto S_j , y las que lo sacan, es decir las reacciones hacia adelante y hacia atrás, para ello usamos las tasas de reacción y junto con las probabilidades evaluamos en esos puntos, con lo que finalmente tenemos nuestra ecuación maestra

$$\begin{aligned} \partial_t P(S_j, t) = \Omega \sum_i & (t_i^-(S_j + \Gamma_{ji})P(S_j + \Gamma_{ji}, t) - t_i^+(S_j)P(S_j, t) \\ & + t_i^+(S_j - \Gamma_{ji})P(S_j - \Gamma_{ji}, t) - t_i^-(S_j)P(S_j, t)). \end{aligned} \quad (3.16)$$

(Aunque en verdad nuestras tasas de propensión son Ωt_i^- o Ωt_i^+ debido a que son los términos que acompañan a $P()$ en la ecuación anterior, pero no lo hemos manejado de esta manera para que el límite determinista sea más fácil de calcular.) Esta ecuación también es conocida como la ecuación química maestra, por lo mismo que hemos mencionado, es la versión estocástica de la ley de acción de masas, una gran ventaja del procedimiento que hemos desarrollado hasta este momento es que; a partir de reacciones deterministas uno determina fácilmente su contra parte estocástica junto con su correspondiente ecuación maestra, la desventaja de esto es la misma que en la ley de acción de masas, es necesario conocer a detalle todas las reacciones que ocurran, aunque este problema es más de carácter experimental.

En algunas ocasiones las cantidades conservadas del sistema nos ayudan a reducir la complejidad del sistema, estas se encuentran a partir de los vectores nulos la matriz Γ_{ij} , estos vectores nulos nos ayuda a encontrar que combinación lineal de la concentración de las especies involucradas no evoluciona en el tiempo, es decir, cuáles son las cantidades conservadas del sistema, optaremos por algo similar en el caso estocástico, usaremos el vector nulo y luego se contrae con un vector formado por S_j , suponemos que la matriz Γ_{ij} tiene l vectores nulos γ_{il} , entonces las cantidades conservadas en el sistema serán

$$\sum_i \gamma_{li} S_i = N_l, \quad (3.17)$$

la ecuación anterior corresponde a la l -ésima cantidad conservada.

3.3.1. Método alternativo para determinar la distribución estacionaria

Cuando (3.16) ya no evoluciona en el tiempo decimos que nos encontramos en los estados estacionarios del sistema, es decir se cumple la siguiente relación

$$\begin{aligned} 0 = \sum_i & (t_i^-(S_j + \Gamma_{ji})P(S_j + \Gamma_{ji}) - t_i^+(S_j)P(S_j) \\ & + t_i^+(S_j - \Gamma_{ji})P(S_j - \Gamma_{ji}) - t_i^-(S_j)P(S_j)). \end{aligned} \quad (3.18)$$

La distribución que cumple la siguiente relación se le conoce como distribución estacionaria. Necesitaremos obtener la distribución estacionaria cuando analicemos al proceso del cual se deriva la función de Hill, debido a que esta se obtiene en el límite estacionario de las ecuaciones diferenciales, entonces intuimos que la distribución estacionaria será un análogo a esto, además cuando

tengamos la distribución estacionaria podremos determinar la magnitud de las fluctuaciones, aunque en general no existe un método para encontrar la distribución estacionaria del sistema, existe un método para una clase de sistemas que es débilmente reversible o reversible [19], antes de continuar definiremos estos dos conceptos.

Definición 1: sea $\mathcal{S} = \{S_i\}$, $\mathcal{C} = \{\nu_i = \alpha_{ij}S_j, \nu'_i = \beta_{ij}S_j\}$ y $\mathcal{R} = \{\mathcal{R}_i : \nu_i \rightarrow \nu'_i\}$ denotan al conjunto de especies, los reactantes y reactivos, y finalmente a las reacciones del sistema. El conjunto $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ es llamado una red de procesos químicos (RPQ).

Definición 2: una red de procesos químicos, $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, es llamada débilmente reversible si para cualquier reacción $\nu_i \rightarrow \nu'_i \in \mathcal{R}$, existe una secuencia de reacciones tal que ν'_i es un reactante y ν_i un reactivo, es decir, existe un conjunto de reactantes y reactivos $\{\nu_1, \nu_2 \dots \nu_r, \nu'_1, \nu'_2, \dots \nu'_r\}$ tal que $\nu'_i \rightarrow \nu_1, \nu_1 \rightarrow \nu'_1 \dots \nu'_r \rightarrow \nu_i \in \mathcal{R}$. Es llamada reversible si para para cualquier reacción $\nu_i \rightarrow \nu'_i \in \mathcal{R}$ también existe $\nu'_i \rightarrow \nu_i \in \mathcal{R}$.

Definición 3: $S = \text{span}_{\nu_i \rightarrow \nu'_i \in \mathcal{R}} \{\nu_i \rightarrow \nu'_i\}$ es el subespacio estequiométrico de la red. Además, denotamos por $s = \text{Dim}(S)$

Definición 4: La deficiencia de una RPQ $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, es $\Delta = |\mathcal{C}| - l - s$, donde $|\mathcal{C}|$ es el número de reactantes y reactivos ($|\mathcal{C}| = \text{número de elementos de } \mathcal{C}$), l es el número de clases de enlace del grafo de red, y s es la dimensión del subespacio estequiométrico de la red. [19]

En [19] ellos emplean los estados estacionarios obtenidos a partir de las ecuaciones diferenciales para determinar la correspondiente distribución estacionaria, aquí desarrollaremos un método alternativo para encontrar la distribución estacionaria directamente de (3.18), para ello supondremos la siguiente hipótesis: tenemos una RPQ que es al menos débilmente reversible, tiene deficiencia cero (para asegurar convergencia) y en el estado estacionario se cumple lo siguiente

$$P(\mathbf{S})_S = \prod_i P_i(S_i), \tag{3.19}$$

lo que nos dice esta hipótesis es que la distribución estacionaria es el producto de distribuciones independientes P_i y cada una depende de una sola variable S_i , lo que sigue ahora es encontrar ecuaciones para iterar (ya que nos aparece una expresión que depende de si misma pero evaluada en un valor diferente) y encontrar cada una de las distribuciones independientes, para encontrar este estado seguiremos el siguiente algoritmo:

- 1.- Formamos dos vectores de la forma

$$\nu_i^+(S_j) = t_i^+(S_j)P(S_j), \quad \nu_i^-(S_j) = t_i^-(S_j)P(S_j). \tag{3.20}$$

- 2.- Construimos un sistema de ecuaciones de la siguiente manera

$$\begin{aligned} \sum_j \Gamma_{ij}(\nu_j^+(S_l)\theta(\Gamma_{ij}) + \nu_j^-(S_l)\theta(-\Gamma_{ij})) = \\ \sum_j \Gamma_{ij}(\nu_j^+(S_l - \Gamma_{lj})\theta(-\Gamma_{ij}) + \nu_j^-(S_l + \Gamma_{lj})\theta(\Gamma_{ij})). \end{aligned} \tag{3.21}$$

Se forma i ecuaciones, una para cada variable del sistema, estas ecuaciones se inspiran en la estructura de las ecuaciones diferenciales de la RPQ cuando se encuentran en el estado estacionario (las ecuaciones diferenciales de las concentraciones están igualadas a cero), por lo que esta última relación nos permite saber que términos de la ecuación maestra (3.18) se anulan entre sí. Se ha introducido una función de Heavinside que está definida como

$$\theta(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}, \quad (3.22)$$

3.- Dado que tenemos un sistema de ecuaciones que usa la matriz estequiométrica, es posible que existan ecuaciones que sean dependientes de otras, entonces se puede reducir la matriz (pudiendo usar el método de Gauss) o simplemente notando la forma de la matriz, por lo que definiremos a la matriz reducida como Γ'_{ij} , si el sistema se reduce entonces existen cantidades conservadas, por lo que también las introducimos en (3.21), por lo que ahora nuestro sistema de ecuaciones será

$$\sum_j \Gamma'_{ij} (\nu_j^+(S_i) \theta(\Gamma'_{ij}) + \nu_j^-(S_i) \theta(-\Gamma'_{ij})) = \sum_j \Gamma'_{ij} (\nu_j^+(S_i - \Gamma'_{ij}) \theta(-\Gamma'_{ij}) + \nu_j^-(S_i + \Gamma'_{ij}) \theta(\Gamma'_{ij})). \quad (3.23)$$

4.- Para cada ecuación que obtuvimos suponemos que se cumple (3.19), entonces tomamos el promedio sobre todas las variables restantes exceptuando a la variable S_i (y también lo denotamos en el promedio) de las ecuaciones (3.23), por lo que obtenemos algo similar a

$$\begin{aligned} & \sum_j \Gamma'_{ij} (\langle t_j^+(S_i) \rangle_{l \neq i} P_i(S_i) \theta(\Gamma'_{ij}) + \langle t_j^-(S_i) \rangle_{l \neq i} P_i(S_i) \theta(-\Gamma'_{ij})) = \\ & \sum_j \Gamma'_{ij} (\langle t_j^+(S_i) \rangle_{l \neq i} P_i(S_i - \Gamma_{ij}) \theta(-\Gamma'_{ij}) + \langle t_j^-(S_i) \rangle_{l \neq i} P_i(S_i + \Gamma_{ij}) \theta(\Gamma'_{ij})). \end{aligned} \quad (3.24)$$

Ya con estas ecuaciones encontramos la forma de la distribución estacionaria de nuestro sistema, finalmente es necesario normalizarla, por lo general nos encontraremos con una expresión similar a [19]

$$P(S_j) = \prod_i M \frac{C_i^{S_i}}{S_i!} \prod_l \delta_{\sum_i \gamma_{li} S_i - N_l, 0}, \quad (3.25)$$

el coeficiente M es una constante de normalización, la delta de Kronecker que aparece se puede modificar para escribir a unas variables en términos de otras, siempre que existan cantidades conservadas en el sistema.

Con este método es posible determinar la distribución estacionaria, en algunas ocasiones aparecen distribuciones de Poisson principalmente debido a que una reacción es reversible, en otros casos distribuciones del tipo binomial, recordar que este método solo vale para procesos que al menos sean débilmente reversibles y tengan deficiencia cero, en algunos casos el sistema no tienen deficiencia cero, entonces se pueden transformar las reacciones como en [20] para obtener deficiencia cero. Es importante notar que la distribución estacionaria que se obtiene de este método, es la distribución estacionaria en el estado estacionario.

Ejemplo

Antes de terminar esta sección daremos un pequeño ejemplo de cómo se usan las relaciones que hemos desarrollado, una gran ventaja del desarrollo realizado hasta este punto se debe a que a

partir de las reacciones deterministas se construye la ecuación maestra. Fijémonos en la siguiente reacción:



hay que recordar que los coeficientes estequiométricos α y β se construyen con ayuda de las reacciones del sistema, en este caso solo tenemos 1 reacción y 2 especies, del lado izquierdo hay una X y una A mientras que de lado derecho dos X , entonces tendremos los siguientes coeficientes estequiométricos

$$\alpha = (1, 1), \quad \beta = (2, 0), \quad (3.27)$$

restándole a β la α y luego calculamos la transpuesta obtenemos la matriz estequiométrica

$$\Gamma = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (3.28)$$

por lo que las tasas de reacción serán:

$$t^+ = k^+ ax \frac{1}{\Omega^2}, \quad t^- = k^- x(x-1) \frac{1}{\Omega^2}, \quad (3.29)$$

a representa al número de moléculas A , x al número de moléculas X , k^+ y k^- las constantes de reacción, con lo que finalmente tendremos la siguiente ecuación maestra

$$\begin{aligned} \Omega \partial_t P(a, x, t) = & k^+(a+1)(x-1)P(a+1, x-1, t) - k^-x(x-1)P(a, x, t) \\ & + k^-x(x+1)P(a-1, x+1, t) - k^+axP(a, x, t), \end{aligned} \quad (3.30)$$

este sistema se encuentra en el estado estacionario cuando la derivada temporal es igual a cero, de donde se puede encontrar la siguiente condición

$$k^+(a+1)P(a+1, x-1) - k^-xP(a, x) = 0, \quad (3.31)$$

este sistema posee una cantidad conservada $x + a = N_0$ (calculado a partir del vector nulo de Γ_{ij}), calculamos la deficiencia de nuestro sistema con $|C| = 2$, $l = 1$ y $s = 1$ obtenemos $\Delta = 0$ y la reacción es reversible, ahora aplicamos el procedimiento que explicamos anteriormente a esta última relación; primero introducimos la cantidad conservada y después se encuentra una relación que nos permite encontrar la distribución estacionaria del sistema, obteniendo finalmente

$$P(a, x) = P(0) \left(\frac{k^+}{k^-} \right)^x \frac{N_0!}{x!(N_0-x)!} \delta_{a, N_0-x}. \quad (3.32)$$

En este ejemplo obtenemos una distribución que es una variante de la distribución binomial, además, se ve que no siempre aparecen distribuciones de Poisson en el límite estacionario.

3.4. Límite determinista del proceso de vida y muerte multivariable

Ya hemos derivado la ecuación maestra para los sistemas estocásticos, y explicado brevemente como aplicarlo a un determinado sistema, además siempre es deseable comprobar que el sistema estocástico converja a uno determinista bajo ciertas condiciones (queremos que los modelos las redes de regulación genética que analizaremos más adelante converjan a los resultados deterministas).

Existen diversos enfoques de cómo entender la distribución de probabilidad, si uno los interpreta como el resultado debido a un ensamble, entonces los resultados deterministas serían el promedio de las variables sobre este. Para encontrar las variables deterministas uno se apoya de la ecuación maestra (3.10) (suponiendo que depende de una sola variable n) a la cual multiplicamos por la variable n y sacamos el promedio para finalmente obtener las ecuaciones deterministas. Un buen ejemplo para aplicar lo que hemos descrito con palabras puede ser la ecuación (3.16), para ello multiplicamos esta ecuación por alguna variable S_j , luego dividimos por Ω

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{S_j}{\Omega} P(S_j, t) \right) &= S_j \sum_i \left(t_i^-(S_j + \Gamma_{ij}) P(S_j + \Gamma_{ij}, t) - t_i^+(S_j) P(S_j, t) \right. \\ &\quad \left. + t_i^+(S_j - \Gamma_{ij}) P(S_j - \Gamma_{ij}, t) - t_i^-(S_j) P(S_j, t) \right), \end{aligned} \quad (3.33)$$

sumamos sobre todos los estados de la variable S_j

$$\begin{aligned} \frac{\partial}{\partial t} \left(\sum_{S_j} \frac{S_j}{\Omega} P(S_j, t) \right) &= \sum_{S_j} S_j \sum_i \left(t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ij}, t) - t_i^+(S_j) P(S_j, t) \right. \\ &\quad \left. + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ij}, t) - t_i^-(S_j) P(S_j, t) \right), \end{aligned} \quad (3.34)$$

reacomodando términos y definiendo el promedio de una función $f(S_j)$ como $\langle f(S_j) \rangle = \sum_{S_j} f(S_j) P(S_j, t)$, notar que esta definición vale para cualquier función, por lo que obtenemos

$$\frac{\partial}{\partial t} \left(\frac{\langle S_j \rangle}{\Omega} \right) = \sum_i \Gamma_{ij} \langle t_i^+(S_j) - t_i^-(S_j) \rangle, \quad (3.35)$$

definimos a nuestras variables deterministas como $s_j = \frac{\langle S_j \rangle}{\Omega}$ (debido a que las variables deterministas tienen unidades de concentración, Ω tendrá unidades de mol sobre volumen en el caso de sistemas químicos o su equivalente según el sistema), ahora la pregunta es como aparecen las tazas de reacción deterministas en nuestra ecuación anterior, para ello primero vemos que se puede hacer un desarrollo en serie alrededor del promedio de la siguiente manera

$$\begin{aligned} \langle f(X_i) \rangle &\approx \langle f(\langle X_i \rangle) + \sum_i (X_i - \langle X_i \rangle) \frac{\partial f(X_i)}{\partial X_i} \Big|_{X_i = \langle X_i \rangle} + \sum_i \sum_j \frac{(X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle)}{2} \frac{\partial^2 f(X_i)}{\partial X_i \partial X_j} \Big|_{X_i = \langle X_i \rangle} \rangle \\ &= f(\langle X_i \rangle) + \sum_i \sum_j \frac{(\sigma^2(X_i, X_j))}{2} \frac{\partial^2 f(X_i)}{\partial X_i \partial X_j} \Big|_{X_i = \langle X_i \rangle}, \end{aligned} \quad (3.36)$$

en esta aproximación solo se realizó hasta el segundo momento ya que es del orden de $\frac{1}{\Omega}$ y además, esta dividido por un factorial, por lo que la contribución de términos de orden superior sería despreciable, el segundo término de la primera igualdad desaparece debido a que al tomar los promedios se eliminan mutuamente los términos, hemos introducido a la covarianza debido a la definición de esta

$$\sigma^2(X_i, X_j) = (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle,$$

más adelante daremos una manera de calcularla, en el caso en el que $S_{l_1} = S_{l_2}$ la covarianza se convierte en la varianza. Inspirados en la estructura de los términos que desaparecen, podemos definir a las fluctuaciones de nuestro sistema como

$$\xi_i = X_i - \langle X_i \rangle, \quad \xi_i^2 = (X_i - \langle X_i \rangle)^2, \quad (3.37)$$

la primera relación sera la definición de fluctuación que usaremos, pero debido a su estructura no es una cantidad que pueda ser medible, sin embargo, no es el caso del segundo término ya que al

tomar su promedio se convierte en la varianza, por lo que la usaremos directamente para cuantificar las fluctuaciones en el sistema,

$$\eta_i^2 = \frac{\langle \xi_i^2 \rangle}{\Omega^2} = \frac{\sigma^2(X_i, X_i)}{\Omega^2}, \quad (3.38)$$

esta cantidad ya es medible en los sistemas estocásticos, y en la literatura por lo general se le nombra fluctuaciones, aunque en realidad es el promedio del cuadrado de las fluctuaciones dividido por Ω^2 tal como hemos mostrado, aunque tomaremos el mismo nombre que se le da en la literatura por lo que a esta última relación la nombraremos fluctuación, lo que nos dicen es que hay aleatoriedad en la interacción de las diferentes especies. Ahora lo que sigue es aplicar estas ideas al caso de los procesos de vida y muerte multivariable, para ello primero analizamos lo siguiente considerando que Ω es grande (notar que no infinita)

$$\begin{aligned} t_i^+ (\langle S_j \rangle) &= k_i^+ \prod_j \frac{\langle S_j \rangle!}{\Omega^{\alpha_{ij}} (\langle S_j \rangle - \alpha_{ij})!} \approx k_i^+ \prod_j \left(\frac{\langle S_j \rangle}{\Omega} \right)^{\alpha_{ij}} \approx k_i^+ \prod_j s_j^{\alpha_{ij}} = R_i^{D+}, \\ t_i^- (\langle S_j \rangle) &= k_i^- \prod_j \frac{\langle S_j \rangle!}{\Omega^{\beta_{ij}} (\langle S_j \rangle - \beta_{ij})!} \approx k_i^- \prod_j \left(\frac{\langle S_j \rangle}{\Omega} \right)^{\beta_{ij}} \approx k_i^- \prod_j s_j^{\beta_{ij}} = R_i^{D-}, \end{aligned} \quad (3.39)$$

estas tasas de reacción son justo las que se dan en la ley de acción de masas determinista. La aproximación que hemos realizado se puede obtener suponiendo que $\langle \frac{S_j}{\Omega} \rangle \gg \frac{\alpha_{ij}}{\Omega}$ como a continuación,

$$\frac{\langle S_j \rangle!}{\Omega^{\alpha_{ij}} (\langle S_j \rangle - \alpha_{ij})!} = \left(\frac{\langle S_j \rangle}{\Omega} \right) \left(\frac{\langle S_j \rangle}{\Omega} - \frac{1}{\Omega} \right) \dots \left(\frac{\langle S_j \rangle}{\Omega} - \frac{\alpha_{ij} - 1}{\Omega} \right) \approx \left(\frac{\langle S_j \rangle}{\Omega} \right)^{\alpha_{ij}}, \quad (3.40)$$

(uno siempre puede tomar en cuenta todos los términos sin realizar ninguna aproximación) por lo que ahora en (3.36) sustituimos esta última relación, obteniendo lo siguiente

$$\begin{aligned} R_i^+ &= \langle t_i^+ (S_j) \rangle \approx k_i^+ \left(\prod_j s_j^{\alpha_{ij}} + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(S_{l_1}, S_{l_2})}{2\Omega^2} \Omega^2 \left[\frac{\partial^2}{\partial S_{l_1} \partial S_{l_2}} \left(\prod_j \left(\frac{S_j}{\Omega} \right)^{\alpha_{ij}} \right) \right]_{S_j = \langle S_j \rangle} \right) \\ &= k_i^+ \left(\prod_j s_j^{\alpha_{ij}} + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(s_{l_1}, s_{l_2})}{2} \left[\frac{\partial^2}{\partial s_{l_1} \partial s_{l_2}} \left(\prod_j s_j^{\alpha_{ij}} \right) \right] \right), \\ R_i^- &= \langle t_i^- (S_j) \rangle \approx k_i^- \left(\prod_j s_j^{\beta_{ij}} + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(S_{l_1}, S_{l_2})}{2\Omega^2} \Omega^2 \left[\frac{\partial^2}{\partial S_{l_1} \partial S_{l_2}} \left(\prod_j \left(\frac{S_j}{\Omega} \right)^{\beta_{ij}} \right) \right]_{S_j = \langle S_j \rangle} \right) \\ &= k_i^- \left(\prod_j s_j^{\beta_{ij}} + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(s_{l_1}, s_{l_2})}{2} \left[\frac{\partial^2}{\partial s_{l_1} \partial s_{l_2}} \left(\prod_j s_j^{\beta_{ij}} \right) \right] \right), \end{aligned} \quad (3.41)$$

estas expresiones son una corrección a las tasas deterministas, sin embargo, al multiplicar los términos dentro del corchete por Ω^2 desaparece por completo Ω de ese lado, mientras que $\frac{\sigma^2(S_l, S_n)}{\Omega^2} = \sigma^2(s_l, s_n)$ es la covarianza de las variables deterministas y además proporcional a $\frac{1}{\Omega}$ (como más adelante veremos), es decir Ω aparece en el denominador, por lo que en el límite en el cual Ω se vuelve muy grande el término que contiene la covarianza se vuelve cero obtenido exactamente las ecuaciones deterministas usuales. Con todo esto reescribimos nuestra ecuación determinista como a continuación

$$\frac{\partial s_j}{\partial t} = \sum_i \Gamma_{ji} \left(R_i^{D+}(s_l) - R_i^{D-}(s_l) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \quad (3.42)$$

($R_i^{D+}(s_i)$ y $R_i^{D-}(s_i)$ denotan a las tasas deterministas) estas son un conjunto de ecuaciones diferenciales, esta es una forma de obtener las ecuaciones deterministas a partir de un modelo estocástico, ya además aparecen correcciones estocásticas, como la covarianza es proporcional al inverso de Ω por lo que cuando Ω se vuelve muy grande o cuando las tasas son lineales desaparecen las correcciones. Una posible interpretación de este término correctivo podría ser que esta es la contribución de las fluctuaciones de las interacciones de las especies involucradas. Debido a que esta es una ecuación diferencial que evoluciona en el tiempo, necesitamos encontrar una ecuación diferencial para la covarianza, aunque esto lo dejaremos para el siguiente capítulo.

También podemos encontrar una contra parte determinista de las cantidades conservadas del sistema, estas se calculan con ayuda de los vectores nulos de Γ_{ij} , suponiendo que esta matriz tiene l vectores nulos, por lo que si dividimos la expresión (3.17) sobre Ω obtenemos

$$\sum_i \gamma_{li} \frac{S_i}{\Omega} = N_l^d, \tag{3.43}$$

(se ha definido $N_l^d = \frac{N_l}{\Omega}$, N_l es la misma que aparece en (3.17)) tomamos el promedio de la expresión anterior y tomamos la definición de las variables deterministas

$$\sum_i \gamma_{li} s_i = N_l^d, \tag{3.44}$$

obteniendo finalmente la cantidad conservada determinista en términos de las variables s_i como usualmente se hace en el caso determinista. Los resultados que obtuvimos son para procesos de vida y muerte multivariable, en el caso que se esté analizando otro tipo de procesos y/o sistemas se realizaría un desarrollo similar tomando en cuenta su correspondiente ecuación maestra.

3.5. Métodos de solución de la ecuación maestra

Siempre es deseable encontrar una solución analítica de (3.10), para esto existen diversos métodos que intentan solucionarla, entre las más conocidas; la función generatriz, transformada de Laplace, representación de Poisson, dependiendo del sistema existen ventajas al usar alguno sobre los otros. A continuación, haremos una breve descripción de cada una de estas, aunque estos métodos principalmente sirven cuando el sistema es lineal (o "casi lineal"). A pesar de contar con estos métodos muchos de los sistemas no se pueden solucionar de manera analítica, por lo que uno puede llegar a emplear diferentes técnicas para encontrar una aproximación a la solución (aproximación lineal de ruido, aproximación de Fokker-Planck) o directamente simular el proceso estocástico (algoritmo de Gillespie).

3.5.1. Operaciones matriciales

Uno puede llegar escribir la ecuación maestra (3.10) en forma matricial de la siguiente manera

$$\frac{d}{dt} P_n(t) = W P_n(t), \tag{3.45}$$

donde $W = \delta_{n,m'} (\sum_{m'} w_{nm'}) - w_{nn}$, además hemos supuesto que $P_n(t)$ es un vector de todos los posibles estados del sistema, la solución de esta ecuación es la siguiente

$$P_n(t) = \exp(Wt) P_n(0), \tag{3.46}$$

la exponencial que aparece es de hecho la exponencial de una matriz, para más detalles de esta expresión se puede ver [26], aunque esta solución parece ser muy fácil de encontrar e incluso de implementar computacionalmente, surgen algunas complicaciones cuando el tamaño del sistema es lo suficientemente grande, ya que se requiere de mucho poder computacional, por lo anterior se recomienda usar este método solo cuando el sistema es pequeño.

3.5.2. Transformada de Laplace

Otra alternativa para solucionar la ecuación maestra es emplear la transformada de Laplace,

$$\mathcal{L}(f(x))(s) = \int_0^{\infty} e^{sx} f(x) dx, \quad (3.47)$$

aplicando la transformada sobre la ecuación maestra y usando sus propiedades

$$s\mathcal{L}(P_n(t))(s) = P_n(0) + W\mathcal{L}(P_n(t))(s), \quad (3.48)$$

operando esta ecuación y calculando la transformada inversa uno encuentra la distribución de probabilidad. Dado que se ha usado de nuevo la matriz W es posible que este método sea contra-productivo cuando el sistema se vuelve lo suficientemente grande.

3.5.3. Función generatriz

Apoyarse de una función generatriz es uno de los métodos más usados al momento de resolver la ecuación maestra, resulta ser una muy buena estrategia resolverla usando este método, debido a que a partir de esta fácilmente se pueden obtener promedios, además convierte a la ecuación maestra en una ecuación diferencial parcial "más fácil" de tratar. Para ello se supone que existe una función generatriz de la forma

$$G(z, t) = \sum_n z^n p_n(t), \quad (3.49)$$

donde $p_n(t)$ es la distribución de probabilidad de la variable n , de esta relación se obtiene una condición de normalización $G(1, t) = 1$, derivando se obtienen los momentos

$$\begin{aligned} \left. \frac{\partial G(z, t)}{\partial z} \right|_{z=1} &= \sum_n n z^n p_n(t) = \langle n(t) \rangle, \\ \left. \frac{\partial^2 G(z, t)}{\partial z^2} \right|_{z=1} &= \langle n^2(t) \rangle - \langle n(t) \rangle^2, \end{aligned} \quad (3.50)$$

por lo que en la ecuación maestra lo que se hace es multiplicarla por z^n y luego sumar sobre n , de tal manera que se obtiene la función generatriz, para ser más precisos una ecuación diferencial parcial para $G(z, t)$. Veremos cómo se trabaja con este método, para eso lo usaremos en algún sistema en particular, entonces aplicaremos este formalismo a la ecuación (3.16), para ello se usa la siguiente función generatriz

$$G(\mathbf{z}, t) = \sum_{S_j} \left(\prod_j z_j^{S_j} \right) P(S_j, t), \quad (3.51)$$

(\mathbf{z} es un vector) sustituyendo esta última expresión en la ecuación (3.16), podemos notar que

$$\partial_t G(\mathbf{z}, t) = \partial_t^+ G(\mathbf{z}, t) + \partial_t^- G(\mathbf{z}, t), \quad (3.52)$$

donde

$$\partial_t^+ G(\mathbf{z}, t) = \Omega \sum_{i,S} k_i^+ \left(\left(\prod_j \frac{(S_j - \Gamma_{ji})!}{\Omega^{\alpha_{ij}} (S_j - \Gamma_{ij} - \alpha_{ij})!} z_j^{S_j} \right) P(S_j - \Gamma_{ij}, t) - \left(\prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!} z_j^{S_j} \right) P(S_j, t) \right), \quad (3.53)$$

reescribiendo

$$\partial_t^+ G(\mathbf{z}, t) = \Omega \sum_{i,S} k_i^+ \left(\prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!} z_j^{S_j + \Gamma_{ij}} - \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!} z_j^{S_j} \right) P(S_j, t), \quad (3.54)$$

se puede notar que

$$\prod_j \frac{z_j^{\alpha_{ij}}}{\Omega^{\alpha_{ij}}} \partial_j^{\alpha_{ij}} z_j^{S_j} = \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!} z_j^{S_j}, \quad (3.55)$$

por lo que finalmente tendremos

$$\partial_t^+ G(\mathbf{z}, t) = \Omega \sum_i \left(\prod_j z_j^{\beta_{ij}} - \prod_j z_j^{\alpha_{ij}} \right) k_i^+ \prod_j \frac{1}{\Omega^{\alpha_{ij}}} \partial_j^{\alpha_{ij}} G(\mathbf{z}, t), \quad (3.56)$$

repetimos para $\partial_t^- G(\mathbf{z}, t)$, y juntamos los dos resultados por lo que finalmente tendremos

$$\partial_t G(\mathbf{z}, t) = \Omega \sum_i \left(\prod_j z_j^{\beta_{ij}} - \prod_j z_j^{\alpha_{ij}} \right) \left(k_i^+ \prod_j \frac{1}{\Omega^{\alpha_{ij}}} \partial_j^{\alpha_{ij}} - k_i^- \prod_j \frac{1}{\Omega^{\beta_{ij}}} \partial_j^{\beta_{ij}} \right) G(\mathbf{z}, t), \quad (3.57)$$

esta expresión es una fórmula para encontrar las ecuaciones diferenciales parciales de la función generatriz de un proceso de vida y muerte multivariable, debido a que las constantes presentes en esta ecuación vienen directamente de la parte determinista no es necesario realizar alguna consideración sobre esta, por lo que al realizar aproximaciones solo será necesario fijarse en Ω .

3.5.4. Representación de Poisson

Otro método alternativo para solucionar la ecuación maestra es la representación de Poisson, este método tiene como objetivo transformar a la ecuación maestra a una similar a la ecuación de Fokker-Planck, para ello se supone que la distribución de probabilidad está dada como superposición no correlacionada de distribuciones de Poisson:

$$P_n(t) = \int \frac{e^{-\mu} \mu^n}{n!} f(\mu, t) d\mu, \quad (3.58)$$

($P_n(t)$) representa la distribución de probabilidad de la variable n y $t \in T$) aquí la función $f(\mu, t)$ es una cuasi distribución de probabilidad, la variable μ es una variable auxiliar, los límites de integración por lo regular están dados en el intervalo 0 a ∞ , además existe una estrecha relación entre la función generatriz y la representación de Poisson

$$G(z, t) = \int e^{(z-1)\mu} f(\mu, t) d\mu, \quad (3.59)$$

(esta relación es para una distribución que depende de una sola variable) algunas relaciones muy útiles que se derivan de este método y particularmente de la relación anterior,

$$\begin{aligned} \langle n(t) \rangle &= \int \mu f(\mu, t) d\mu, \\ f(\mu, 0) &= \sum_n P_n(0) \frac{e^{-\mu} \mu^n}{n!} \delta(\mu - n). \end{aligned} \quad (3.60)$$

Podemos intentar encontrar una expresión para (3.16), y dado que existe una relación similar a (3.59) (considerando más variables)

$$G(\mathbf{z}, t) = \int \exp \left(\sum_j (z_j - 1) u_j \right) f(\mathbf{u}, t) d\mathbf{u}, \quad (3.61)$$

sustituyendo esta última expresión en (3.57) obtenemos

$$\begin{aligned} \partial_t G(\mathbf{z}, t) &= \Omega \sum_i \int d\mathbf{u} \left(\prod_j (\partial_{u_j} + 1)^{\beta_{ij}} - \prod_j (\partial_{u_j} + 1)^{\alpha_{ij}} \right) * \\ &\quad \left(k_i^+ \prod_j \frac{u_j^{\alpha_{ij}}}{\Omega^{\alpha_{ij}}} - k_i^- \prod_j \frac{u_j^{\beta_{ij}}}{\Omega^{\beta_{ij}}} \right) \exp \left(\sum_j (z_j - 1) u_j \right) f(\mathbf{u}, t), \end{aligned} \quad (3.62)$$

integramos por partes y eliminamos los términos de la frontera, finalmente tendremos

$$\partial_t f(\mathbf{u}, t) = \Omega \sum_i \left(\prod_j (1 - \partial_{u_j})^{\beta_{ij}} - \prod_j (1 - \partial_{u_j})^{\alpha_{ij}} \right) \left(k_i^+ \prod_j \frac{u_j^{\alpha_{ij}}}{\Omega^{\alpha_{ij}}} - k_i^- \prod_j \frac{u_j^{\beta_{ij}}}{\Omega^{\beta_{ij}}} \right) f(\mathbf{u}, t), \quad (3.63)$$

definiendo $J_i(\mathbf{u}) = k_i^+ \prod_j \frac{u_j^{\alpha_{ij}}}{\Omega^{\alpha_{ij}}} - k_i^- \prod_j \frac{u_j^{\beta_{ij}}}{\Omega^{\beta_{ij}}}$ y la ecuación (3.63) será

$$\partial_t f(\mathbf{u}, t) = \Omega \sum_i \left(\prod_j (1 - \partial_{u_j})^{\beta_{ij}} - \prod_j (1 - \partial_{u_j})^{\alpha_{ij}} \right) J_i(\mathbf{u}) f(\mathbf{u}, t), \quad (3.64)$$

bajo ciertas condiciones está ecuación se parece a la ecuación de Fokker-Planck, para ello definimos la siguiente variable $\mu_j = \frac{u_j}{\Omega}$, sustituyendo en la relación (3.64) obtenemos

$$\begin{aligned} \partial_t f(\mu, t) &= \Omega \sum_i \left(\prod_j (1 - \Omega^{-1} \partial_{\mu_j})^{\beta_{ij}} - \prod_j (1 - \Omega^{-1} \partial_{\mu_j})^{\alpha_{ij}} \right) J_i(\mu) f(\mu, t), \\ J_i(\mu) &= k_i^+ \prod_j \mu_j^{\alpha_{ij}} - k_i^- \prod_j \mu_j^{\beta_{ij}}, \end{aligned} \quad (3.65)$$

en el caso en el cual se cumplen las siguientes condiciones

$$\sum_j \alpha_{ij} \leq 2, \quad \sum_j \beta_{ij} \leq 2, \quad (3.66)$$

se obtiene

$$\begin{aligned} \partial_t f(\mu, t) &= - \sum_j \frac{\partial}{\partial \mu_j} [A_j f(\mu, t)] + \frac{\Omega^{-1}}{2} \sum_{jl} \frac{\partial^2}{\partial \mu_j \partial \mu_l} [B_{jl} f(\mu, t)], \\ A_j &= \sum_i \Gamma_{ji} J_i(\mu), \\ B_{jl} &= \sum_i J_i(\mu) (\beta_{ij} \beta_{il} - \alpha_{ij} \alpha_{il} - \delta_{jl} \Gamma_{ji}), \end{aligned} \quad (3.67)$$

esta última expresión es justo la ecuación de Fokker-Planck, este enfoque resulta ser bastante conveniente para los procesos de vida-muerte multivariable debido a que esta última relación coincide con el resultado de aplicar la aproximación lineal de ruido a la ecuación maestra con la que se partió, para más detalles y algunos ejemplos ilustrativos ver [18].

3.5.5. Operadores

Este último método resulta ser bastante útil en los procesos de vida y muerte, pero solo cuando el sistema es lineal, para ello este método se apoyan de los operadores de creación y aniquilación, similar a las ideas presentadas en [11], los operadores creación y aniquilación resultan ser bastante útiles e intuitivos al momento de analizar estos sistemas, debido a que la creación sería equivalente al

nacimiento o creación de una molécula, mientras que la aniquilación a la degradación, explicaremos este método para el caso de una variable, para ello realizamos la siguiente convención

$$|G(t)\rangle = \sum_{x=0}^{\infty} p_x(t) |x\rangle, \quad (3.68)$$

se ha adoptado una notación muy común en mecánica cuántica [21], se puede recuperar la la función generatriz de la siguiente manera (estos resultados se obtienen de acuerdo con la notación de Dirac)

$$\begin{aligned} \langle z|G(t)\rangle &= G(z, t), & \langle x|G(t)\rangle &= p_x(t) \\ \langle z|x\rangle &= z^x, & \langle x|z\rangle &= \frac{1}{z^{x+1}}, \\ \langle x|x'\rangle &= \delta_{x,x'}. \end{aligned} \quad (3.69)$$

Ahora aplicamos la ecuación maestra sobre algún estado $|x\rangle$ y luego sumamos sobre esta variable obteniendo una relación muy similar a la ecuación de Schrödinger, similar a la siguiente expresión

$$\frac{\partial}{\partial t} |G(t)\rangle = (a^+ + a^- + b) |G(t)\rangle, \quad (3.70)$$

en la parte izquierda de esta expresión aparecen operadores de creación y aniquilación, en algunos casos redefiniendo estos operadores, aparecerán operadores escalera similar al problema del oscilador armónico, para más detalles de cómo se usa este método ver apéndice D donde se ha desarrollado para el proceso de vida y muerte. Aunque hemos expuesto estos métodos para el caso que la distribución dependa de una sola variable, este método puede extenderse a más variables, específicamente para la ecuación maestra (3.16) encontraremos su representación con operadores, dado que los operadores de creación y aniquilación tienen diferentes representaciones dependiendo de la base, existe la siguiente correspondencia

$$a^+ \rightarrow z, \quad a^- \rightarrow \partial_z,$$

entonces (3.16) escrito en términos de operadores de creación y aniquilación será

$$\frac{\partial}{\partial t} |G(t)\rangle = \Omega \sum_i \left(\prod_j (a_j^+)^{\beta_{ij}} - \prod_j (a_j^+)^{\alpha_{ij}} \right) \left(k_i^+ \prod_j \left(\frac{a_j^-}{\Omega} \right)^{\alpha_{ij}} - k_i^- \prod_j \left(\frac{a_j^-}{\Omega} \right)^{\beta_{ij}} \right) |G(t)\rangle, \quad (3.71)$$

dependiendo de los coeficientes estequiométricos es la forma de la ecuación anterior, los operadores que aparecen en el lado izquierdo de la relación anterior se pueden escribir como operadores escalera, similar a los que aparecen en el problema del oscilador armónico, aunque no siempre será el caso.

3.6. Métodos de simulación: Algoritmo de Gillespie

Dado que en algunos casos es prácticamente imposible determinar una solución analítica de la ecuación maestra, uno se apoya de algoritmos y herramientas computacionales para intentar encontrar como se comportaría el proceso estocástico y de esta manera determinar el promedio de las variables involucradas, así como las fluctuaciones, de hecho, la mayoría de los sistemas que llegaremos a analizar serán no lineales por lo que llegaremos a usar métodos de simulación. El comienzo de las simulaciones de procesos estocásticos empezó hace cerca de 40 años con el trabajo de Gillespie [22], posteriormente se hicieron algunas modificaciones o se implementaron nuevos modelos [23]. A continuación explicaremos algunos de estos algoritmos.

El algoritmo de Gillespie es uno de los más usados para modelar y/o simular procesos químicos o bioquímicos [22], aunque existen otros métodos [23] mucho más eficientes y que más adelante comentaremos. La idea principal detrás de este algoritmo es la siguiente: si nuestro sistema se encuentra en algún estado \mathbf{X} , entonces

1. ¿Cuándo sería la siguiente reacción?
2. ¿Qué tipo de reacción sería?

dada la naturaleza aleatoria de los procesos químicos, estas preguntas solo pueden ser resueltas en términos de probabilidad, para ello introducimos la función $P(\mu, \tau)$ definida como la probabilidad que dado algún estado \mathbf{X} ocurra una reacción en el intervalo $(t + \tau, t + \tau + d\tau)$ y la reacción sea R_μ . Los valores específicos de la variable τ y μ darían respuesta a las preguntas anteriores. Gillespie mostró que es posible derivar una expresión para $P(\mu, \tau)$ y de ahí encontrar τ y μ . La hipótesis fundamental de la cual se basó es que $P(\mu, \tau)$ se puede escribir como el producto de que ocurra una reacción R_μ en el intervalo $(t + \tau, t + \tau + d\tau)$, $a_\mu d\tau$ ($a_\mu = \{\Omega t_i^+, \Omega t_i^-\}$), por la probabilidad que no ocurra $P_0(\tau)$ para un determinado estado \mathbf{X} en el tiempo t [23], por lo que tendremos

$$P(\mu, \tau) d\tau = P_0(\tau) a_\mu d\tau, \quad (3.72)$$

además, $P_0(\mu, \tau)$ tiene la siguiente forma

$$P_0(\tau + d\tau) = P_0(\tau) \left[1 - \sum_{i=1}^{2M} a_i \right], \quad (3.73)$$

el termino dentro del corchete nos habla sobre la probabilidad de que no ocurra ninguna reacción, por lo que resolviendo se obtiene lo siguiente

$$P_0(\tau) = e^{-\sum_{i=1}^{2M} a_i \tau}, \quad (3.74)$$

sustituyendo (3.74) en (3.72) obtenemos

$$P(\mu, \tau) = \begin{cases} a_\mu e^{-a_0 \tau} & \text{si } 0 \leq \tau < \infty \\ 0 & \text{en otro caso} \end{cases}, \quad (3.75)$$

donde $a_\mu = \{\Omega t_i^+, \Omega t_i^-\}$ y a_0 está definido como

$$a_0 = \sum_{i=1}^{2M} a_i = \sum_{i=1}^M \Omega (t_i^+ + t_i^-), \quad (3.76)$$

la expresión (3.75) es tan solo otra versión de la ecuación maestra, sin embargo, esta solo vale para encontrar la siguiente reacción dependiendo del estado \mathbf{X} del sistema. Existen varias formas de implementar computacionalmente este algoritmo, nosotros solo mostraremos dos de estos a continuación.

3.6.1. Método directo

Este método es que Gillespie describió originalmente, para implementarlo es necesario que tener dos números aleatorios r_1 y r_2 uniformemente distribuidos entre (0,1) para cada paso, el tiempo para que la siguiente reacción ocurra está dada por $t + \tau$, donde τ es de la siguiente manera

$$\tau = \frac{1}{a_0} \ln \left(\frac{1}{r_1} \right), \quad (3.77)$$

μ se elige como el menor coeficiente que cumple la siguiente relación

$$\sum_j^{\mu} a_j > a_0 r_2, \quad (3.78)$$

entonces el estado del sistema es actualizado a $X(t + \tau) = X(t) + S_{\mu}$, donde S es la matriz estequiométrica. A continuación, describiremos con más detalle la estructura del algoritmo.

Algoritmo

1. Inicialización: poner el número inicial de las especies químicas; colocar el valor de las constantes de reacción k_i , escoger el tiempo inicial t igual a cero y escoger el tiempo final T en el que terminara el algoritmo.
2. Calcular la función de propensión a_j y así como a_0 .
3. Generar dos números aleatorios r_1 y r_2 entre $(0,1)$.
4. Calcular τ de acuerdo con (3.77) y encontrar μ con ayuda de (3.78).
5. Actualizar el número de especies químicas de acuerdo a $X(t + \tau) \rightarrow X(t) + S_{\mu}$ y t como $t \rightarrow t + \tau$ y
6. Si $t < T$ entonces regresar al paso 2, en otro caso terminar.

Cuando se regresa al paso número 2 se recalculan las cantidades que se necesitan y se vuelven a generar los números aleatorios. Se puede notar que los números aleatorios τ y μ son generados de acuerdo con la densidad de probabilidad (3.74), para más detalles de los desarrollos ver [23].

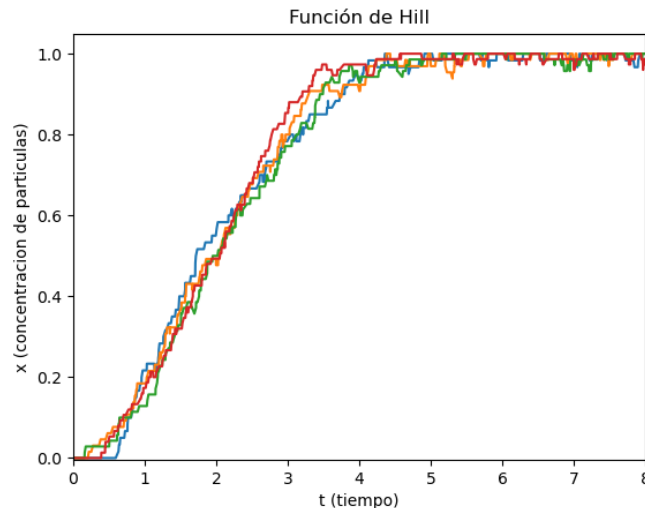


Figura 3.2: Proceso de Hill. Se modela la dinámica estocástica de las reacciones de donde se deriva la función de Hill.

Construimos un programa en Python asociado a este algoritmo (a lo largo del trabajo usamos este algoritmo debido a su sencillez). El programa es el siguiente (en este ejemplo en particular hemos usado al proceso de Hill, debido a que lo usaremos en el próximo capítulo), hemos

introducido una variación al algoritmo explicado, el cual consiste en que los datos solo se guarden después de un determinado tiempo t_1 esto con el objetivo de reducir el número de datos que se mantienen en la memoria RAM de la computadora, esto resulta particularmente beneficioso cuando el tamaño del sistema se vuelve grande. El archivo asociado a este programa se encuentra en la carpeta de Drive que mencionamos en la introducción, particularmente corresponde al archivo llamado `gillespiedt.py`.

Una observación sobre este programa es que uno solo define las matrices de los coeficientes estequiométricos, las condiciones iniciales, el tamaño del sistema, y ya este se encarga de realizar las simulaciones para cada una de las variables, por lo que es fácilmente aplicable a cualquier proceso químico o a algún sistema similar, además la mayoría de las estructuras usadas son tomadas directamente de la sección de procesos de vida y muerte multivariable. De este programa obtenemos la imagen 3.2 donde se describe la concentración de alguna especie química. En esta figura se nota claramente la presencia de fluctuaciones, y como la concentración va aumentando de manera discreta.

3.6.2. Primera reacción

Para este método se generan M números aleatorios r_i uniformemente distribuidos entre $(0,1)$, de aquí se generan τ_i tiempos de acuerdo con

$$\tau_i = \frac{1}{a_0} \ln \left(\frac{1}{r_i} \right), \quad (3.79)$$

entonces se escoge a τ como

$$\tau = \min\{\tau_1, \tau_2 \dots \tau_M\}, \quad (3.80)$$

se elige a la μ asociada al τ que se elige,

$$\mu = \text{sub índice de } \min\{\tau_1, \tau_2 \dots \tau_M\}. \quad (3.81)$$

Algoritmo

1. Inicialización: poner el número inicial de las especies químicas; colocar el valor de las constantes de reacción k_i , escoger el tiempo inicial t igual a cero y escoger el tiempo final T en el que terminara el algoritmo.
2. Calcular la función de propensión a_j y a_0 .
3. Generar M números aleatorios r_i entre $(0,1)$.
4. Calcular τ de acuerdo con (3.80) y encontrar μ con ayuda de (3.81).
5. Actualizar el número de especies químicas de acuerdo a $X(t + \tau) \rightarrow X(t) + S_\mu$ y t como $t \rightarrow t + \tau$
6. Si $t < T$ entonces regresar al paso 2, en otro caso terminar.

Cuando se regresa al paso 2 se vuelven a realizar de nuevo todos los cálculos con el número de especies químicas actualizadas. Para este caso no escribiremos un programa para calcular las reacciones químicas ya que para los propósitos de la tesis es suficiente el que hemos desarrollado en la sección anterior.

Existen otros métodos para simular los procesos estocásticos, muchas de estas son variantes del algoritmo de Gillespie [23], estas variantes se realizan de acuerdo a las necesidades y consideraciones del sistema a analizar, tal puede ser el caso de un sistema de muchas especies químicas o incluso considerar algunas distribución dependiente de la posición, de igual manera se pueden llegar a emplear métodos computacionales para resolver algunas aproximaciones, de hecho una muy usada es la ecuación de Langevin y la ecuación de Fokker-Planck.

Capítulo 4

Fluctuaciones intrínsecas en una red bioquímica estocástica

Debido a la naturaleza estocástica de algunos sistemas es indispensable cuantificar las fluctuaciones, así como determinar de qué tipo son (intrínsecas o extrínsecas), cuantificar las fluctuaciones estocásticas de una red de procesos bioquímicos en uno de los principales objetivos de este trabajo, por lo que en este capítulo desarrollaremos herramientas para esto. Primero discutiremos algunos conceptos relacionados a las fluctuaciones y después mostraremos algunas expresiones de como calcularla, encontramos una variante del teorema de fluctuación disipación derivado desde un formalismo diferente al que es comúnmente empleado, incluso nos aventuramos a calcular una para términos de orden superior. Posteriormente discutiremos sobre las diversas maneras de definir a las fluctuaciones intrínsecas. Además, discutiremos brevemente la importancia de las fluctuaciones en las redes genéticas. Terminaremos el capítulo con un análisis de la retroactividad con correcciones estocásticas. Todas las herramientas obtenidas en este capítulo nos ayudaran a construir un programa de Python que cuantifique las fluctuaciones de alguna red de procesos químicos, lo aplicaremos al analizar a las redes genéticas y al proceso de fosforilación en los próximos capítulos.

4.1. Ruido vs fluctuación

En la literatura el ruido y las fluctuaciones se tratan como conceptos prácticamente equivalentes, esto podría ser cierto en algunos casos, sin embargo, para evitar esta ambigüedad, nosotros definiremos de manera adecuada estos conceptos:

Ruido.

Esta cantidad es cuantificada a través de una función y principalmente se usa en sistemas que están descritos por ecuaciones deterministas (ecuaciones diferenciales), por lo general esta función es ajena al sistema y entra al sistema tal como se puede observar en la figura 4.1, uno de los objetivos de esta es introducir y cuantificar como el sistema puede ser afectado por condiciones externas tales como la temperatura.

Fluctuación.

Es una cantidad se cuantifica regularmente a partir de una distribución de probabilidad, por lo general se hace uso de la varianza y el promedio para definirla tal como en (4.4) o (4.6), esta cantidad nos habla sobre el cuanto se aleja una determinada variable de su promedio. Debido a la naturaleza estadística de las fluctuaciones esta se usa en los procesos estocásticos.

Dada las características que tienen cada uno de estos conceptos, en lo que sigue del trabajo optaremos por usar el termino de fluctuación al momento de cuantificar la desviación de las variables del promedio.

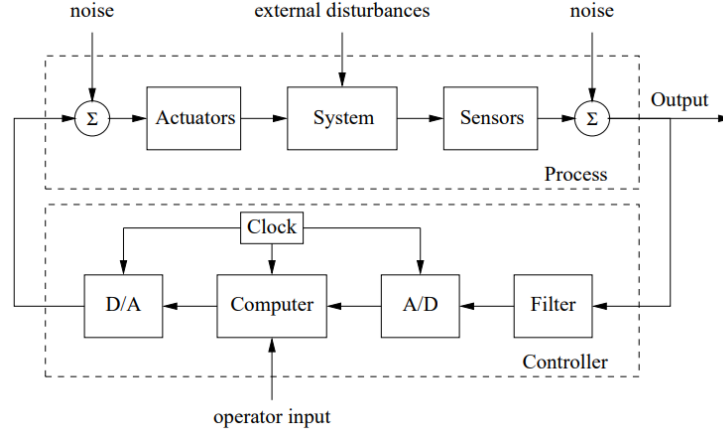


Figura 4.1: Ejemplo de un sistema con control [6].

4.2. Fluctuaciones

Ya hemos mostrado como la ecuación maestra se convierte en la determinista bajo ciertos criterios, también hemos explicado algunos métodos para solucionar la ecuación maestra (cuando esta es "casi lineal"), además de algunas técnicas computacionales para modelar los sistemas estocásticos, en el caso en el que hemos hecho la aproximación al límite determinista es deseable conocer las fluctuaciones que están presentes en el sistema. Cuando se encuentra una expresión explícita para la distribución de probabilidad de la ecuación maestra se vuelve mucho más sencillo encontrar las fluctuaciones, mientras que con las técnicas computacionales se tiene que inferir de las simulaciones.

Existen varias maneras de determinar las fluctuaciones en el sistema, pero estas siempre hacen uso de la varianza y el promedio. La varianza se usa como una cantidad que mide la dispersión de los valores con respecto de la media, la varianza es una muy buena manera de determinar la dispersión siempre y cuando no se tenga una distribución con una cola muy pesada, ya que es posible que no se capture correctamente la dispersión, pero para este trabajo la mayoría de las distribuciones serán del tipo Poissonianas o variantes de la binomial por lo que usaremos la varianza como medida de dispersión ya que captura muy bien esta información, la otra cantidad que se usa es el promedio, el promedio de una variables es equivalente a la que se obtiene de la solución determinista, tal como hemos explicado anteriormente, de hecho dentro de la varianza de alguna variable aleatoria $x(t)$ también aparecerá el promedio como en la siguiente relación

$$\sigma_{x(t)}^2 = \langle x(t)^2 \rangle - \langle x(t) \rangle^2, \quad (4.1)$$

para los procesos estocásticos, por lo regular se encuentra una ecuación diferencial para el promedio a partir de la ecuación maestra, tal como hemos explicado en la sección del límite determinista, donde por se encuentra algo parecido a la siguiente expresión

$$\frac{\partial \langle x(t) \rangle}{\partial t} = f_1(\langle x(t) \rangle), \quad (4.2)$$

para calcular la varianza nos apoyamos de la relación (4.1), introduciendo los términos dentro de la ecuación maestra y se obtendría algo similar a lo siguiente

$$\frac{\partial(\sigma_{x(t)}^2)}{\partial t} = f_2(\langle x(t) \rangle), \quad (4.3)$$

esta es una ecuación diferencial para la varianza, que por lo general es una expresión similar al teorema de fluctuación disipación (TFD) aunque en este caso para una sola variable, es importante tener una ecuación diferencial para la varianza debido a que en la aproximación que empleamos la varianza también es una cantidad que evoluciona en el tiempo. Ya tenemos ahora los elementos suficientes para definir las fluctuaciones, en este caso hay dos maneras [24] y otra propuesta que se usa mucho en la literatura:

i) **Factor de Fano.** En este caso se definen las fluctuaciones como

$$\eta^2 = \frac{\sigma^2}{\langle x \rangle}, \quad (4.4)$$

esta definición de la fluctuación tiene algunos inconvenientes cuando la distribución es de Poisson ya que el factor de Fano será igual a uno. En esta definición se usa como dispersión a la varianza y se normaliza con el promedio.

ii) **Alternativa I.** La otra forma de calcular las fluctuaciones es de la siguiente manera

$$\eta^2 = \frac{\sigma^2}{\langle x \rangle^2}, \quad (4.5)$$

A lo largo del texto nosotros usaremos esta última opción ya que es mucho más ventajosa y clara para los modelos que analizaremos. Lo que nos dice esta fórmula es que se a la dispersión que en este caso es la varianza se divide entre el promedio al cuadrado (muy similar a lo que se hace en el cálculo de errores), de esta manera tenemos una cantidad adimensional que cuantifica las fluctuaciones.

iii) **Alternativa II.** Tal como hemos derivado en el capítulo de análisis estocástico hemos definido a las fluctuaciones del sistema como a continuación

$$\eta^2 = \frac{\sigma^2}{\Omega^2}, \quad (4.6)$$

esta definición esta más acorde a lo que se emplea en la literatura y es derivable a partir de la forma en la que obtuvimos el límite determinista y no solo es una propuesta como los otros dos casos anteriores, por lo que tiene más fundamento teórico.

Hay que recordar que en las tres definiciones de las fluctuaciones no se calcula directamente la fluctuación sino el promedio de la magnitud de esta. Las primeras dos propuestas para la definición de fluctuación se usan en el sentido de querer normalizar esas cantidades, mientras que la última la encontramos a partir del límite determinista que discutimos en el capítulo anterior.

Como estamos particularmente interesados en los sistemas de vida y muerte multivariable, calcularemos las fluctuaciones presentes para este sistema, donde nos encontraremos con una expresión similar al teorema de fluctuación-disipación, que es calculada a directamente de la ecuación maestra. Recordamos que ya hemos calculado la ecuación diferencial para el promedio la cual es

$$\frac{\partial(\langle S_j \rangle)}{\partial t} = \Omega \sum_i \Gamma_{ji} \langle t_i^+(S_l) - t_i^-(S_l) \rangle, \quad (4.7)$$

ahora calcularemos una ecuación diferencial para la varianza, para ello primero calculamos el promedio de $\langle S_{l_1} S_{l_2} \rangle$, entonces multiplicamos la ecuación maestra por las variables $S_{l_1} S_{l_2}$,

$$\begin{aligned} \frac{\partial}{\partial t} (S_{l_1} S_{l_2} P(S_j, t)) &= \Omega S_{l_1} S_{l_2} \sum_i (t_i^-(S_j + \Gamma_{ij}) P(S_j + \Gamma_{ij}, t) - t_i^+(S_j) P(S_j, t) \\ &\quad + t_i^+(S_j - \Gamma_{ij}) P(S_j - \Gamma_{ij}, t) - t_i^-(S_j) P(S_j, t)), \end{aligned} \quad (4.8)$$

sumamos sobre todos los estados de la variable S_j

$$\begin{aligned} \frac{\partial}{\partial t} \left(\sum_{S_j} S_{l_1} S_{l_2} P(S_j, t) \right) &= \Omega \sum_{S_j} S_{l_1} S_{l_2} \sum_i (t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+(S_j) P(S_j, t) \\ &\quad + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^-(S_j) P(S_j, t)), \end{aligned} \quad (4.9)$$

reordenando términos e insertando la definición del promedio de una función obtenemos

$$\begin{aligned} \frac{\partial}{\partial t} (\langle S_{l_1} S_{l_2} \rangle) &= \Omega \sum_{S_j} \sum_i ((S_{l_1} - \Gamma_{l_1 i})(S_{l_2} - \Gamma_{l_2 i}) t_i^-(S_j) P(S_j, t) - S_{l_1} S_{l_2} t_i^+(S_j) P(S_j, t) \\ &\quad + (S_{l_1} + \Gamma_{l_1 i})(S_{l_2} + \Gamma_{l_2 i}) t_i^+(S_j) P(S_j, t) - S_{l_1} S_{l_2} t_i^-(S_j) P(S_j, t)), \end{aligned} \quad (4.10)$$

simplificando

$$\frac{\partial}{\partial t} (\langle S_{l_1} S_{l_2} \rangle) = \Omega \sum_i (\langle (S_{l_1} \Gamma_{l_2 i} + S_{l_2} \Gamma_{l_1 i})(t_i^+(S_j) - t_i^-(S_j)) \rangle + \langle (\Gamma_{l_1 i} \Gamma_{l_2 i})(t_i^+(S_j) + t_i^-(S_j)) \rangle). \quad (4.11)$$

Recordar que la covarianza está definida como $\sigma^2(S_{l_1}, S_{l_2}) = \langle S_{l_1} S_{l_2} \rangle - \langle S_{l_1} \rangle \langle S_{l_2} \rangle$, por lo que con ayuda de la expresión anterior la derivada temporal de la covarianza se convierte en

$$\begin{aligned} \frac{\partial}{\partial t} \sigma^2(S_{l_1}, S_{l_2}) &= \Omega \sum_i (\langle (S_{l_1} \Gamma_{l_2 i} + S_{l_2} \Gamma_{l_1 i})(t_i^+(S_j) - t_i^-(S_j)) \rangle + \langle (\Gamma_{l_1 i} \Gamma_{l_2 i})(t_i^+(S_j) + t_i^-(S_j)) \rangle \\ &\quad - \langle (S_{l_1} \Gamma_{l_2 i} + S_{l_2} \Gamma_{l_1 i}) \rangle \langle (t_i^+(S_j) - t_i^-(S_j)) \rangle), \end{aligned} \quad (4.12)$$

cuando tenemos el promedio de una función podemos expandirla hasta el segundo momento (las funciones que nos aparecen son polinomiales) de la siguiente manera

$$\begin{aligned} \langle f(X_l) \rangle &\approx f(\langle X_l \rangle) + \sum_i \sum_j \frac{\sigma^2(X_i, X_j)}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle}, \\ \langle X_k f(X_l) \rangle &\approx \langle X_k \rangle \langle f(X_l) \rangle + \sum_i \sigma^2(X_i, X_k) \frac{\partial f(X_l)}{\partial X_i} \Big|_{X_l = \langle X_l \rangle}, \end{aligned} \quad (4.13)$$

(para más detalles de cómo se toma el límite determinista de las tasas de propensión ver el capítulo 3 de análisis estocástico) usando esto en la relación anterior ahora tendremos lo siguiente

$$\begin{aligned} \frac{\partial}{\partial t} \sigma^2(S_{l_1}, S_{l_2}) &= \Omega \sum_i (\Gamma_{l_1 i} \Gamma_{l_2 i} (t_i^+(\langle S_j \rangle) + t_i^-(\langle S_j \rangle)) \\ &\quad + \sum_{j_1} \left(\left(\sigma^2(S_{l_1}, S_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial \langle S_{j_1} \rangle} + \sigma^2(S_{j_1}, S_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial \langle S_{j_1} \rangle} \right) (t_i^+(\langle S_j \rangle) - t_i^-(\langle S_j \rangle)) \right) \\ &\quad + \sum_{j_1} \sum_{j_2} \left(\frac{\sigma^2(S_{j_1}, S_{j_2}) \Gamma_{l_1 i} \Gamma_{l_2 i}}{2} \frac{\partial^2}{\partial \langle S_{j_1} \rangle \partial \langle S_{j_2} \rangle} (t_i^+(\langle S_j \rangle) + t_i^-(\langle S_j \rangle)) \right). \end{aligned} \quad (4.14)$$

Este resultado de hecho una variante del TFD, algo que por lo regular se obtiene de la aproximación lineal de ruido o la ecuación de Fokker-Planck [24] (para una derivación del TFD usual ver apéndice A), o a través de aproximaciones similares a la nuestra [25] aunque en este solo analizan

hasta reacciones de segundo orden, por lo que nuestro resultado es más general. Nosotros hemos mostrado otra forma de obtenerlo sin tener que realizar muchos desarrollos y/o suposiciones como habitualmente se hace en las otras aproximaciones, para que su estructura se asemeje más al TFD usual primero escribimos a la relación anterior en términos de las variables deterministas,

$$\begin{aligned} \frac{\partial}{\partial t} \frac{\sigma^2(S_{l_1}, S_{l_2})}{\Omega^2} &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D^+}(s_j) + R_i^{D^-}(s_j))}{\Omega} \right. \\ &+ \sum_{j_1} \left(\frac{\sigma^2(S_{l_1}, S_{j_1})}{\Omega^2} \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \frac{\sigma^2(S_{j_1}, S_{l_2})}{\Omega^2} \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D^+}(s_j) - R_i^{D^-}(s_j)) \\ &\left. + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(S_{j_1}, S_{j_2})}{\Omega^2} \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D^+}(s_j) + R_i^{D^-}(s_j)) \right), \end{aligned} \quad (4.15)$$

notar que $t_i^\pm(\langle S_j \rangle) \approx R_i^{D^\pm}(s_j)$, donde $R_i^{D^\pm}(s_j)$ son las tasas deterministas, ahora definimos lo siguiente

$$\begin{aligned} \sigma^2 &= \frac{\sigma^2(S_{l_1}, S_{l_2})}{\Omega^2}, & \mathbf{A} &= \sum_i \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} (R_i^{D^+}(s_j) - R_i^{D^-}(s_j)), \\ \mathbf{B} &= \sum_i \Gamma_{l_1 i} \Gamma_{l_2 i} (R_i^{D^+}(s_j) + R_i^{D^-}(s_j)), & \Delta \mathbf{B} &= \sum_i \Gamma_{l_1 i} \Gamma_{l_2 i} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D^+}(s_j) + R_i^{D^-}(s_j)), \end{aligned} \quad (4.16)$$

introduciendo estos términos en (4.15) recuperamos el TFD usual más un término extra

$$\frac{\partial}{\partial t} \sigma^2 = \frac{\mathbf{B}}{\Omega} + \mathbf{A} \sigma^2 + \sigma^2 \mathbf{A}^T + \sigma^2 \frac{1}{2\Omega} \Delta \mathbf{B}, \quad (4.17)$$

por las propiedades de $\Delta \mathbf{B}$ y σ podemos reescribirla definiendo $\mathbf{A}' = \mathbf{A} + \frac{1}{4\Omega} \Delta \mathbf{B}$

$$\frac{\partial}{\partial t} \sigma^2 = \frac{\mathbf{B}}{\Omega} + \mathbf{A}' \sigma^2 + \sigma^2 \mathbf{A}'^T, \quad (4.18)$$

el resultado que hemos obtenido se puede considerar como una corrección al TFD usual, debido a que aparece el término $\Delta \mathbf{B}$, cuando las tasas en el sistema sean lineales este término desaparece del sistema, entonces podría pasar desapercibido, y como depende del inverso de Ω su contribución se vuelve pequeña, por lo que en sistemas que tengan tasas no lineales y sean pequeños empieza a cobrar importancia, se puede realizar una expansión hasta órdenes superiores pero esto lo dejamos para la siguiente sección. En el caso estacionario tendremos lo siguiente

$$\frac{\mathbf{B}}{\Omega} = -\mathbf{A}' \sigma^2 - \sigma^2 \mathbf{A}'^T. \quad (4.19)$$

Debido a que aparece explícitamente Ω en nuestra ecuación podemos tomar el límite cuando esta tienda a ser muy grande entonces el sistema ya no presentaría fluctuaciones en el estado estacionario. De esta ecuación uno puede interpretar que \mathbf{B} es una especie de fuente para las fluctuaciones. En el caso en cual se obtenga una expresión explícita para la distribución, se calcula el promedio y la varianza, después de sustituyen estas cantidades dentro de la expresión (4.6) aunque en el caso multivariable no es nada simple encontrar la distribución en el estado estacionario (tal como hemos presentado anteriormente esto es posible solo para sistemas débilmente reversibles), en el caso en el que se esté usando un sistema computacional, emplearemos la misma fórmula, pero aquí los promedios y varianzas se calcularán en el entorno del algoritmo.

Fluctuaciones intrínsecas en una red bioquímica estocástica
4.3 Ley de acción de masas hasta el tercer momento

Dada las fórmulas que hemos obtenido se puede modelar la dinámica de los sistemas estocásticos a partir de las siguientes ecuaciones diferenciales

$$\begin{aligned} \frac{\partial s_j}{\partial t} &= \sum_i \Gamma_{ji} \left(R_i^{D+}(s_l) - R_i^{D-}(s_l) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \\ \frac{\partial}{\partial t} \sigma^2(s_{l_1}, s_{l_2}) &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D+}(s_j) + R_i^{D-}(s_j))}{\Omega} \right. \\ &\quad + \sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_j) - R_i^{D-}(s_j)) \\ &\quad \left. + \sum_{j_1} \sum_{j_2} \sigma^2(s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_j) + R_i^{D-}(s_j)) \right), \end{aligned} \quad (4.20)$$

($\sigma^2(s_{j_1}, s_{j_2}) = \frac{\sigma^2(s_{l_1}, s_{l_2})}{\Omega^2}$ denota a la covarianza de las variables deterministas, así como $R_i^{D\pm}(s_j)$ denotan a las tasas deterministas) el número de ecuaciones diferenciales en un sistema que tenga n diferentes especies bioquímicas independientes será $\frac{n^2+3n}{2}$, debido a la no linealidad de este sistema de ecuaciones se llegan a emplear métodos numéricos para resolverla, en la parte final de este documento emplearemos esta ecuación para modelar algunos de los sistemas estocásticos, debido a que es computacionalmente más eficiente que emplear directamente el algoritmo de Gillespie. Hay que recordar que los términos diagonales de la covarianza son el cuadrado de la magnitud de las fluctuaciones, por lo que de ahí cuantificaremos las fluctuaciones.

4.3. Ley de acción de masas hasta el tercer momento

Hemos usado una aproximación ligeramente diferente a la usual para determinar el límite determinista, en donde nos hemos apoyado directamente de un desarrollo en serie, aunque hasta ahora solo hemos hecho esta aproximación hasta el segundo momento central, con el objetivo de conocer que tanto contribuyen términos de orden superior al sistema, nos daremos a la tarea de encontrar una expresión similar al del TFD pero considerando términos de orden superior, por lo que primero definiremos al tercer momento central como $\zeta^3(X_i, X_j, X_k) = -\langle (X_i - \langle X_i \rangle) \langle (X_j - \langle X_j \rangle) \langle (X_k - \langle X_k \rangle) \rangle \rangle$, para los desarrollos que necesitamos usaremos la siguiente relación que es fácilmente demostrable

$$\begin{aligned} \zeta^3(X_i, X_j, X_k) &= -\langle X_i \rangle \sigma^2(X_j, X_k) - \langle X_k \rangle \sigma^2(X_i, X_j) - \langle X_j \rangle \sigma^2(X_k, X_i) - \langle X_i \rangle \langle X_j \rangle \langle X_k \rangle + \langle X_i X_j X_k \rangle, \\ &= -\epsilon_{ijk}^+ \langle X_i \rangle \sigma^2(X_j, X_k) - \langle X_i \rangle \langle X_j \rangle \langle X_k \rangle + \langle X_i X_j X_k \rangle \end{aligned} \quad (4.21)$$

hemos usado un término parecido al tensor de Levi-Civita [26] para simplificar a nuestro sistema de tal manera que absorbe tres términos, para considerar a ζ^3 en la dinámica del sistema trabajaremos con las siguientes aproximaciones

$$\begin{aligned} \langle f(X_l) \rangle &\approx f(\langle X_l \rangle) + \sum_i \sum_j \frac{\sigma^2(X_i, X_j)}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle} + \sum_i \sum_j \sum_k \frac{\zeta^3(X_i, X_j, X_k)}{3!} \frac{\partial^3 f(X_l)}{\partial X_i \partial X_j \partial X_k} \Big|_{X_l = \langle X_l \rangle}, \\ \langle X_s f(X_l) \rangle &\approx \langle X_s \rangle \langle f(X_l) \rangle + \sum_i \sigma^2(X_i, X_s) \frac{\partial f(X_l)}{\partial X_i} \Big|_{X_l = \langle X_l \rangle} + \sum_i \sum_j \frac{\zeta^3(X_i, X_j, X_s)}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle}, \end{aligned}$$

$$\begin{aligned}
 \langle X_s X_r f(X_l) \rangle &\approx \langle X_s X_r \rangle \langle f(X_l) \rangle - \sigma^2(X_s, X_r) \sum_i \sum_j \frac{\sigma^2(X_i, X_j)}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l=\langle X_l \rangle} \\
 &+ \sum_i \sigma^2(X_r, X_i) \langle X_s \rangle \frac{\partial f(X_l)}{\partial X_i} \Big|_{X_l=\langle X_l \rangle} + \sum_i \sigma^2(X_s, X_i) \langle X_r \rangle \frac{\partial f(X_l)}{\partial X_i} \Big|_{X_l=\langle X_l \rangle} \\
 &- \sigma^2(X_s, X_r) \sum_i \sum_j \sum_k \frac{\zeta^3(X_i, X_j, X_k)}{3!} \frac{\partial^3 f(X_l)}{\partial X_i \partial X_j \partial X_k} \Big|_{X_l=\langle X_l \rangle} + \sum_i \zeta^3(X_r, X_s, X_i) \frac{\partial f(X_l)}{\partial X_i} \Big|_{X_l=\langle X_l \rangle} \\
 &+ \sum_i \sum_j \left(\frac{\zeta^3(X_r, X_i, X_j)}{2} \langle X_s \rangle + \frac{\zeta^3(X_s, X_i, X_j)}{2} \langle X_r \rangle \right) \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l=\langle X_l \rangle}, \quad (4.22)
 \end{aligned}$$

(se han tachado dos términos porque su contribución es despreciable) usaremos estas aproximaciones para tener una corrección a tercer momento de la ley de acción de masas y del TFD, de donde obtenemos

$$\begin{aligned}
 \frac{\partial s_j}{\partial t} &= \sum_i \Gamma_{ji} \left(R_i^{D+}(s_l) - R_i^{D-}(s_l) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right. \\
 &\quad \left. + \sum_{j_1} \sum_{j_2} \sum_{j_3} \frac{\zeta^3(s_{j_1}, s_{j_2}, s_{j_3})}{3!} \frac{\partial^3}{\partial s_{j_1} \partial s_{j_2} \partial s_{j_3}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \\
 \frac{\partial}{\partial t} \sigma^2(s_{l_1}, s_{l_2}) &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D+}(s_l) + R_i^{D-}(s_l))}{\Omega} + \sum_{j_1} \sum_{j_2} \sigma^2(s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) + R_i^{D-}(s_l)) \right. \\
 &\quad + \sum_{j_1} \sum_{j_2} \sum_{j_3} \zeta^3(s_{j_1}, s_{j_2}, s_{j_3}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{3! \Omega} \frac{\partial^3}{\partial s_{j_1} \partial s_{j_2} \partial s_{j_3}} (R_i^{D+}(s_l) + R_i^{D-}(s_l)) \\
 &\quad + \sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \\
 &\quad \left. + \sum_{j_1} \sum_{j_2} \left(\zeta^3(s_{l_1}, s_{j_1}, s_{j_2}) \frac{\Gamma_{l_2 i}}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} + \zeta^3(s_{l_2}, s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i}}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} \right) (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \quad (4.23)
 \end{aligned}$$

donde $R_i^{D\pm}(s_j)$ son las tasas deterministas, estas nuevas ecuaciones diferenciales consideran el tercer momento central, además, el término que acompaña a este tercer momento son derivadas de orden 3 y 2 respectivamente, por lo que cobra relevancia solo para tasas que sean no-lineales y/o de orden superior a 2 y 3. Ahora lo que sigue es encontrar también una ecuación diferencial para este tercer momento, para ello primero calculamos el promedio $\langle S_{l_1} S_{l_2} S_{l_3} \rangle$, para ello multiplicamos la ecuación maestra por las variables $S_{l_1} S_{l_2} S_{l_3}$,

$$\begin{aligned}
 \frac{\partial}{\partial t} (S_{l_1} S_{l_2} S_{l_3} P(S_j, t)) &= \Omega S_{l_1} S_{l_2} S_{l_3} \sum_i (t_i^-(S_j + \Gamma_{ij}) P(S_j + \Gamma_{ij}, t) - t_i^+(S_j) P(S_j, t) \\
 &\quad + t_i^+(S_j - \Gamma_{ij}) P(S_j - \Gamma_{ij}, t) - t_i^-(S_j) P(S_j, t)), \quad (4.24)
 \end{aligned}$$

sumamos sobre todos los estados de la variable S_j

$$\begin{aligned}
 \frac{\partial}{\partial t} \left(\sum_{S_j} S_{l_1} S_{l_2} S_{l_3} P(S_j, t) \right) &= \Omega \sum_{S_j} S_{l_1} S_{l_2} S_{l_3} \sum_i (t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+(S_j) P(S_j, t) \\
 &\quad + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^-(S_j) P(S_j, t)), \quad (4.25)
 \end{aligned}$$

reordenando términos y sacando el promedio de los dos lados obtenemos

$$\begin{aligned}
 \frac{\partial}{\partial t} (\langle S_{l_1} S_{l_2} S_{l_3} \rangle) &= \Omega \sum_{S_j} \sum_i ((S_{l_1} - \Gamma_{l_1 i})(S_{l_2} - \Gamma_{l_2 i})(S_{l_3} - \Gamma_{l_3 i}) t_i^-(S_j) P(S_j, t) - S_{l_1} S_{l_2} S_{l_3} t_i^+(S_j) P(S_j, t) \\
 &\quad + (S_{l_1} + \Gamma_{l_1 i})(S_{l_2} + \Gamma_{l_2 i})(S_{l_3} - \Gamma_{l_3 i}) t_i^+(S_j) P(S_j, t) - S_{l_1} S_{l_2} S_{l_3} t_i^-(S_j) P(S_j, t)), \quad (4.26)
 \end{aligned}$$

simplificando

$$\begin{aligned}
\frac{\partial}{\partial t} \langle (S_{l_1} S_{l_2} S_{l_3}) \rangle &= \Omega \sum_i \left(\langle (S_{l_1} \Gamma_{l_2 i} \Gamma_{l_3 i} + S_{l_2} \Gamma_{l_1 i} \Gamma_{l_3 i} + S_{l_3} \Gamma_{l_1 i} \Gamma_{l_2 i}) (t_i^+(S_j) + t_i^-(S_j)) \rangle \right. \\
&\quad + \langle (\Gamma_{l_1 i} S_{l_2} S_{l_3} + \Gamma_{l_3 i} S_{l_1} S_{l_2} + \Gamma_{l_2 i} S_{l_3} S_{l_1} + \Gamma_{l_1 i} \Gamma_{l_2 i} \Gamma_{l_3 i}) (t_i^+(S_j) - t_i^-(S_j)) \rangle \left. \right) \\
&= \Omega \sum_i \left(\langle (\epsilon_{l_1 l_2 l_3}^+ S_{l_1} \Gamma_{l_2 i} \Gamma_{l_3 i}) (t_i^+(S_j) + t_i^-(S_j)) \rangle \right. \\
&\quad \left. + \langle (\epsilon_{l_1 l_2 l_3}^+ \Gamma_{l_1 i} S_{l_2} S_{l_3} + \Gamma_{l_1 i} \Gamma_{l_2 i} \Gamma_{l_3 i}) (t_i^+(S_j) - t_i^-(S_j)) \rangle \right). \tag{4.27}
\end{aligned}$$

Recordando la igualdad para el tercer momento que derivamos al principio de esta sección (4.21) y con ayuda de la expresión anterior, la derivada temporal del tercer momento se convierte en

$$\begin{aligned}
\frac{\partial}{\partial t} \zeta^3(S_{l_1}, S_{l_2}, S_{l_3}) &= \Omega \sum_i \left(\langle (\epsilon_{l_1 l_2 l_3}^+ S_{l_1} \Gamma_{l_2 i} \Gamma_{l_3 i}) (t_i^+(S_j) + t_i^-(S_j)) \rangle \right. \\
&\quad + \langle (\epsilon_{l_1 l_2 l_3}^+ \Gamma_{l_1 i} S_{l_2} S_{l_3} + \Gamma_{l_1 i} \Gamma_{l_2 i} \Gamma_{l_3 i}) (t_i^+(S_j) - t_i^-(S_j)) \rangle \\
&\quad - \epsilon_{l_1 l_2 l_3}^+ \langle S_{l_3} \rangle \left(\langle (S_{l_1} \Gamma_{l_2 i} + S_{l_2} \Gamma_{l_1 i}) (t_i^+(S_j) - t_i^-(S_j)) \rangle + \langle (\Gamma_{l_1 i} \Gamma_{l_2 i}) (t_i^+(S_j) + t_i^-(S_j)) \rangle \right) \\
&\quad - \langle (S_{l_1} \Gamma_{l_2 i} + S_{l_2} \Gamma_{l_1 i}) \rangle \langle (t_i^+(S_j) - t_i^-(S_j)) \rangle - \epsilon_{l_1 l_2 l_3}^+ \sigma^2(S_{l_1}, S_{l_2}) \Gamma_{l_3 i} \langle (t_i^+(S_j) - t_i^-(S_j)) \rangle \\
&\quad \left. - \epsilon_{l_1 l_2 l_3}^+ \langle S_{l_1} \rangle \langle S_{l_2} \rangle \Gamma_{l_3 i} \langle (t_i^+(S_j) - t_i^-(S_j)) \rangle \right), \tag{4.28}
\end{aligned}$$

simplificando, usando (4.22) e introduciendo las variables deterministas $s_j = \frac{\langle S_j \rangle}{\Omega}$, obtenemos

$$\begin{aligned}
\frac{\partial}{\partial t} \zeta^3(s_{l_1}, s_{l_2}, s_{l_3}) &= \epsilon_{l_1 l_2 l_3}^+ \sum_i \left(\frac{1}{3\Omega^2} \Gamma_{l_1 i} \Gamma_{l_2 i} \Gamma_{l_3 i} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right. \\
&\quad + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \\
&\quad + \sum_{j_1} \sum_{j_2} \sum_{j_3} \frac{\zeta^3(s_{j_1}, s_{j_2}, s_{j_3})}{3!} \frac{\partial^3}{\partial s_{j_1} \partial s_{j_2} \partial s_{j_3}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \left. \right) \\
&\quad + \frac{1}{\Omega} \left(\sum_{j_1} \left(\sigma^2(s_{j_1}, s_{l_1}) \Gamma_{l_2 i} \Gamma_{l_3 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_l) + R_i^{D-}(s_l)) \right. \\
&\quad + \sum_{j_1} \sum_{j_2} \left(\zeta^3(s_{l_1}, s_{j_1}, s_{j_2}) \frac{\Gamma_{l_2 i} \Gamma_{l_3 i}}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} \right) (R_i^{D+}(s_l) + R_i^{D-}(s_l)) \left. \right) \\
&\quad - s_{l_3} \left(\sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right. \\
&\quad + \sum_{j_1} \sum_{j_2} \left(\zeta^3(s_{l_1}, s_{j_1}, s_{j_2}) \frac{\Gamma_{l_2 i}}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} + \zeta^3(s_{j_1}, s_{j_2}, s_{l_2}) \frac{\Gamma_{l_1 i}}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} \right) (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \left. \right) \\
&\quad + \Gamma_{l_1 i} \left(\sum_{j_1} \left(\sigma^2(s_{l_2}, s_{j_1}) s_{l_3} + \sigma^2(s_{l_3}, s_{j_1}) s_{l_2} + \zeta^3(s_{l_2}, s_{l_3}, s_{j_1}) \right) \frac{\partial}{\partial s_{j_1}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right. \\
&\quad \left. + \sum_{j_1} \sum_{j_2} \left(\frac{\zeta^3(s_{l_2}, s_{j_1}, s_{j_2})}{2} s_{l_3} + \frac{\zeta^3(s_{l_3}, s_{j_1}, s_{j_2})}{2} s_{l_2} \right) \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right) \left. \right). \tag{4.29}
\end{aligned}$$

Por lo que para modelar a sistemas considerando correcciones estocásticas podemos emplear a (4.23) y (4.29) como correcciones estocásticas al sistema determinista, en esta aproximación hemos usado hasta el tercer momento, sin embargo como se puede observar el tercer momento central ya depende del inverso de Ω^2 , por lo que su contribución a la dinámica del sistema es practicante despreciable para la primera ecuación de (4.23), mientras que para el TFD el tercer momento central desaparecería si el sistema ya no es lineal, y una de las razones por la cual se omite su uso. Usar (4.29) hace que en nuestro sistema aparezcan $\frac{n^3+2n}{3}$ nuevas ecuaciones diferenciales

independientes, por lo que ahora tendremos un sistema con $\frac{n^2+3n}{2} + \frac{n^3+2n}{3}$ ecuaciones diferenciales independientes, podemos decir que esta nueva ecuación diferencial corresponde a una variante a segundo orden del TFD.

4.4. Importancia de las fluctuaciones en los circuitos genéticos

Con ayuda de los modelos estocásticos es posible visualizar y cuantificar las fluctuaciones de los circuitos genéticos, esto con ayuda de las expresiones de la sección anterior, sin embargo, en este punto es importante preguntarse cuál es el papel de esta en tales sistemas, su importancia y consecuencias en la dinámica del sistema, algunas de estas cuestiones las iremos respondiendo a continuación, para ello nos apoyaremos de [27, 28, 29], se pueden analizar las fluctuaciones en tres niveles dentro de los organismos:

1. Fisiológico

Las fluctuaciones pueden permitir ciertos mecanismos de regulación fisiológica útiles, como la coordinación de la expresión de un gran conjunto de genes.

- a) **Fluctuaciones en los circuitos genéticos:** Debido a que el número de elementos bioquímicos involucrados es pequeño, las fluctuaciones son parte fundamental de la dinámica del sistema.
- b) **Fluctuaciones y regulación en los circuitos genéticos:** Aunque por lo general se encuentra que las fluctuaciones alteran la dinámica del sistema, es posible encontrar sistemas en el cual sucede lo contrario.

2. Nivel poblacional

Las fluctuaciones permiten una amplia gama de estrategias de diferenciación probabilística de organismos microbianos a pluricelulares. **Diferenciación probabilística:** una de las ventajas funcionales clave de las fluctuaciones es su capacidad para permitir la diferenciación probabilística de células que de otro modo serían idénticas. Estos sistemas tienden a clasificarse de acuerdo con sus características dinámicas, algunos ejemplos son:

- a) Los sistemas estocásticos de conmutación de estado cambian entre estados metaestables y a menudo se basan en bucles de retroalimentación positiva.
- b) Los sistemas de diferenciación excitable activados por fluctuaciones permiten a las células entrar probabilísticamente en un estado, pero vuelven al estado original después de un tiempo definido. Estos sistemas utilizan una combinación de bucles de retroalimentación positiva y negativa.
- c) En los sistemas de diferenciación procrastinadora, las células individuales construyen de manera gradual y variablemente el nivel de un regulador clave para generar una amplia distribución de retrasos antes de comprometerse con uno nuevo, marcadamente diferente.

3. Evolución

Las fluctuaciones pueden facilitar la adaptación y el desarrollo evolutivo. Las fluctuaciones también pueden tener un papel integral en la evolución al ampliar la gama de fenotipos que pueden resultar de un genotipo dado.

- a) **Evolución y fluctuación:** Muchas mutaciones causan cambios cualitativos en el desarrollo, pero lo hacen solo en una fracción de los individuos, incluso en una población isogénica. Cuando la señalización se ve atenuada por la mutación, las fluctuaciones permiten secuencias de eventos alternativas y, por lo tanto, fenotipos alternativos en algunas células.

Estos son solo algunos de los roles de las fluctuaciones en los circuitos genéticos, para más detalles de cada uno de estos puntos ver las referencias mencionadas, existen otros usos de estas en las células que aún están siendo investigadas, pero con tan solo las anteriores nos damos cuenta de que estos llegan a jugar papeles claves en el proceso de transcripción/traducción.

4.5. Fluctuaciones extrínsecas e intrínsecas

Las autocorrelaciones resumen convenientemente tanto la magnitud como la frecuencia de las fluctuaciones y son bastante manejables analíticamente. Una manera de hacerlo es en términos de la varianza sobre el promedio al cuadrado, lo que permite una separación limpia de las diferentes fuentes de ruido siempre que los modelos sean débilmente no lineales. Otra medida común es el factor de Fano, la varianza sobre el promedio, que es igual a uno para las distribuciones de Poisson. Pero la comparación con el Poissoniano solo funciona bien para procesos aleatorios discretos univariados, donde la varianza a menudo es proporcional al promedio con una constante de proporcionalidad que refleja la naturaleza general del proceso. Para los procesos aleatorios multivariados, la distribución de Poisson no tiene una posición especial y el uso del factor Fano puede ser engañoso. Es posible que ninguna medida funciona bien para todos los tipos de fluctuaciones; las fluctuaciones espontáneas dependen del número de moléculas, mientras que las fluctuaciones forzadas no. Sin embargo, la varianza sobre el promedio al cuadrado es sin duda una base más adecuada para las interpretaciones experimentales. Primero, en la mayoría de los estudios experimentales realizados hasta ahora, el número promedio de proteínas por célula es demasiado alto para contribuir a fluctuaciones espontáneas sustanciales. En segundo lugar, al graficar la varianza sobre el promedio al cuadrado como una función del promedio inverso, cualquier comportamiento de escala univariante se identifica fácilmente sin introducir problemas de escala para ningún ruido extrínseco. En tercer lugar, la relevancia de una fluctuación depende típicamente del tamaño normal del sistema. Debido a que la varianza es un momento de segundo orden, debe normalizarse mediante el promedio al cuadrado [30].

A continuación, describiremos algunas maneras en la cual se diferencian las fluctuaciones intrínsecas de las extrínsecas [30]:

- 1.- **Naturaleza estadística de las fluctuaciones.** Este punto de vista es desde una perspectiva física, aquí el definir intrínseco y extrínseco no es otra cosa que el definir el adentro y afuera de un sistema, por lo que dependiendo de que consideremos adentro y afuera etiquetaremos a las fluctuaciones. Por ejemplo, para un MTT si decimos que la proteína es nuestro sistema (o lo único que se puede medir), entonces las fluctuaciones debidas a la proteína son las intrínsecas y las fluctuaciones debido a los otros elementos serán las extrínsecas.
- 2.- **Que tan importante es un proceso.** Este punto de vista es desde una perspectiva biológica, aquí uno se pregunta que tan importante es un componente, y en base a esto definir las fluctuaciones intrínsecas y extrínsecas. Por ejemplo, podemos decir que el ARm es lo importante en el sistema, entonces las fluctuaciones debidas al ARm son las intrínsecas y las fluctuaciones debido a los otros elementos serán las extrínsecas.
- 3.- **Correlaciones entre diferentes elementos.** En el caso del MTT uno puede ver la correlación entre diferentes proteínas, por lo que pueden existir proteínas que afecten a muchos genes, mientras que otros afectan a pocos o a uno, y dependiendo de esto se realiza la clasificación, con estas ideas las fluctuaciones del ARm serían intrínsecas.
- 4.- **A parir de experimentos.** Una estrategia particularmente inteligente para separar las fuentes de las fluctuaciones se basa en las correlaciones entre la expresión de dos genes con indicadores fluorescentes, donde estos están físicamente separados pero regulados de manera idéntica. [31]

Uno de los objetivos del trabajo es cuantificar las fluctuaciones en las redes genéticas, específicamente determinar las fluctuaciones intrínsecas, y como hemos explicado en los 4 puntos anteriores el decidir que es una fluctuación intrínseca es un tanto arbitraria, por lo que optaremos por la siguiente definición

Fluctuación intrínseca: estas son determinadas por la estructura, las tasas de reacción, y la concentración de especies de las redes bioquímicas subyacentes [28].

Dado que uno de los objetivos de este trabajo es cuantificar las fluctuaciones intrínsecas en nuestro sistema, para ello usaremos la expresión $\eta^2 = \frac{\sigma^2}{\Omega^2}$, usamos esta expresión debido a que va en concordancia con la definición de fluctuación intrínseca que hemos elegido.

4.6. Cuantificar fluctuaciones

En esta sección nos dedicaremos a analizar una red genética estocástica a través de herramientas computacionales, esto con el fin de entender cómo funcionan y como son los resultados a partir del algoritmo de Gillespie y la aproximación determinista y la variante del teorema de fluctuación-disipación (TFD), y así aplicar los programas a los sistemas que analizaremos en los últimos capítulos.

Para el caso en el que empleemos el algoritmo de Gillespie, haremos un ensamble y sobre este determinaremos el promedio de la concentración de las especies bioquímicas involucradas, así como las fluctuaciones de estas, nos apoyaremos de la desviación estándar para introducir las fluctuaciones en el sistema como a continuación

$$x(t) \pm 2\eta_{x(t)} = x(t) \pm 2\sigma_{x(t)}, \quad (4.30)$$

hemos elegido a 2η para que así aparezca la desviación estándar $2\sigma_{x(t)}$, debido a que en el intervalo de $4\sigma_{x(t)}$ alrededor del promedio se encuentra la mayoría de las fluctuaciones (a veces más del 90%), en otras palabras, en esta banda se encuentran los resultados que se obtendrían de las simulaciones estocásticas.

De la misma manera construiremos un programa de Python para poder simular algunas reacciones bioquímicas usando a la variante del TFD, nuevamente usaremos a (4.30) para construir las bandas en donde se encontrarán las fluctuaciones del sistema, para ver lo que se obtiene y cómo funciona el programa veremos un ejemplo a continuación. El sistema que llegaremos a analizar será el toggle switch.

4.6.1. Toggle switch

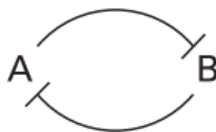
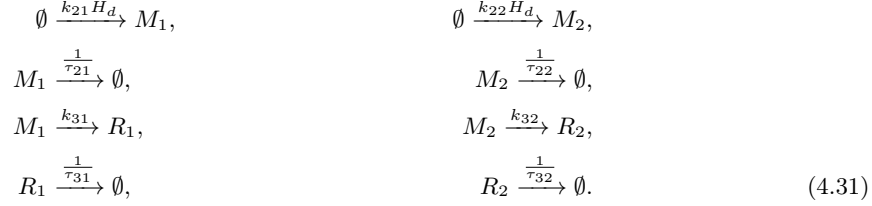


Figura 4.2: Toggle switch.

Para mostrar lo que se obtiene usando el algoritmo de Gillespie analizaremos el “toggle switch”, aquí las proteínas actúan como represores al momento de la síntesis del ARm, es decir (r_1 , r_2 el número de proteínas, y m_1 , m_2 el número de ARm), si r_1 aumenta, entonces la síntesis de m_2

disminuye, e inversamente, si r_2 aumenta, m_1 disminuye, además todas estas cantidades sufren una degradación o disolución. Una simple red bioquímica que describe este sistema son las siguientes 8 reacciones.



A partir de estos procesos encontramos las siguientes tasas de propensión

$$\begin{array}{ll}
 t_1^+ = k_{21}H_d(r_2/\Omega), & t_1^- = \frac{1}{\tau_{21}} \frac{m_1}{\Omega}, \\
 t_2^+ = k_{31} \frac{m_1}{\Omega}, & t_2^- = \frac{1}{\tau_{31}} \frac{r_1}{\Omega}, \\
 t_3^+ = k_{22}H_d(r_1/\Omega), & t_3^- = \frac{1}{\tau_{22}} \frac{m_2}{\Omega}, \\
 t_4^+ = k_{32} \frac{m_2}{\Omega}, & t_4^- = \frac{1}{\tau_{32}} \frac{r_2}{\Omega}
 \end{array} \tag{4.32}$$

aquí se ha introducido la ecuación de Hill determinista H_d , definida por

$$H_d(x) = \frac{\left(\frac{x}{K_R}\right)^n}{1 + \left(\frac{x}{K_R}\right)^n}, \tag{4.33}$$

además, se han introducido unas constantes de reacción, las k_{ij} están relacionadas a la síntesis, mientras que t_{ij}^- a los tiempos de degradación y disolución. La matriz estequiométrica es

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.34}$$

Nos apoyamos de la siguiente relación para escribir la ecuación maestra del sistema

$$\partial_t P(S_j, t) = \Omega \sum_i \left(t_i^- (S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+ (S_j) P(S_j, t) + t_i^+ (S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^- (S_j) P(S_j, t) \right).$$

Finalmente, la ecuación maestra de este sistema está dado por

$$\begin{aligned}
 \frac{d}{dt} P(m_1, r_1, m_2, r_2, t) = & \Omega \sum_{i=1}^2 k_{2i} H_d\left(\frac{r_1}{\Omega}\right) P(m_i - 1, r_i, t) - k_{2i} H_d\left(\frac{r_1}{\Omega}\right) P(m_i, r_i, t) \\
 & + \frac{m_i + 1}{\tau_{2i}} P(m_i + 1, r_i, t) - \frac{m_i}{\tau_{2i}} P(m_i, r_i, t) \\
 & + k_{3i} m_i P(m_i, r_i - 1, t) - k_{3i} m_i P(m_i, r_i, t) \\
 & + \frac{r_i + 1}{\tau_{3i}} P(m_i, r_i + 1, t) - \frac{r_i}{\tau_{3i}} P(m_i, r_i, t) \Big).
 \end{aligned} \tag{4.35}$$

Con estas expresiones ya estamos preparados para realizar las simulaciones estocásticas de nuestro sistema para determinar los promedios y las fluctuaciones, de hecho, solo emplearemos a los vectores de propensión del sistema y a la matriz estequiométrica. Nos apoyaremos de un programa muy similar al que describimos en la parte final del capítulo anterior, para más detalles del código ver la carpeta de Drive en donde se encuentra el archivo completo (*redgetswitch.py*).

Fluctuaciones intrínsecas en una red bioquímica estocástica

4.6 Cuantificar fluctuaciones

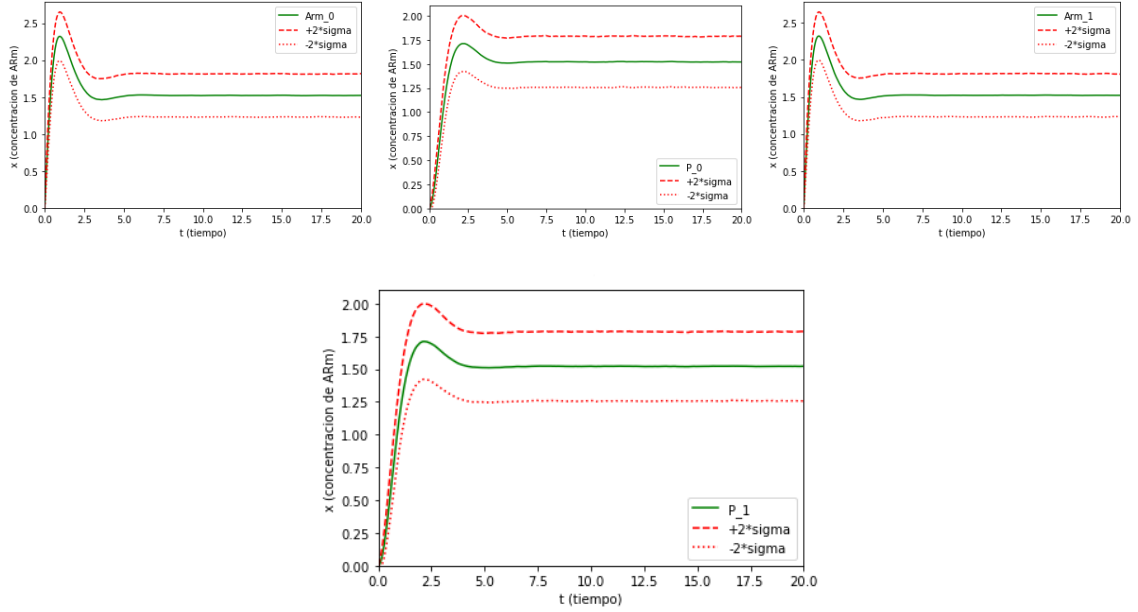


Figura 4.3: Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de simulaciones estocásticas, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_1(0) = m_2(0) = 0$, $n = 2$, $\Omega = 80$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$, tamaño del ensamble= 10000.

Lo que se obtiene del programa de Python está en la figura 4.3, en las figuras de la parte izquierda se observa la concentración del ARm y a en las de la derecha las de la proteína, la línea verde muestra el promedio de las variables sobre el ensamble que hemos utilizado, mientras que las líneas punteadas nos dicen que en este intervalo se encontraran la mayoría de las fluctuaciones del sistema, una característica muy interesante de esta es que estas fluctuaciones tienden a mantenerse en un determinado valor conforme el sistema se acerca al estado estacionario, además las fluctuaciones son relativamente considerables con respecto a la concentración de las especies bioquímicas del sistema debido principalmente a que el sistema es pequeño. Aunque hemos usado un ensamble de tamaño 10000 los resultados que aparecen en las gráficas ya son lo suficientemente uniformes.

También usaremos la variante del TFD para modelar nuestro sistema, para esto realizamos la aproximación determinista del sistema, es decir, usaremos el siguiente conjunto de ecuaciones diferenciales

$$\begin{aligned} \frac{\partial s_j}{\partial t} &= \sum_i \Gamma_{ji} \left(R_i^{D+}(s_l) - R_i^{D-}(s_l) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \\ \frac{\partial}{\partial t} \sigma^2(s_{l_1}, s_{l_2}) &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D+}(s_j) + R_i^{D-}(s_j))}{\Omega} \right. \\ &\quad + \sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_j) - R_i^{D-}(s_j)) \\ &\quad \left. + \sum_{j_1} \sum_{j_2} \sigma^2(s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_j) + R_i^{D-}(s_j)) \right), \end{aligned}$$

Fluctuaciones intrínsecas en una red bioquímica estocástica
4.6 Cuantificar fluctuaciones

uno de los elementos importantes son las tasas deterministas y la matriz estequiométrica, la matriz estequiométrica está en (4.34) y las tasas de reacción son

$$\begin{aligned}
 t_1^{D+} &= k_{21}H_d(R_2), & t_1^{D-} &= \frac{1}{\tau_{21}}M_1, \\
 t_2^{D+} &= k_{31}M_1, & t_2^{D-} &= \frac{1}{\tau_{31}}R_1, \\
 t_3^{D+} &= k_{22}H_d(R_1), & t_3^{D-} &= \frac{1}{\tau_{22}}M_2, \\
 t_4^{D+} &= k_{32}M_2, & t_4^{D-} &= \frac{1}{\tau_{32}}R_2,
 \end{aligned} \tag{4.36}$$

con estos elementos ya tenemos un sistema de ecuaciones diferenciales, y para resolverlo usaremos un programa de Python (el archivo se encuentra en la carpeta de Drive con nombre *atswitch2.py*).

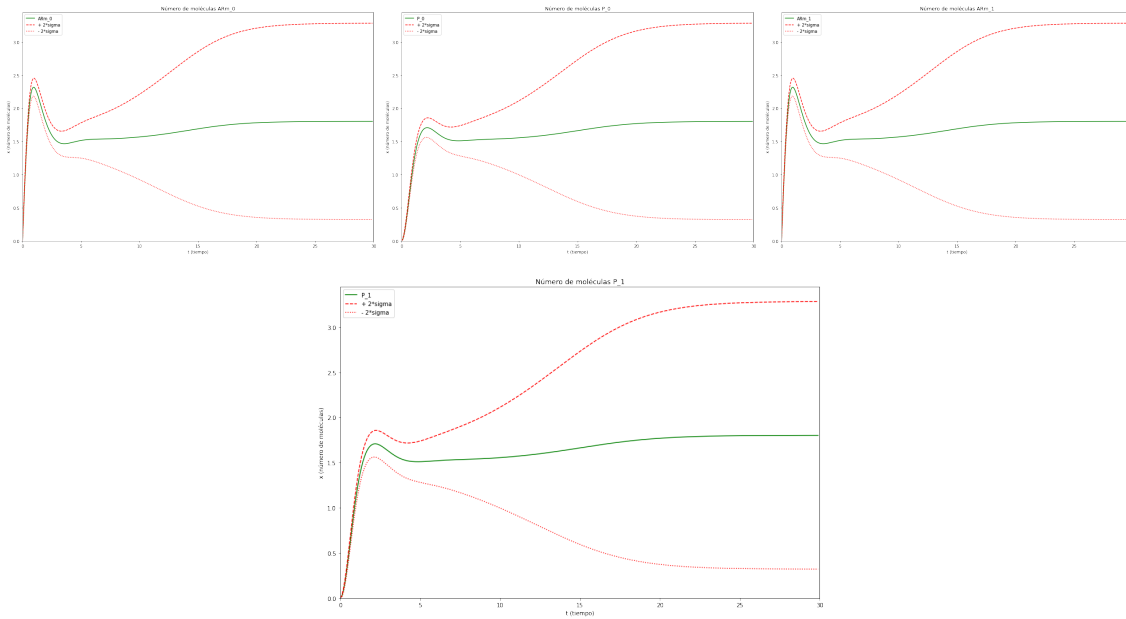


Figura 4.4: Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_1(0) = m_2(0) = 0$, $n = 2$, $\Omega = 80$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$.

Lo que se obtiene de este programa se encuentra en la figura 4.4, entre los resultados importantes que nos gustaría recalcar el hecho que usar al TFD para modelar al sistema implica resolver un conjunto de ecuaciones diferenciales, lo que se traduce en un tiempo de computación mucho más corto. En la figura la línea verde representa al promedio determinista y las bandas rojas punteadas a las fluctuaciones del sistema, se observa que después de un determinado tiempo las fluctuaciones crecen mucho más que lo obtenido del algoritmo de Gillespie, una posible interpretación de que las bandas estén muy anchas es el hecho que este punto de equilibrio tiene un potencial asociado muy pequeño, por lo que el sistema estocástico tiene la posibilidad de saltar y alejarse muy fácilmente de este punto de equilibrio, esta posibilidad no se captura a través del algoritmo de Gillespie debido a que el tamaño del ensambles es solo de 10000, por lo que no se

capturarían todos los posibles eventos.

A partir de los resultados uno se pregunta si eligiendo unas condiciones iniciales diferentes el sistema convergería al otro punto de equilibrio, y en este caso como se comportarían las fluctuaciones del sistema. Los resultados de esto se pueden observar en la figura 4.5, aquí el sistema converge al otro punto de equilibrio, las líneas verdes representan a las concentraciones deterministas y las líneas rojas punteadas a las fluctuaciones del sistema, podemos observar que la banda de las fluctuaciones es pequeña cuando el sistema se acerca a este punto de equilibrio, esto se debería al hecho que este punto de equilibrio tiene asociado un potencial grande. Otra posible causa puede ser el tipo de función Hill que hemos usado, aunque esto lo veremos más adelante al analizar las redes de regulación genética.

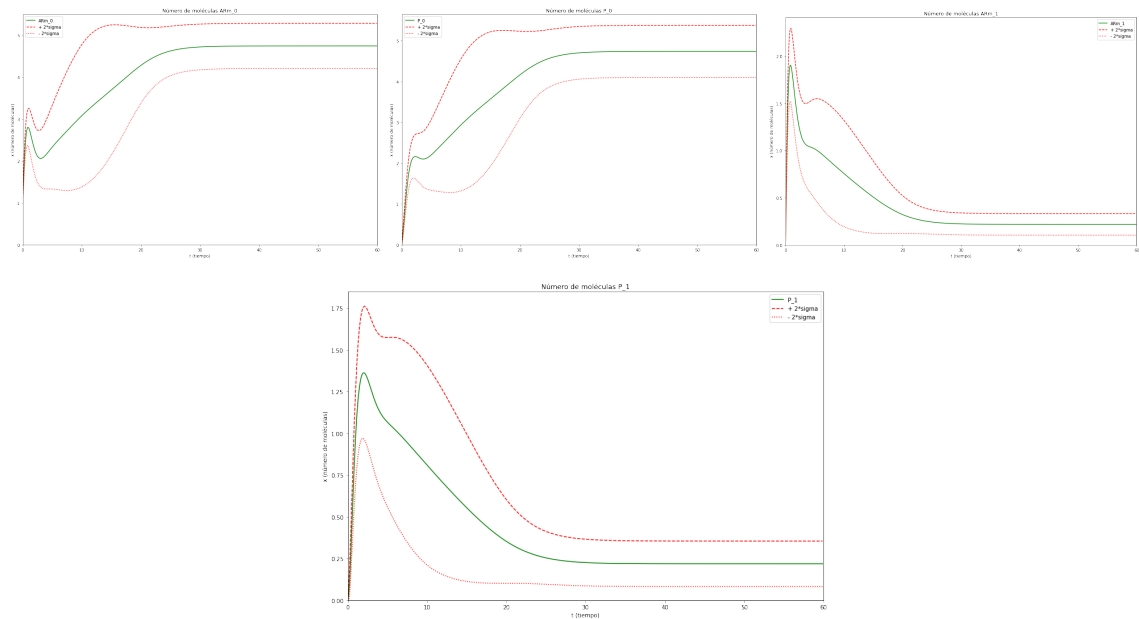


Figura 4.5: Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, se ha elegido una condición inicial diferente, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = m_2(0) = 0$, $m_1(0) = 1$, $n = 2$, $\Omega = 80000$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$.

A partir de las simulaciones nos dimos cuenta de otro efecto interesante, los resultados se pueden observar en la figura 4.6, aquí se han usado las mismas condiciones iniciales que en la figura 4.4 aunque el tamaño del sistema ahora es de 80000, aquí se observa que el tiempo en el que el sistema empieza a alejarse del punto de equilibrio es mucho mayor, por lo que cuando el tamaño del sistema se vuelve lo suficientemente grande el tiempo en el que pasa este efecto es también muy grande, y es por eso que esto no se notan en el límite determinista.

Fluctuaciones intrínsecas en una red bioquímica estocástica

4.7 Retroactividad con correcciones estocásticas

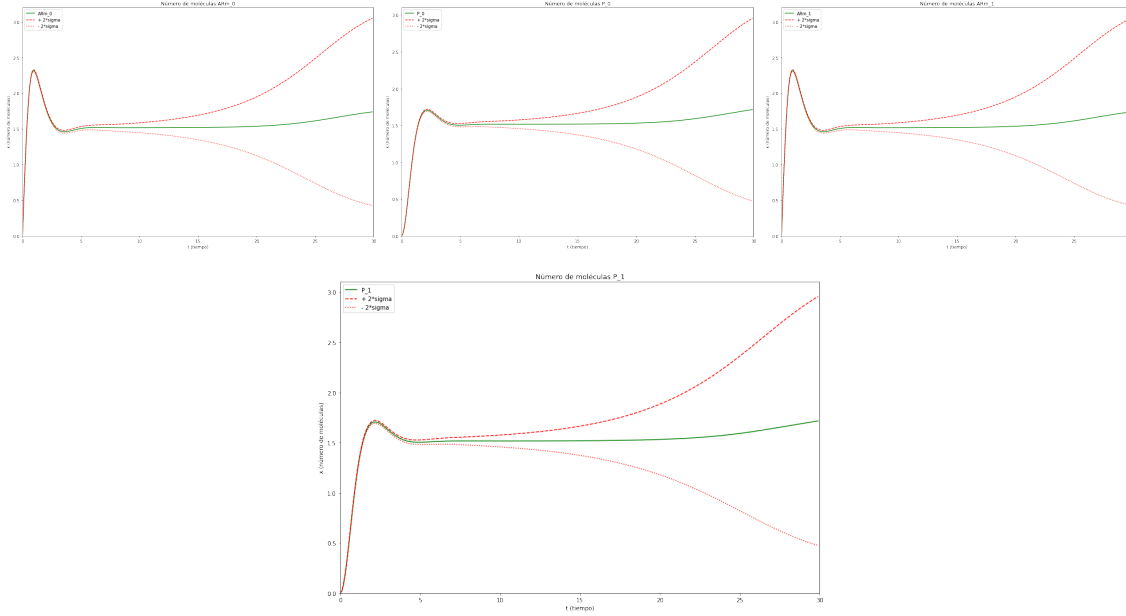


Figura 4.6: Toggle switch. Se ha modelado la dinámica de las proteínas y el ARm a partir de la aproximación determinista, en todas las figuras las líneas verdes representan la dinámica de las concentraciones deterministas de las especies bioquímicas involucradas, mientras que las líneas rojas punteadas a la banda en donde se encuentran las fluctuaciones. Con ayuda de los siguientes parámetros: $K = 1$, $r_1(0) = r_2(0) = 0$, $n = 2$, $\Omega = 80000$, $\tau_{2i} = \tau_{3i} = k_{3i} = 1$, $k_{2i} = 10$.

4.7. Retroactividad con correcciones estocásticas

Otra de las cuestiones en la que estamos interesados es la retroactividad en los sistemas modulares, debido a que estos sistemas también pueden ser pequeños (ya que podrían emplearse en circuitos genéticos o sistemas similares), por lo que al momento de cuantificar la retroactividad de los sistemas es indispensable considerar las fluctuaciones y sus posibles efectos. Ya hemos desarrollado en el capítulo 2 un método para encontrar la retroactividad en sistemas bioquímicos (recomendamos leer esa sección para recordar algunos conceptos y desarrollos). Con lo desarrollado en este capítulo nos damos cuenta de que, si consideramos las fluctuaciones, ahora la dinámica de nuestros sistemas bioquímicos estará descrito por el siguiente conjunto de ecuaciones diferenciales

$$\begin{aligned} \frac{\partial s_j}{\partial t} &= \sum_i \Gamma_{ji} \left(R_i^{D+}(s_i) - R_i^{D-}(s_i) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_i) - R_i^{D-}(s_i)) \right), \\ \frac{\partial}{\partial t} \sigma^2(s_{l_1}, s_{l_2}) &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D+}(s_j) + R_i^{D-}(s_j))}{\Omega} \right. \\ &\quad + \sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_j) - R_i^{D-}(s_j)) \\ &\quad \left. + \sum_{j_1} \sum_{j_2} \sigma^2(s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_j) + R_i^{D-}(s_j)) \right), \end{aligned}$$

podemos fijarnos que entre la ley de acción de masas deterministas y la que tiene correcciones estocásticas la única diferencia son las tasas de reacción, en donde aparecen las fluctuaciones, por lo que podemos recalculer la retroactividad de nuestros sistemas con correcciones estocásticas, para

esto pensemos que las tasas de reacción del sistema se transforman como a continuación

$$R_i^{D\pm}(x_l) \rightarrow R_i^{\pm D}(x_l) + R_i^{\pm S}(x_l), \quad (4.37)$$

el primer término corresponde a las tasas deterministas y la segunda a la corrección estocástica, siguiendo con esta idea, y la notación que hemos empleado en el capítulo 2 podemos escribir a las ecuaciones de movimiento como

$$\begin{aligned} Q_{ij}(R_j^+(x_l) - R_j^-(x_l)) &\rightarrow QR_i^{D'} + QR_i^{S'}, \\ QR_i^{D'} &= Q_{ij}(R_j^{D+}(x_l) - R_j^{D-}(x_l)), \\ QR_i^{S'} &= Q_{ij}(R_j^{S+}(x_l) - R_j^{S-}(x_l)), \end{aligned} \quad (4.38)$$

es decir, existe una clara separación entre la parte determinista $QR_i^{D'}$ y la parte estocástica $QR_i^{S'}$, recordamos que en el capítulo 2 dimos una expresión para la retroactividad dada como

$$\begin{aligned} \frac{dx_i}{dt} &= Q_{ii}(R_i^+(x_l) - R_i^-(x_l)) \left(1 + \sum_j A_{ij} \mathcal{R}_j^i \right), \\ A_{ij} \mathcal{R}_j^i &= A_{ij} \left(\frac{dw_j}{d\Omega_i} \right) = -A_{ij} \frac{\left(Q_{ij} \frac{\partial(R_j^+(\Omega_l + \sum_j A_{lj} w_j) - R_j^-(\Omega_l + \sum_j A_{lj} w_j))}{\partial \Omega_i} \right)}{\left(Q_{ij} \frac{\partial(R_j^+(\Omega_l + \sum_j A_{lj} w_j) - R_j^-(\Omega_l + \sum_j A_{lj} w_j))}{\partial w_i} \right)}, \end{aligned}$$

(A_{ij} es la conexión entre módulos), aplicamos la notación dada en (4.38) a \mathcal{R}_j^i , introduciendo las correcciones estocásticas la expresión para la retroactividad se modificaría como

$$\mathcal{R}_i^j = - \frac{\left(\frac{\partial(QR_i^{D'} + QR_i^{S'})}{\partial \Omega_i} \right)}{\left(\frac{\partial(QR_i^{D'} + QR_i^{S'})}{\partial w_i} \right)}, \quad (4.39)$$

definimos a la retroactividad determinista \mathcal{R}_i^{jD} y estocástica \mathcal{R}_i^{jS} como

$$\mathcal{R}_i^{jD} = \frac{\left(\frac{\partial(QR_i^{D'})}{\partial \Omega_i} \right)}{\left(\frac{\partial(QR_i^{D'})}{\partial w_i} \right)}, \quad \mathcal{R}_i^{jS} = \frac{\left(\frac{\partial(QR_i^{S'})}{\partial \Omega_i} \right)}{\left(\frac{\partial(QR_i^{S'})}{\partial w_i} \right)}, \quad (4.40)$$

por lo que ahora la retroactividad será de la forma

$$\mathcal{R}_i^j = - \mathcal{R}_i^{jD} \left(\frac{1 + \frac{\partial QR_i^{S'}}{\partial \Omega_i} \left(\frac{\partial QR_i^{D'}}{\partial \Omega_i} \right)^{-1}}{1 + \frac{\partial QR_i^{S'}}{\partial w_i} \left(\frac{\partial QR_i^{D'}}{\partial w_i} \right)^{-1}} \right), \quad (4.41)$$

el término que sacamos del paréntesis corresponde a la retroactividad determinista, mientras que los otros términos se podrían considerar como correcciones estocásticas, podemos ver que el término $\frac{\partial R_i^S}{\partial \Omega_i} \left(\frac{\partial R_i^D}{\partial \Omega_i} \right)^{-1}$ va aumentando la retroactividad, mientras que $\frac{\partial R_i^S}{\partial w_i} \left(\frac{\partial R_i^D}{\partial w_i} \right)^{-1}$ lo va atenuando, por lo que con ayuda de las fluctuaciones es posible que la retroactividad sea menor en comparación con la determinista por lo que se obtendría una mayor estabilidad, también es posible que ocurra lo contrario, las contribuciones estocásticas aparecen en el sistema cuando las tasas de reacción son no lineales y/o de orden 2. Es importante mencionar que la retroactividad se calcula pensando que algunas de las variables están en equilibrio, por lo que la contribución estocástica en la retroactividad será la misma que cuando los sistemas se encuentren aislados.

Es posible que uno se pregunte si existe una manera en la cual a pesar de tener correcciones estocásticas la retroactividad no dependa de estos términos, para ello suponemos que $\frac{\partial Q R_i^S}{\partial \Omega_i} \left(\frac{\partial Q R_i^D}{\partial \Omega_i} \right)^{-1}$ y $\frac{\partial Q R_i^S}{\partial w_i} \left(\frac{\partial Q R_i^D}{\partial w_i} \right)^{-1}$ son muy pequeños, por lo que realizando expansión en serie a primer orden de (4.41) obtenemos

$$\mathcal{R}_i^j \approx -\mathcal{R}_i^{jD} \left(1 + \frac{\partial Q R_i^S}{\partial \Omega_i} \left(\frac{\partial Q R_i^D}{\partial \Omega_i} \right)^{-1} - \frac{\partial Q R_i^S}{\partial w_i} \left(\frac{\partial Q R_i^D}{\partial w_i} \right)^{-1} \right), \quad (4.42)$$

para que en el sistema desaparezcan las contribuciones estocásticas se tiene que cumplir la siguiente condición

$$\frac{\partial Q R_i^S}{\partial \Omega_i} \left(\frac{\partial Q R_i^D}{\partial \Omega_i} \right)^{-1} - \frac{\partial Q R_i^S}{\partial w_i} \left(\frac{\partial Q R_i^D}{\partial w_i} \right)^{-1}, \quad (4.43)$$

despejando

$$\mathcal{R}_i^{jD} = \frac{\partial Q R_i^D}{\partial \Omega_i} \left(\frac{\partial Q R_i^D}{\partial w_i} \right)^{-1} = \frac{\partial Q R_i^S}{\partial \Omega_i} \left(\frac{\partial Q R_i^S}{\partial w_i} \right)^{-1} = \mathcal{R}_i^{jS}, \quad (4.44)$$

es decir, la retroactividad determinista y estocástica tienen que ser iguales para que desaparezca toda la contribución estocástica a la retroactividad de todo el sistema, notar que esta misma fórmula es válida para la expresión (4.41), haciendo que desaparezcan las contribuciones estocásticas.

Con estos resultados ya tenemos las herramientas suficientes para analizar la modularidad en las redes bioquímicas, de hecho para modelar estos sistemas se usaría un sistema de ecuaciones diferenciales similar al que se presentó al principio de esta sección, ya que ayuda a cuantificar las fluctuaciones del sistema, como ya hemos mencionado, el término \mathcal{R}_i^j solo nos dice si nuestro sistema seguirá presentado modularidad al conectarlo a un nuevo módulo, o si nuestro sistema se modificara al conectarle un nuevo módulo.

Capítulo 5

Fosforilación

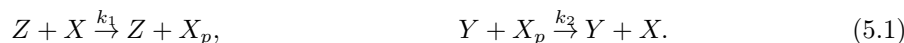
Ya tenemos todas las técnicas y herramientas estocásticas para analizar a fondo cualquier sistema bioquímico que sigue a la ley de acción de masas, incluyendo los programas que nos ayudan a simularlos a partir del algoritmo de Gillespie, así como con la ayuda del teorema de fluctuación disipación (TFD), por lo que ahora nos daremos a la tarea de analizar otro sistema que es relativamente más complejo, analizaremos el proceso de fosforilación/desfosforilación con el fin de entender y visualizar como se modifica la retroactividad cuando están presentes fluctuaciones, para ello nos apoyaremos de las reacciones que aparecen en [9], ya lo hemos mencionado en el capítulo 2. En la primera parte construiremos su ecuación maestra y determinaremos la distribución estacionaria, así como las fluctuaciones del sistema. En la parte final de este capítulo implementaremos el TFD para modelar nuestro sistema y también cuantificar las fluctuaciones durante toda la dinámica del sistema ya que el algoritmo de Gillespie tiene un tiempo muy alto de computación. Recordar que los códigos que usamos para simular se encuentran en la carpeta de Drive (https://drive.google.com/drive/folders/1DcDw7D0B-hq-9gxt-8z4P3_K63m1eU01?usp=sharing).

5.1. Análisis a partir del algoritmo de Gillespie

En esta primera sección analizaremos al sistema de fosforilación/desfosforilación usando el algoritmo de Gillespie, para ello usaremos un programa de Python muy similar al empleado en el capítulo anterior, aunque debido a que el coste computacional para inferir las fluctuaciones a partir de este es muy alto, en la siguiente sección emplearemos otro enfoque para este fin.

5.1.1. Caso I

Vamos a considerar de nuevo el sistema que fue presentado por Del Vecchio, para ello nos apoyamos de la figura 5.1. La reacción que está a la izquierda corresponde a un proceso de fosforilación y desfosforilación en cual participan las proteínas Z y Y , para ello vamos a considerar un modelo de una reacción de un paso para las reacciones de fosforilación y desfosforilación analizadas por [8]



Dado que, para las reacciones químicas, es este caso bioquímicas, se hace uso de los procesos de vida y muerte multivariable, escribiremos también la reacción de la derecha de la imagen, en el cual una molécula fosforilada se une a una promotora P



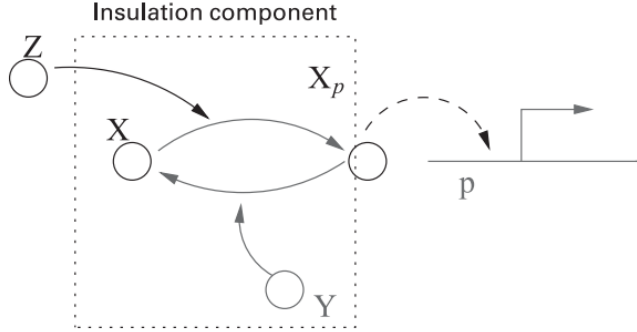


Figura 5.1: Componente aislado [9]. En la figura se observa como está aislada la región en donde ocurre el proceso de fosforilación/desfosforilación.

en este sistema la proteína Z está experimentando cambios en el tiempo, más específicamente está siendo sintetizada por una función externa, esto lo capturamos en la parte derecha de la ecuación anterior con la función $k(t)$. Con ayuda de todas las reacciones anteriores encontramos la matriz de coeficientes estequiométricos y la matriz estequiométrica

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Gamma_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.3)$$

apoyándonos de estas calculamos las tasas de propensión que están definidas como

$$t_i^+(S_j) = k_i^+ \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!}, \quad t_i^-(S_j) = k_i^- \prod_j \frac{S_j!}{\Omega^{\beta_{ij}} (S_j - \beta_{ij})!},$$

de donde obtenemos

$$\begin{aligned} t_1^+ &= k(t), & t_1^- &= \delta z \frac{1}{\Omega}, \\ t_2^+ &= k_1 x z \frac{1}{\Omega^2}, & t_2^- &= 0, \\ t_3^+ &= k_2 x_p y \frac{1}{\Omega^2}, & t_3^- &= 0, \\ t_4^+ &= k_{on} x_p p \frac{1}{\Omega^2}, & t_4^- &= k_{off} c \frac{1}{\Omega}. \end{aligned} \quad (5.4)$$

Hemos denotado a $\{z, x, x_p, y, p, c\}$ como el número de moléculas bioquímicas de $\{Z, X, X_p, Y, P, X_p p\}$ respectivamente. Nos apoyamos de la siguiente expresión para calcular la ecuación maestra

$$\partial_t P(S_j, t) = \Omega \sum_i \left(t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+(S_j) P(S_j, t) + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^-(S_j) P(S_j, t) \right).$$

Usando los elementos que hemos calculado evaluamos, así finalmente tendremos la siguiente ecuación maestra

$$\begin{aligned}
 \frac{d}{dt}P(z, x, x_p, p, c, t) &= k_1(x+1)\frac{z}{\Omega}P(z, x+1, x_p-1, p, c, t) - k_1(x)\frac{z}{\Omega}P(z, x, x_p, p, c, t) \\
 &+ k_2(x_p+1)\frac{y}{\Omega}P(z, x-1, x_p+1, p, c, t) - k_2(x_p)\frac{y}{\Omega}P(z, x, x_p, p, c, t) \\
 &+ k_{on}(x_p+1)\frac{(p+1)}{\Omega}P(z, x, x_p+1, p+1, c-1, t) - k_{off}cP(z, x, x_p, p, c, t) \\
 &+ k_{off}(c+1)P(z, x, x_p-1, p-1, c+1, t) - k_{on}x_p\frac{p}{\Omega}P(z, x, x_p, p, c, t) \\
 &+ k(t)\Omega P(z-1, x, x_p, p, c, t) - \delta zP(z, x, x_p, p, c, t) \\
 &+ \delta(z+1)P(z+1, x, x_p, p, c, t) - k(t)\Omega P(z, x, x_p, p, c, t).
 \end{aligned} \tag{5.5}$$

De la matriz estequiométrica encontramos tres cantidades conservadas la primera es $x + x_p + c = x_0$, la segunda es $p + c = c_0$, existe otra cantidad conservada en el sistema el cual es el número de proteínas Y se mantiene constante, aunque esto es debido a las reacciones del sistema.

Para determinar la distribución estacionaria de nuestro sistema igualaremos la ecuación (5.5) anterior a cero y luego usamos el método que explicamos en el capítulo de análisis estocástico, es importante notar que las reacciones (5.1) y (5.2) no tienen deficiencia cero, pero siguiendo al artículo [20] se pueden transformar de tal manera que si la tengan (ver apéndice C), obteniendo lo siguiente

$$\begin{aligned}
 P(z, x, x_p, p, c)_S &= \frac{\left(\frac{k(t)\Omega}{\delta}\right)^z e^{-\frac{k(t)\Omega}{\delta}}}{z!} \frac{\left(\frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}\right)^c}{\left(1 + \frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}\right)^{p_0}} \binom{p_0}{c} \frac{\left(\frac{k_1\langle z(t) \rangle}{k_2 y}\right)^{x_p}}{\left(1 + \frac{k_1\langle z(t) \rangle}{k_2 y}\right)^{x_0 - \langle c \rangle}} \binom{x_0 - \langle c \rangle}{x_p} \\
 &* \delta_{x, x_0 - x_p - c} * \delta_{p, p_0 - c}.
 \end{aligned} \tag{5.6}$$

Donde hemos definido $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ este es el coeficiente binomial, aunque el promedio $\langle c \rangle$ no siempre es un valor entero, por lo que al momento de calcular el coeficiente binomial cambiaríamos al factorial por la función gamma. Podemos observar que el sistema se compone de una distribución de Poisson, y dos que son del tipo binomial, las de tipo binomial nos dicen que hay variables que están estrechamente relacionadas como es el caso de p y c .

Con la distribución estacionaria podemos determinar los promedios de las variables en el estado estacionario,

$$\langle z(t) \rangle = \frac{k(t)\Omega}{\delta}, \quad \langle x_p \rangle = \frac{\frac{k_1\langle z(t) \rangle}{k_2 y}}{1 + \frac{k_1\langle z(t) \rangle}{k_2 y}} (x_0 - \langle c \rangle), \quad \langle c \rangle = \frac{\frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}}{1 + \frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}} p_0. \tag{5.7}$$

A continuación simularemos este sistema usando el algoritmo de Gillespie (para más detalles del código revisar la carpeta de Drive en donde se encuentra el archivo *fosfo1.ipynb*), hemos elegido a $k(t) = 0,01(1 + \sin(0,005t))$ tal como se ha presentado en el artículo de Del Vecchio, para ello usaremos a (5.3) y a (5.4) para simular el sistema.

Hemos omitido presentar a algunas variables debido a que no nos interesa mucho su dinámica ya que se puede inferir de las demás. Hemos graficado la concentración de algunas de las moléculas del sistema (ver figura 5.2), dado que el sistema es modelado a partir de una descripción estocástica se presentan fluctuaciones, a pesar de esto se puede ver claramente la tendencia global del sistema en las gráficas que hemos realizado, el cual es de carácter oscilatorio.

Dado que estamos usando un modelo estocástico para describir este sistema es deseable conocer las fluctuaciones, como ya hemos calculado nuestra distribución en el estado estacionario, entonces podemos calcular a partir de ella las fluctuaciones de nuestro sistema en estado estacionario, usamos

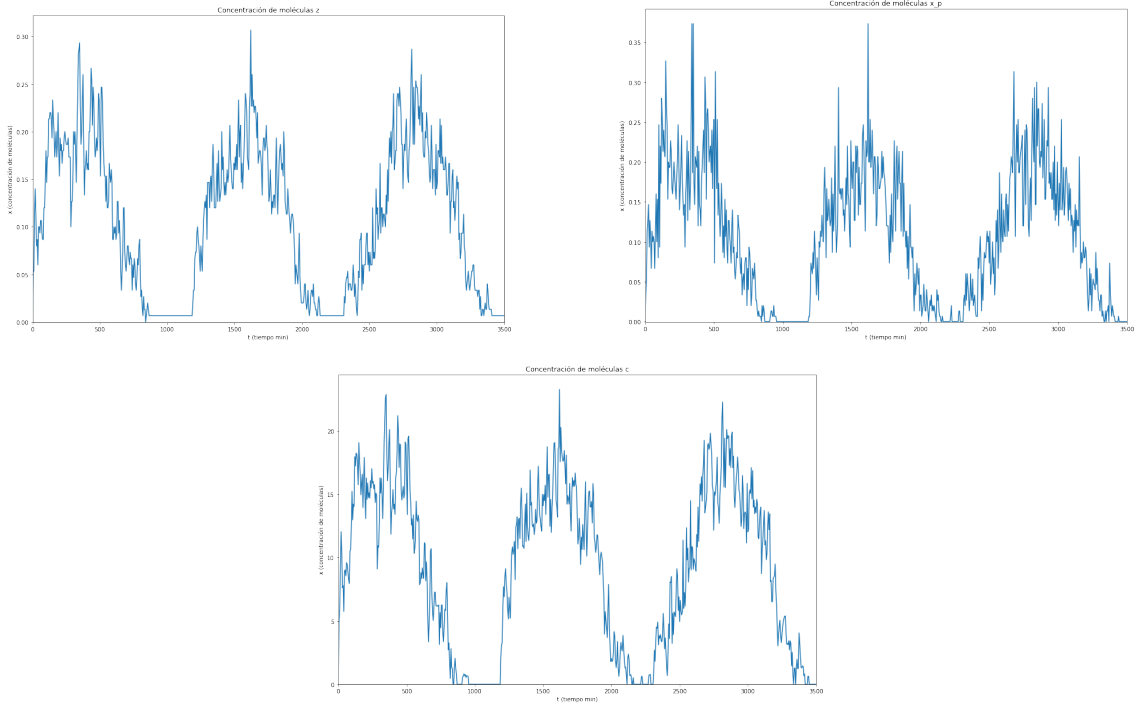


Figura 5.2: Dinámica estocástica del proceso de fosforilación y desfosforilación, con $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$.

a la distribución estacionaria para obtener las varianzas y luego dividimos entre Ω^2 , obtenemos lo siguiente

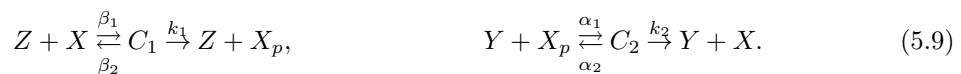
$$\eta_z^2 = \frac{1}{\Omega} \frac{\langle z(t) \rangle}{\Omega}, \quad \eta_{x_p}^2 = \frac{1}{\Omega} \frac{\frac{\langle x_p \rangle}{\Omega}}{1 + \frac{k_1 \langle z \rangle}{k_2 y}}, \quad \eta_c^2 = \frac{1}{\Omega} \frac{\frac{\langle c \rangle}{\Omega}}{1 + \frac{k_{on} \langle x_p \rangle}{k_{off} \Omega}}, \quad (5.8)$$

En estas expresiones podemos ver que dependen del inverso de Ω , ya que cuando un promedio se divide entre Ω corresponde a la concentración determinista, por lo que cuando el sistema se vuelve lo suficientemente grande las fluctuaciones desaparecen.

Simular este sistema empleando el algoritmo de Gillespie tiene un gran coste computacional, esto sería aún más si queremos usar este algoritmo para determinar las fluctuaciones de este sistema, por lo que en la siguiente sección usaremos al TFD que hemos encontrado anteriormente para simular la dinámica del sistema e inferir las fluctuaciones de esta.

5.1.2. Caso II

Ahora seguiremos analizando al proceso de fosforilación/desfosforilación, aunque en este caso supondremos que existen estados intermedios al momento que la molécula se fosforilación y desfosforilación tal como a continuación



cuando la molécula fosforilada se une a una promotora seguiremos teniendo las siguientes reacciones



usaremos estas reacciones para construir un modelo estocástico del sistema, por lo que ahora la matriz estequiométrica y las matrices de coeficientes estequiométricos del sistema serán

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\Gamma_{ij} = \begin{pmatrix} 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.11)$$

con ayuda de estas calculamos las tasas de propensión que están definidas como

$$t_i^+(S_j) = k_i^+ \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!}, \quad t_i^-(S_j) = k_i^- \prod_j \frac{S_j!}{\Omega^{\beta_{ij}} (S_j - \beta_{ij})!},$$

de donde obtenemos

$$\begin{aligned} t_1^+ &= k(t), & t_1^- &= \delta z \frac{1}{\Omega}, \\ t_2^+ &= \beta_1 x z \frac{1}{\Omega^2}, & t_2^- &= \beta_2 c_1 \frac{1}{\Omega}, \\ t_3^+ &= k_1 c_1 \frac{1}{\Omega}, & t_3^- &= 0, \\ t_4^+ &= \alpha_1 x_p y \frac{1}{\Omega^2}, & t_4^- &= \alpha_2 c_2 \frac{1}{\Omega}, \\ t_5^+ &= k_2 c_2 \frac{1}{\Omega}, & t_5^- &= 0, \\ t_6^+ &= k_{on} x_p p \frac{1}{\Omega^2}, & t_6^- &= k_{off} c \frac{1}{\Omega}. \end{aligned} \quad (5.12)$$

Hemos denotado a $\{z, x, x_p, y, p, c, c_1, c_2\}$ como el número de moléculas bioquímicas de $\{Z, X, X_p, Y, P, X_p p, XZ, X_p Y\}$ respectivamente. Nos apoyamos de la siguiente expresión para calcular la ecuación maestra

$$\partial_t P(S_j, t) = \Omega \sum_i \left(t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+(S_j) P(S_j, t) + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^-(S_j) P(S_j, t) \right).$$

Evaluando todos los elementos que hemos calculado obtenemos la ecuación maestra de nuestro sistema,

$$\begin{aligned}
 \dot{P}(z, x, x_p, y, p, c, c_1, c_2, t) = & \beta_1(x+1) \frac{(z+1)}{\Omega} P(z+1, x+1, x_p, y, p, c, c_1-1, c_2, t) - \beta_2 c_1 P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + \beta_2(c_1+1) P(z-1, x-1, x_p, y, p, c, c_1+1, c_2, t) - \beta_1 x \frac{z}{\Omega} P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + k_1(c_1+1) P(z-1, x, x_p-1, y, p, c, c_1+1, c_2, t) - k_1 c_1 P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + \alpha_1(x_p+1) \frac{(y+1)}{\Omega} P(z, x, x_p+1, y+1, p, c, c_1, c_2-1, t) - \alpha_2 c_2 P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + \alpha_2(c_2+1) P(z-1, x-1, x_p, y, p, c, c_1, c_2+1, t) - \alpha_1 x_p \frac{y}{\Omega} P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + k_2(c_2+1) P(z, x, x_p-1, y-1, p, c, c_1, c_2+1, t) - k_2 c_2 P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + k_{on}(x_p+1) \frac{(p+1)}{\Omega} P(z, x, x_p+1, y, p+1, c-1, c_1, c_2, t) - k_{off} c P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + k_{off}(c+1) P(z, x, x_p-1, y, p-1, c+1, c_1, c_2, t) - k_{on} x_p \frac{p}{\Omega} P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + k(t) \Omega P(z-1, x, x_p, y, p, c, c_1, c_2, t) - \delta z P(z, x, x_p, y, p, c, c_1, c_2, t) \\
 & + \delta(z+1) P(z+1, x, x_p, y, p, c, c_1, c_2, t) - k(t) \Omega P(z, x, x_p, y, p, c, c_1, c_2, t). \tag{5.13}
 \end{aligned}$$

De la matriz estequiométrica encontramos tres cantidades conservadas la primera es $x + x_p + c + c_1 + c_2 = x_0$, la segunda es $y + c_2 = y_0$ y por último $p + c = p_0$, podemos apoyarnos de estas cantidades conservadas para calcular la ecuación maestra de nuestro sistema en el estado estacionario, es importante notar que las reacciones (5.9) y (5.9) no tienen deficiencia cero, pero siguiendo al artículo [20] se pueden transformar de tal manera que si la tengan (ver apéndice C), entonces nos apoyamos del método ya presentado y obtenemos

$$\begin{aligned}
 P(z, x_p, c)_S = & \frac{\left(\frac{k(t)\Omega}{\delta}\right)^z e^{-\frac{k(t)\Omega}{\delta}}}{z!} \frac{\left(\frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}\right)^c}{\left(1 + \frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}\right)^{p_0}} \binom{p_0}{c} \frac{\left(\frac{\beta_1\langle z \rangle}{\beta_2+k_1}\frac{1}{\Omega}\right)^{c_1}}{\left(1 + \frac{\beta_1\langle z \rangle}{\beta_2+k_1}\frac{1}{\Omega}\right)^{x_0-\langle x_p+c+c_2 \rangle}} \binom{x_0-\langle x_p+c+c_2 \rangle}{c_1} \\
 & \frac{\left(\frac{\alpha_1\langle x_p \rangle}{\alpha_2+k_2}\frac{1}{\Omega}\right)^{c_2}}{\left(1 + \frac{\alpha_1\langle x_p \rangle}{\alpha_2+k_2}\frac{1}{\Omega}\right)^{y_0}} \binom{y_0}{c_2} \frac{\left(\frac{k_1\langle c_1 \rangle + \alpha_2\langle c_2 \rangle}{\alpha_1\frac{y_0-\langle c_2 \rangle}{\Omega}}\right)^{x_p} e^{-\frac{k_1\langle c_1 \rangle + \alpha_2\langle c_2 \rangle}{\alpha_1\frac{y_0-\langle c_2 \rangle}{\Omega}}}}{x_p!} \\
 & * \delta_{x, x_0-x_p-c-c_1-c_2} * \delta_{c_2, y_0-y} * \delta_{p, p_0-c}. \tag{5.14}
 \end{aligned}$$

Donde hemos definido $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, este es el coeficiente binomial, aunque el promedio $\langle c \rangle$ no siempre es un valor entero, por lo que al momento de calcular el coeficiente binomial cambiaríamos al factorial por la función gamma. Podemos notar que la distribución estacionaria se compone de dos de tipo Poisson y tres del tipo binomial, las binomiales aparecen debido a que algunas de nuestras variables están relacionadas.

Con la distribución estacionaria podemos determinar los promedios de las variables en el estado estacionario,

$$\begin{aligned}
 \langle z(t) \rangle = \frac{k(t)\Omega}{\delta}, \quad \langle x_p \rangle = \frac{k_1\langle c_1 \rangle + \alpha_2\langle c_2 \rangle}{\alpha_1\frac{y_0-\langle c_2 \rangle}{\Omega}}, \quad \langle c \rangle = \frac{\frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}}{1 + \frac{k_{on}\langle x_p \rangle}{k_{off}\Omega}} p_0, \\
 \langle c_2 \rangle = \frac{\alpha_1\frac{\langle x_p \rangle}{\Omega}}{\alpha_2 + k_2 + \alpha_1\frac{\langle x_p \rangle}{\Omega}} y_0, \quad \langle c_1 \rangle = \frac{\beta_1\frac{\langle z \rangle}{\Omega}}{\beta_2 + k_1 + \beta_1\frac{\langle z \rangle}{\Omega}} (x_0 - \langle x_p + c + c_2 \rangle). \tag{5.15}
 \end{aligned}$$

A continuación, simularemos este sistema usando el algoritmo de Gillespie (para más detalles del programa ver la carpeta de Drive en donde se encuentra el archivo *fosfo2.ipynb*), para implementar el programa hemos usado (5.11) y (5.12), hemos elegido a $k(t) = 0,01(1 + \sin(0,005t))$. Los

resultados que se obtienen se pueden observar en la figura 5.3, en este sistema podemos ver claramente la presencia de oscilaciones, aunque hemos omitido presentar a las variables y , p , c_1 y c_2 , debido a que no nos interesa mucho su dinámica y otras se pueden obtener a partir de las cantidades conservadas, otra característica que podemos ver de la simulación de un solo proceso es que el introducir estados intermedios hacen que nuestras gráficas sean más uniformes, es decir hay menos fluctuaciones en el sistema, aunque esto lo cuantificaremos mejor cuando determinemos las fluctuaciones en el sistema.

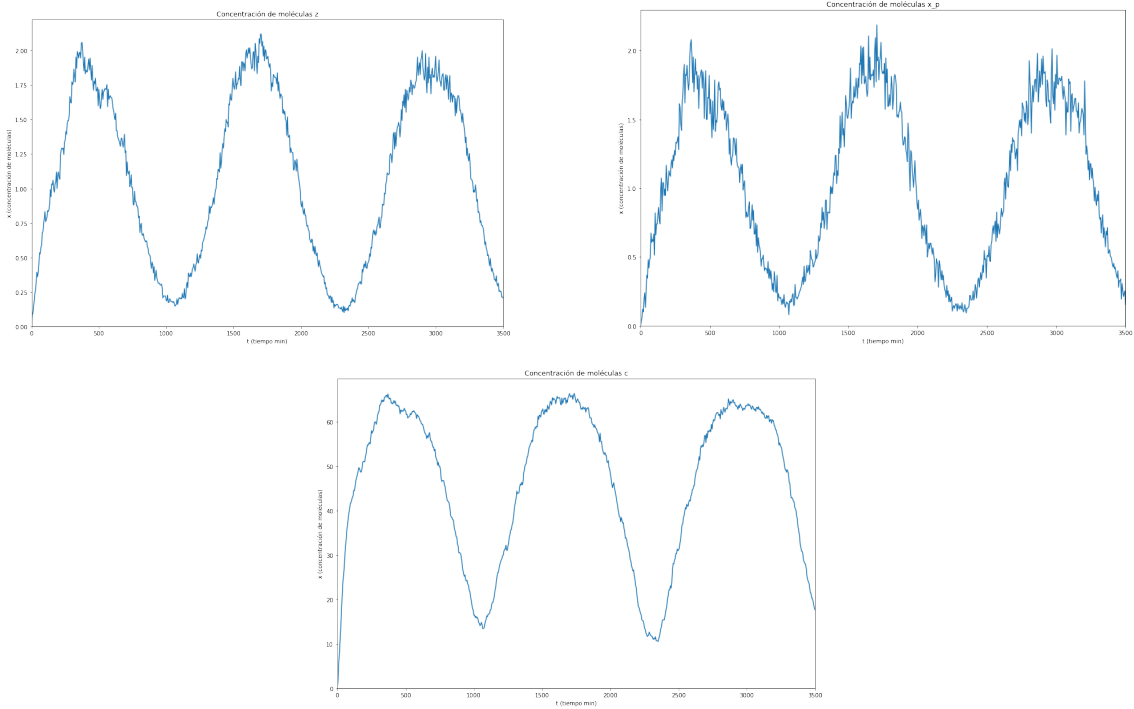


Figura 5.3: Dinámica estocástica del proceso de fosforilación y desfosforilación con procesos intermedios, con $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$.

Dado que estamos usando un modelo estocástico para describir este sistema es deseable conocer las fluctuaciones, como ya hemos calculado nuestra distribución en el estado estacionario, entonces podemos calcular a partir de ella las fluctuaciones de nuestro sistema en el estado estacionario,

$$\begin{aligned}
 \eta_z^2 &= \frac{1}{\Omega} \frac{\langle z(t) \rangle}{\Omega}, & \eta_{x_p}^2 &= \frac{1}{\Omega} \frac{\langle x_p \rangle}{\Omega}, & \eta_c^2 &= \frac{1}{\Omega} \frac{\frac{\langle c \rangle}{\Omega}}{1 + \frac{k_{on}}{k_{off}} \frac{\langle x_p \rangle}{\Omega}}, \\
 \eta_{c_1}^2 &= \frac{1}{\Omega} \frac{\frac{\langle c_1 \rangle}{\Omega}}{1 + \frac{\beta_1}{\beta_2 + k_1} \frac{\langle z \rangle}{\Omega}}, & \eta_{c_2}^2 &= \frac{1}{\Omega} \frac{\frac{\langle c_2 \rangle}{\Omega}}{1 + \frac{\alpha_1}{\alpha_2 + k_2} \frac{\langle x_p \rangle}{\Omega}}. & & (5.16)
 \end{aligned}$$

En estas expresiones podemos ver que dependen del inverso de Ω , ya que cuando un promedio se divide entre Ω corresponde a la concentración determinista, por lo que cuando el sistema se vuelve lo suficientemente grande las fluctuaciones desaparecen. Simular este sistema empleando el algoritmo de Gillespie tiene un gran coste computacional, esto sería aún más si queremos usar este algoritmo para determinar las fluctuaciones de este sistema, por lo que en la siguiente sección usaremos al TFD que hemos encontrado para simular la dinámica del sistema e inferir las fluctuaciones de esta.

5.2. Análisis alternativo

Debido al alto coste computacional de emplear el algoritmo de Gillespie, ahora analizaremos los dos casos de fosforilación/desfosforilación empleando a la aproximación determinista y al teorema de fluctuación disipación (TFD)

$$\begin{aligned} \frac{\partial s_j}{\partial t} &= \sum_i \Gamma_{ji} \left(R_i^{D+}(s_l) - R_i^{D-}(s_l) + \sum_{j_1} \sum_{j_2} \frac{\sigma^2(s_{j_1}, s_{j_2})}{2} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \right), \\ \frac{\partial}{\partial t} \sigma^2(s_{l_1}, s_{l_2}) &= \sum_i \left(\Gamma_{l_1 i} \Gamma_{l_2 i} \frac{(R_i^{D+}(s_l) + R_i^{D-}(s_l))}{\Omega} \right. \\ &\quad + \sum_{j_1} \left(\sigma^2(s_{l_1}, s_{j_1}) \Gamma_{l_2 i} \frac{\partial}{\partial s_{j_1}} + \sigma^2(s_{j_1}, s_{l_2}) \Gamma_{l_1 i} \frac{\partial}{\partial s_{j_1}} \right) (R_i^{D+}(s_l) - R_i^{D-}(s_l)) \\ &\quad \left. + \sum_{j_1} \sum_{j_2} \sigma^2(s_{j_1}, s_{j_2}) \frac{\Gamma_{l_1 i} \Gamma_{l_2 i}}{2\Omega} \frac{\partial^2}{\partial s_{j_1} \partial s_{j_2}} (R_i^{D+}(s_l) + R_i^{D-}(s_l)) \right), \end{aligned} \quad (5.17)$$

donde $R_i^{D+}(s_l)$ y $R_i^{D-}(s_l)$ denotan a las tasas deterministas, estas tasas se calculan con ayuda de las matrices de coeficientes estequiométricos α y β

$$R_i^{D+}(s_l) = \prod_j s_j^{\alpha_{ij}}, \quad R_i^{D-}(s_l) = \prod_j s_j^{\beta_{ij}}, \quad (5.18)$$

resolveremos las ecuaciones diferenciales que están arriba, como el sistema es no-lineal escribiremos un programa de Python para resolverlo de manera numérica. De la varianza se puede inferir las fluctuaciones del sistema, el programa que emplearemos será muy similar al presentado en el capítulo de fluctuaciones por lo que para más detalles de esta revisar la carpeta de Drive.

5.2.1. Caso I

Las reacciones de este caso son las siguientes,



dado que en este sistema la proteína Z está experimentando cambios en el tiempo, más específicamente está siendo sintetizada por una función externa, esto lo capturamos en la parte derecha de la ecuación anterior con la función $k(t)$. Con ayuda de todas las reacciones anteriores encontramos la matriz de coeficientes estequiométricos y la matriz estequiométrica

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \Gamma_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

con ayuda de estas calculamos las tasas de reacción del sistema, y obtenemos

$$\begin{aligned} R_1^{D+} &= k(t), & R_1^{D-} &= \delta \hat{z}, \\ R_2^{D+} &= k_1 \hat{x} \hat{z}, & R_2^{D-} &= 0, \\ R_3^{D+} &= k_2 \hat{x}_p \hat{y}, & R_3^{D-} &= 0, \\ R_4^{D+} &= k_{on} \hat{x}_p \hat{p}, & R_4^{D-} &= k_{off} \hat{c}. \end{aligned} \quad (5.19)$$

Hemos denotado a $\{\hat{z}, \hat{x}, \hat{x}_p, \hat{y}, \hat{p}, \hat{c}, \hat{c}_1, \hat{c}_2\}$ como la concentración de moléculas bioquímicas de $\{Z, X, X_p, Y, P, X_pP\}$ respectivamente. Con ayuda de estos elementos escribimos nuestro programa de Python (*afosl.ipynb*), para encontrar la dinámica del sistema, los resultados se pueden observar en la figura 5.4.

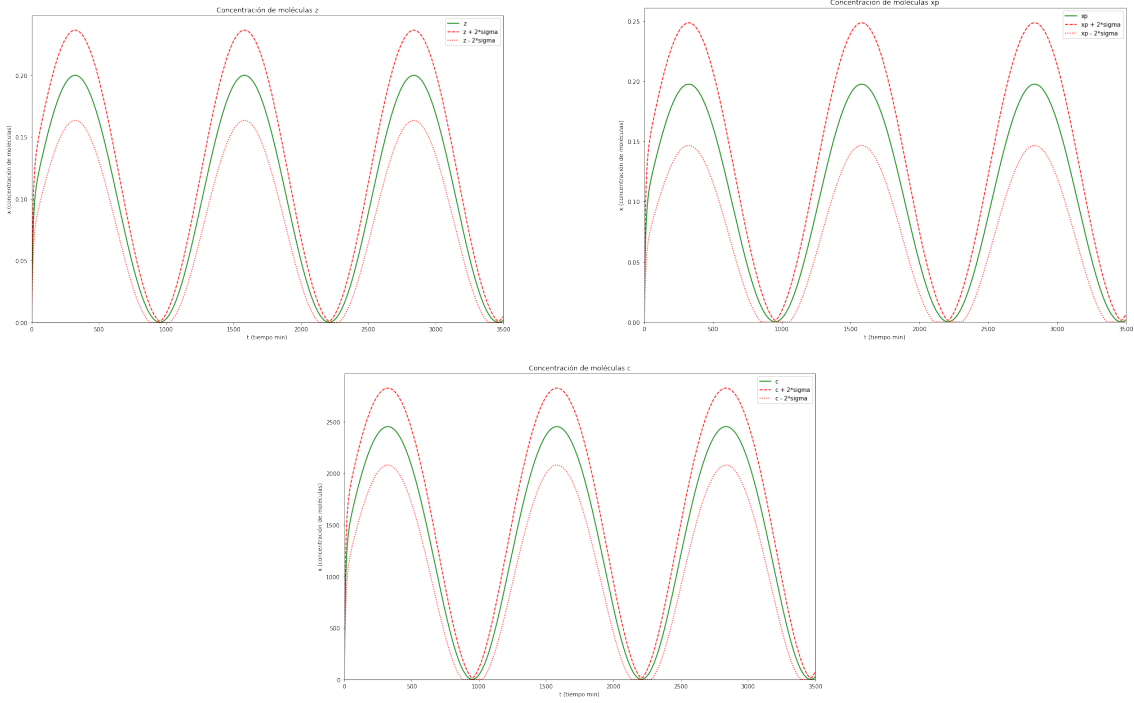


Figura 5.4: Fosforilación con fluctuaciones. Se modela la dinámica del proceso de fosforilación/-desfosforilación a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$.

Solo graficamos a tres de las variables del sistema, ya que las otras dependen de estas, podemos observar que en las tres figuras la línea verde corresponde a la concentración de moléculas y esta se encuentran entre dos bandas marcadas por las líneas rojas punteadas que representan las fluctuaciones del sistema, por lo que la mayoría de las simulaciones estocásticas que llegasen a realizarse se encontraría entre estas dos bandas, de nuevo se encuentra una dinámica oscilatoria en el sistema, tanto en z y c se observa claramente que el sistema tiene de fluctuaciones, esto es aún más marcado en x_p , el cual tiene una banda mucho más ancha. Con estos resultados se observa que hemos cuantificado las fluctuaciones durante toda la dinámica del sistema.

5.2.2. Caso II

Ahora analizaremos el caso II en el cual existen estados intermedios entre la fosforilación y desfosforilación, las reacciones bioquímicas son





usaremos estas reacciones para construir un modelo estocástico del sistema, por lo que ahora la matriz estequiométrica y las matrices de coeficientes estequiométricos del sistema serán

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\Gamma_{ij} = \begin{pmatrix} 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

con ayuda de estas calculamos las tasas de reacción del sistema, obteniendo

$$\begin{aligned} R_1^{D+} &= k(t), & R_1^{D-} &= \delta \hat{z}, \\ R_2^{D+} &= \beta_1 \hat{x} \hat{z}, & R_2^{D-} &= \beta_2 \hat{c}_1, \\ R_3^{D+} &= k_1 \hat{c}_1, & R_3^{D-} &= 0, \\ R_4^{D+} &= \alpha_1 \hat{x}_p \hat{y}, & R_4^{D-} &= \alpha_2 \hat{c}_2, \\ R_5^{D+} &= k_2 \hat{c}_2, & R_5^{D-} &= 0, \\ R_6^{D+} &= k_{on} \hat{x}_p \hat{p}, & R_6^{D-} &= k_{off} \hat{c}. \end{aligned}$$

Hemos denotado a $\{\hat{z}, \hat{x}, \hat{x}_p, \hat{y}, \hat{p}, \hat{c}, \hat{c}_1, \hat{c}_2\}$ como la concentración de moléculas bioquímicas de $\{Z, X, X_p, Y, P, X_p p, XZ, X_p Y\}$. Usando estos elementos en el sistema de ecuaciones diferenciales que presentamos al principio de esta sección escribimos un programa de Python (*(afos2.ipynb)*) para que lo resuelva de manera numérica, los resultados se pueden observar en la figura 5.5.

Nuevamente solo hemos graficado a tres de las variables del sistema, ya que las otras dependen de estas, la línea verde corresponde al promedio de la concentración de moléculas y estas se encuentran entre dos bandas marcadas por las líneas rojas punteadas que representan las fluctuaciones del sistema, por lo que la mayoría de las simulaciones estocásticas que llegasen a realizarse se encontraría entre estas dos bandas, de nuevo se encuentra una dinámica oscilatoria en el sistema, tanto en z y c se observa claramente que el sistema tiene fluctuaciones, esto es aún más marcado en x_p , el cual tiene una banda mucho más ancha, aunque en comparación con lo que se obtuvo en el caso I este sistema tiene unas bandas más pequeñas, quizá debido a que el sistema tiende a tener una más alta concentración de cada una de las especies bioquímicas involucradas, otro resultado importante de esto se debe a que a comparación con el sistema determinista las fluctuaciones pueden llegar a producir una dinámica muy similar cuando hay retroactividad, ya que elegimos un conjunto de parámetros en donde hay retroactividad.

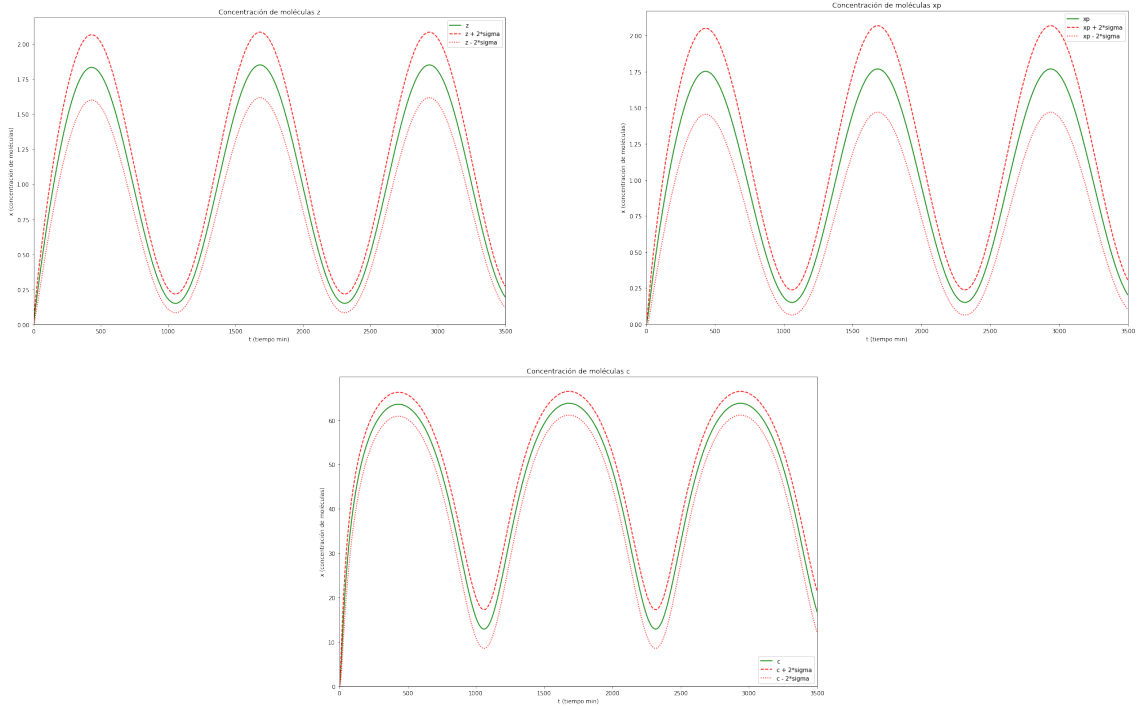


Figura 5.5: Fosforilación con procesos intermedios y fluctuaciones. Se modela la dinámica del proceso de fosforilación/desfosforilación con procesos intermedios a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 50$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$.

Tal como se ha tratado en el trabajo de Del Vecchio vamos a modelar al sistema usando también un conjunto de parámetros en donde se pierde la retroactividad del sistema, los resultados están en la figura 5.6. Podemos notar que la retroactividad se pierde en el sistema, solo después de mucho tiempo la variable x_p y c empiezan a oscilar pero con periodos y amplitudes pequeñas, además podemos observar que este conjunto de variables hace que nuestro sistema tenga fluctuaciones mucho más grandes, lo que haría también que en la contra parte estocástica de estos sistemas no se noten las oscilaciones, otras característica relevante que nos gustaría comentar es el hecho que cuando la retroactividad es menor las fluctuaciones en el sistema tienden a ser más pequeñas, esto se debe al hecho que la ley de acción de masas es un formalismo que describe la interacción entre diferentes especies, por lo que cuando los parámetros son lo más o menos del mismo orden hay más interacción, lo que equivaldría a decir que se transmiten más fluctuaciones entre los módulos del sistema (las especies bioquímicas del sistema). Esto último que acabamos de comentar se puede observar más claramente si recordamos que la retroactividad con correcciones estocásticas está definida como

$$\mathcal{R}_i^j = -\mathcal{R}_i^{jD} \left(\frac{1 + \frac{\partial Q R_i^S}{\partial \Omega_i} \left(\frac{\partial Q R_i^D}{\partial \Omega_i} \right)^{-1}}{1 + \frac{\partial Q R_i^S}{\partial w_i} \left(\frac{\partial Q R_i^D}{\partial w_i} \right)^{-1}} \right),$$

donde \mathcal{R}_i^{jD} es la retroactividad determinista, aunque esta es una cantidad que ya hemos calculado, como estamos interesados particularmente en conocer la retroactividad en el estado estacionario y como las correcciones estocásticas solo aparecen cuando tenemos tasas de orden 2, pero

como en este caso solo los términos diagonales de la covarianza son diferentes de cero, entonces en el estado estacionario no hay ninguna contribución estocástica a la retroactividad, por lo que es suficiente usar la retroactividad determinista en este sistema, por lo que podemos decir que las contribuciones estocásticas serán las mismas que cuando el módulo aún no se ha conectado.

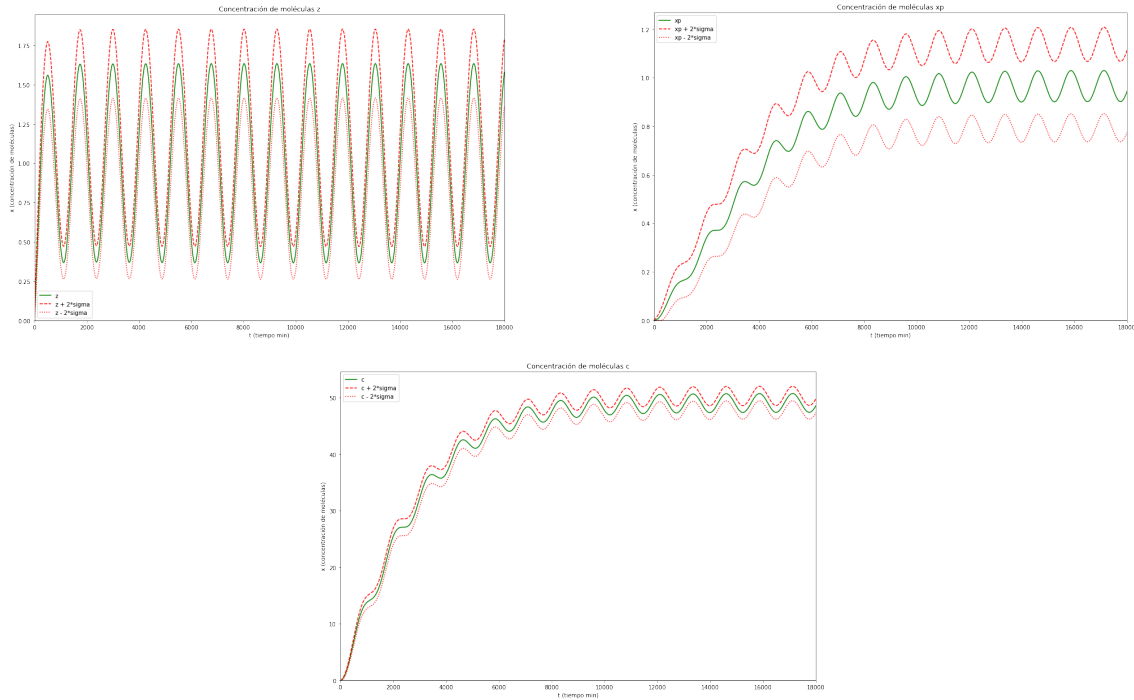


Figura 5.6: Fosforilación con procesos intermedios, mayor retroactividad y fluctuaciones. Se modela la dinámica del proceso de fosforilación/desfosforilación con procesos intermedios a partir de la aproximación determinista y el TFD, las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han empleado los siguientes parámetros: $p_{Tot} = 100$, $k_{off} = k_{on} = 10$, $\delta = 0,01$, $k_1 = k_2 = 0,01$, $\alpha_1 = \beta_1 = 0,01$, $\alpha_2 = \beta_2 = 10$, $x_{Tot} = y_{Tot} = 1500$, $\Omega = 150$.

Una gran ventaja de emplear a las ecuaciones que escribimos al principio de la sección es que el tiempo de computación del sistema se reduce dramáticamente, siendo cientos de veces más rápido y obteniendo resultados satisfactorios, ya que en esta aproximación se supone que el tamaño del ensamble es infinito (por así decirlo), y tan solo depende del tamaño del sistema Ω , por lo que conforme Ω se vuelve grande esta aproximación tiende a ser exacta.

Capítulo 6

Función de Hill

Para explicar la dinámica detrás de los procesos enzimáticos o de unión de algunas moléculas a otras se hace uso de las funciones de entrada, se han presentado algunas en la literatura [4], uno de los más usados es la función de Hill, un caso particular de esta función son las ecuaciones de Michaels-Menten el cual es un modelo bastante sencillo para explicar la unión de una enzima a una proteína (o un proceso similar), incluso puede usarse en modelos booleanos al tomar el límite del coeficiente de Hill tendiendo al infinito, esta se vuelve una función escalón. [4]

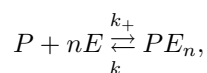
La mayoría de los modelos que se estudian son deterministas, por lo que para encontrar la función de Hill se hace uso de una derivación determinista para luego usarla, sin embargo, dada que varios sistemas son pequeños, tales como los módulos de transcripción/traducción (MTT) y donde se hace uso de la función de Hill (para más detalles ver capítulo 2), entonces es necesaria una descripción estocástica. Por lo que el usar funciones deterministas dentro de un proceso estocástico puede llegar a ser contraproducente y no capturar correctamente la dinámica del sistema como se ve en [32], en este capítulo nos daremos a la tarea de derivar y explicar la versión estocástica de la función de Hill. Pero para llegar a eso primero abordaremos el caso determinista, aunque la explicación este dada en términos de enzimas y proteínas, este sistema se puede extrapolar a otros en los cuales ocurren reacciones similares.

6.1. Derivación determinista

Una de las razones por la cual se usa la función de Hill es para conocer la proporción de activación de alguna cantidad en términos de otra, como veremos a continuación esta ayuda a describir el número de proteínas activas en términos del número de enzimas presentes en el sistema. A continuación, derivaremos la función de Hill desde la ley de acción de masas, esto desde una perspectiva determinista, por lo que el sistema estará descrito por un conjunto de ecuaciones diferenciales, bajo la suposición que el sistema se encuentra en equilibrio se encuentra la función de Hill como veremos a continuación.

6.1.1. Función de Hill usual

Lo primero será derivar la función de Hill que es bien conocida en la literatura, para ello suponemos que tenemos una proteína P a la cual se le pueden unir (o despegar) n enzimas E de tal manera que cuando se han unido se dice que la proteína se ha activado, es decir nuestro sistema esta descrito por el siguiente proceso reversible



la dinámica estará descrita por las siguientes ecuaciones diferenciales, con p y e siendo las concentraciones de la proteína y de la enzima respectivamente, y s la concentración de PE_n (proteínas activas)

$$\begin{aligned}\frac{dp}{dt} &= k_-s - k_+pe^n, \\ \frac{ds}{dt} &= -k_-s + k_+pe^n.\end{aligned}\tag{6.1}$$

Cuando se está en equilibrio se cumple la siguiente relación

$$K^n = \frac{k_-}{k_+} = \frac{pe^n}{s},\tag{6.2}$$

la concentración del número de proteínas activas será

$$H = \frac{\text{proteínas activas}}{\text{proteínas totales}} = \frac{s}{p + s},\tag{6.3}$$

de la relación (6.2), despejamos s , por lo que ahora nuestra función H es

$$H = \frac{s}{p + s} = \frac{\frac{pe^n}{K^n}}{p + \frac{pe^n}{K^n}} = \frac{e^n}{K^n + e^n}.\tag{6.4}$$

Esta es justo la función de Hill, la n se le conoce como el coeficiente de Hill, cuando el valor de la n es muy grande nuestra función se parece a una función escalón. En la figura 6.1 se puede observar cómo se comporta la función de Hill para el caso de un activador y diferentes valores de n .

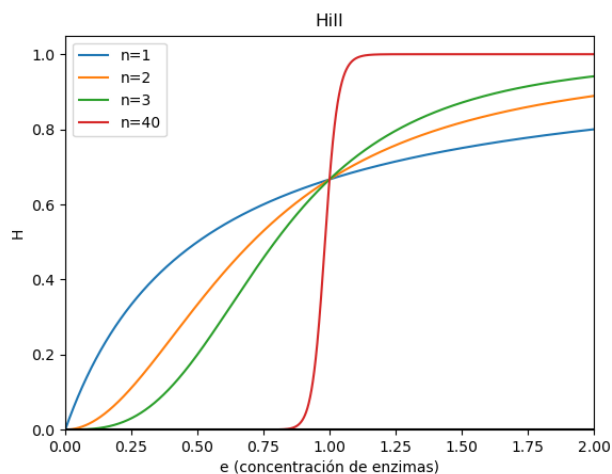


Figura 6.1: Función de Hill determinista. En la figura se gráfica a la función de Hill determinista para un activador a diferentes valores de n y con $K = 1$.

La función de Hill es una concentración por lo que nunca supera el valor de 1, tal como se observa en la figura anterior, mientras el valor de n se hace cada vez más grande la función se parece más a una función escalón. El proceso que hemos descrito anteriormente se podría generalizar como más adelante veremos, para saber el número de proteínas desactivas, lo único que tenemos que hacer es definir la siguiente función

$$D = 1 - H.\tag{6.5}$$

En nuestro trabajo H sería una función de activación, mientras que D sería una de desactivación. En la figura 6.2 se puede observar cómo se comporta la función de Hill en el caso en el que esta actúa como supresora.

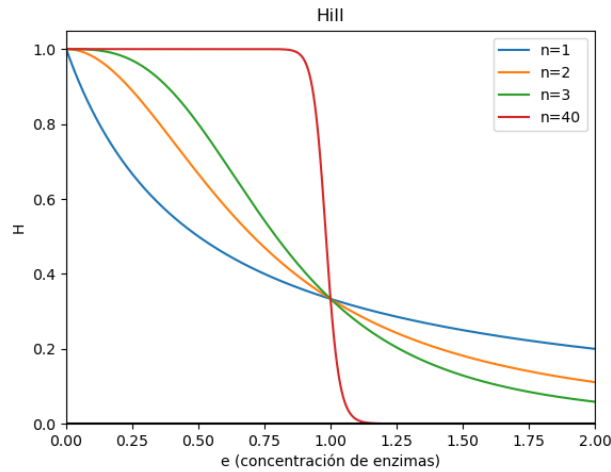


Figura 6.2: Función de Hill determinista. En la figura se gráfica a la función de Hill determinista para un supresor a diferentes valores de n y con $K = 1$.

La función de Hill tal y como las hemos presentado en hasta ahora es la más conocida y usada los procesos bioquímicos, el desarrollo que hicimos para derivar esta función puede llegar a ser una descripción muy poco realista de lo que verdad ocurre en los sistemas bioquímicos, sin embargo, sigue empleándose porque uno de los objetivos es solo describir como es la concentración de las proteínas unidas con enzimas a través de una sigmoide, en muchos casos el coeficiente de Hill se vuelve un valor decimal, por lo que entonces el modelo que hemos presentado tan solo describe la unión enzimática, más adelante veremos cómo resolver alguna de estas cuestiones relacionadas al coeficiente de Hill.

6.1.2. Función de Hill con varios tipos de enzimas

En la sección anterior hemos derivado la función de Hill en el caso en el que solo un tipo de enzima se une a la proteína, sin embargo, existe el caso en el que varios tipos de enzimas se pueden unir, este tipo de sistemas los encontraremos en las redes genéticas, cuando en un MTT participen varias enzimas para su activación o desactivación.

Tipos de enzimas		
Caso	Enzima 1	Enzima 2
1	Activador	Activador
2	Represor	Represor
3	Represor	Activador

Tabla 6.1: Tabla de tipo de enzimas. En el caso que dos tipos de enzimas se unan a una proteína, se esquematiza los posibles tipos de enzimas.

Empecemos por analizar un sistema en el cual dos enzimas de diferente tipo (activadora o represora) se unen a una proteína (algún número de estas n_1 o n_2) en este caso analizaremos tres posibles casos como en la tabla anterior.

Caso 1

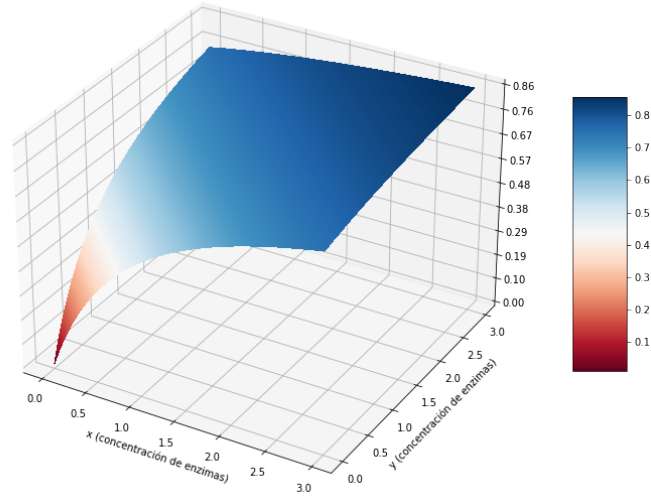


Figura 6.3: Caso 1: función de Hill determinista. Se gráfica el valor de la función de Hill cuando las dos enzimas son activadoras, con $K_1 = K_1 = 1$ y $n_1 = n_2 = 1$.

Primero analizaremos el caso 1, cuando las dos enzimas actúan como activadores, por lo que tendríamos el siguiente conjunto de reacciones



la dinámica del sistema estará descrita por las siguientes ecuaciones diferenciales, con p , e_1 y e_2 siendo las concentraciones de la proteína y de las enzimas respectivamente, mientras que s_1 y s_2 la concentración de PE_{1n_1} y PE_{2n_2} , estas cantidades son las concentraciones de proteínas activas,

$$\begin{aligned} \frac{dp}{dt} &= k_{1-} s_1 - k_{1+} p e_1^{n_1} + k_{2-} s_2 - k_{2+} p e_2^{n_2}, \\ \frac{ds_1}{dt} &= -k_{1-} s_1 + k_{1+} p e_1^{n_1}, \\ \frac{ds_2}{dt} &= -k_{2-} s_2 + k_{2+} p e_2^{n_2}. \end{aligned} \tag{6.6}$$

Cuando se está en equilibrio se cumplen las siguientes relaciones

$$\begin{aligned} K_1^{n_1} &= \frac{k_{1-}}{k_{1+}} = \frac{p e_1^{n_1}}{s_1}, \\ K_2^{n_2} &= \frac{k_{2-}}{k_{2+}} = \frac{p e_2^{n_2}}{s_2}, \end{aligned} \tag{6.7}$$

en este caso la función de Hill se definirá como la razón entre proteínas activas y todas las proteínas del sistema

$$H = \frac{\text{proteínas activas}}{\text{proteínas totales}} = \frac{s_1 + s_2}{p + s_1 + s_2} = \frac{\frac{e_1^{n_1}}{K_1^{n_1}} + \frac{e_2^{n_2}}{K_2^{n_2}}}{1 + \frac{e_1^{n_1}}{K_1^{n_1}} + \frac{e_2^{n_2}}{K_2^{n_2}}}, \quad (6.8)$$

esta función de Hill es el caso en el cual las dos enzimas actúan como activadoras, podemos ver que si alguna de las dos variables e_1 o e_2 valen cero recuperamos a la función de Hill para un solo activador. Graficamos esta última expresión que obtuvimos para un caso en particular, el resultado se puede ver en la figura 6.3. En este caso que las dos enzimas actúan como activadoras, conforme aumenta la concentración de las dos nuestra función se va acercando cada vez más a 1.

Caso 2

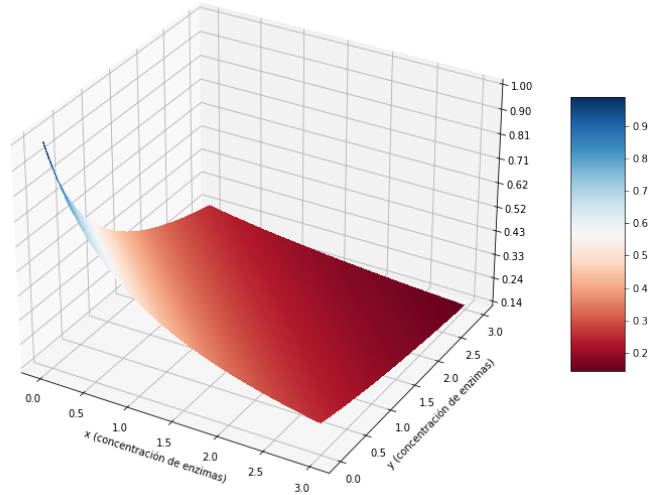


Figura 6.4: Caso 2: función de Hill determinista. Se gráfica el valor de la función de Hill cuando las dos enzimas son supresoras, con $K_1 = K_2 = 1$ y $n_1 = n_2 = 1$.

Ahora analizaremos el caso en el que las dos enzimas actúan como represoras, por lo que tendremos unas reacciones similares a lo siguiente



con p , e_1 y e_2 siendo las concentraciones de la proteína y de las enzimas respectivamente, mientras que s_1 y s_2 la concentración de PE_{1n_1} y PE_{2n_2} , estas cantidades son las concentraciones de proteínas desactivas. De estas reacciones se derivarán un conjunto de ecuaciones diferenciales similar a las que calculamos para el caso anterior, por lo que cuando las derivadas temporales sean cero se encontrarían las siguientes relaciones

$$\begin{aligned} K_1^{n_1} &= \frac{k_{1-}}{k_{1+}} = \frac{pe_1^{n_1}}{s_1}, \\ K_2^{n_2} &= \frac{k_{2-}}{k_{2+}} = \frac{pe_2^{n_2}}{s_2}, \end{aligned} \quad (6.9)$$

en este caso la función de Hill también se definirá como la razón entre proteínas activas y todas las proteínas del sistema, pero en este caso solo conocemos a las desactivas, por lo que restaremos a las desactivas de las totales del sistema

$$\begin{aligned}
 H &= \frac{\text{proteínas activas}}{\text{proteínas totales}} = \frac{\text{proteínas totales} - \text{proteínas desactivas}}{\text{proteínas totales}} \\
 &= \frac{p}{p + s_1 + s_2} = \frac{1}{1 + \frac{e_1^{n_1}}{K_1^{n_1}} + \frac{e_2^{n_2}}{K_2^{n_2}}}.
 \end{aligned} \tag{6.10}$$

Podemos observar que en el caso en el que e_1 o e_2 valen cero recuperamos a la función de Hill para un solo supresor. Graficamos esta última expresión que obtuvimos para un caso en particular, el resultado se puede observar en la figura 6.4. En esta imagen las enzimas actúan como supresoras, por lo que en las regiones en el cual la concentración de las enzimas es grande el valor de la función se acerca a cero.

Caso 3

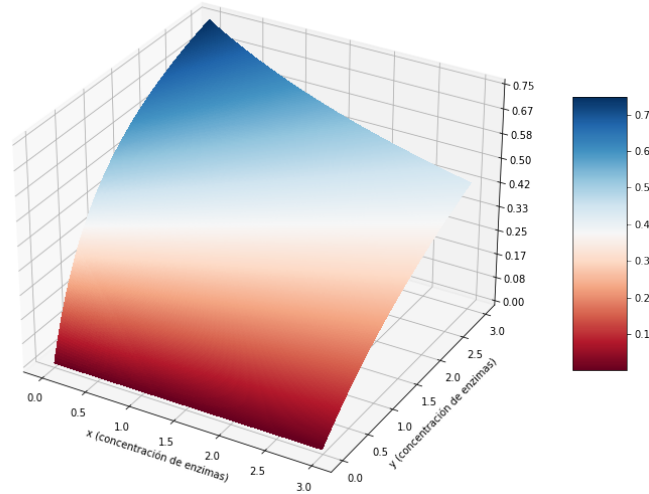


Figura 6.5: Caso 3: función de Hill determinista. Se gráfica el valor de la función de Hill cuando una de las enzimas es activadora y la otra supresora, con $K_1 = K_2 = 1$ y $n_1 = n_2 = 1$.

Para el otro caso suponemos que la enzima E_1 actúa como supresora mientras que E_2 como activadora. Con p , e_1 y e_2 siendo las concentraciones de la proteína y de las enzimas respectivamente, mientras que s_1 y s_2 la concentración de PE_{1n_1} y PE_{2n_2} , estas cantidades son las concentraciones de proteínas desactivas y activas respectivamente. De estas reacciones se derivarán un conjunto de ecuaciones diferenciales similar a las que calculamos para el caso anterior, por lo que cuando las derivadas temporales sean cero se encontrarían las siguientes relaciones

$$\begin{aligned}
 K_1^{n_1} &= \frac{k_{1-}}{k_{1+}} = \frac{pe_1^{n_1}}{s_1}, \\
 K_2^{n_2} &= \frac{k_{2-}}{k_{2+}} = \frac{pe_2^{n_2}}{s_2},
 \end{aligned} \tag{6.11}$$

en este caso la función de Hill también se definirá como la razón entre proteínas activas y todas las proteínas del sistema, este caso conocemos a las desactivas, por lo que restaremos a las desactivas

de las totales del sistema, obtendríamos la siguiente función de Hill

$$H = \frac{\text{proteínas activas}}{\text{proteínas totales}} = \frac{s_2}{p + s_1 + s_2} = \frac{\frac{e_2^{n_2}}{K_2^{n_2}}}{1 + \frac{e_1^{n_1}}{K_1^{n_1}} + \frac{e_2^{n_2}}{K_2^{n_2}}}. \quad (6.12)$$

podemos ver que si e_1 vale cero recuperamos a la función de Hill para un solo activador, mientras que si e_2 vale cero nuestra función de Hill vale cero, es decir, que ahora en nuestro sistema solo existe una supresión, por lo que no puede activarse ninguna proteína. Graficamos esta última expresión que obtuvimos para un caso en particular, el resultado se puede observar en la figura 6.5. En esta imagen una de las enzimas actúa como activadora y otra como supresora por lo que existen regiones en la cual la función se acerca a 1 y en otras a 0.

Es posible derivar algunas expresiones más complejas para este tipo de reacciones, aunque esto último no lo analizaremos a detalle. Nosotros estamos particularmente interesados en redes de regulación genética, por lo que para considerar la actividad de la región promotora se puede realizar la correspondiente derivación considerando todas las reacciones bioquímicas involucradas, de tal manera que si se considera un sistema en la cual existen m diferentes enzimas y que actúan como activadoras o supresoras, por lo que tendríamos la siguiente función de Hill

$$H_q = \frac{\beta_0 K_0 + \sum_j \beta_{1j} \left(\frac{p_j^{n_j}}{K_j^{n_j}} \right)^q}{1 + K_0 + \sum_j \beta_{2j} \frac{p_j^{n_j}}{K_j^{n_j}}}, \quad (6.13)$$

aquí $q = \{0, 1\}$, toma el valor de 0 en el caso en el que la enzima actúa como supresora, mientras que toma el valor de 1 cuando la enzima actúa como activadora, esta expresión es una generalización de la función de Hill, una expresión similar se puede encontrar en [4], sin embargo, ahí solo se trata el caso en el cual todas las n_j (coeficientes de Hill) son iguales a 1, K_0 está relacionado con la transcripción basal, las K_i son las constantes de Michelson-Menten, las constantes β_{ij} están relacionadas a la capacidad asociativa de las moléculas. Este tipo de funciones nos será muy útil cuando analizaremos las redes genéticas, particularmente cuando varias proteínas actúan sobre un módulo de transcripción/traducción (MTT) (en el apéndice B explicamos una expresión más general y detallada de cómo sería la función de Hill en un MTT).

6.2. Derivación estocástica

En la sección anterior hemos derivado la función de Hill en un entorno determinista bajo la ley de acción de masas, pero como estamos interesados en sistemas que son pequeños, por lo que ahora derivaremos una equivalente en el entorno estocástico (aunque solo para una función de Hill activadora), usando la ecuación maestra para procesos de vida y muerte multivariable, la cual es una versión estocástica de la ley de acción de masas como hemos explicado.

Cuando se trabajan con sistemas deterministas solo se introduce la función de Hill, mientras que cuando uno analiza la versión estocástica de las redes de regulación genética por lo general se sigue usando la función de Hill determinista como a continuación [8]

$$\frac{\hat{e}^n}{K^n + \hat{e}^n} \rightarrow \frac{\left(\frac{e}{\Omega} \right)^n}{K^n + \left(\frac{e}{\Omega} \right)^n}, \quad (6.14)$$

donde \hat{e} representa las concentraciones deterministas, e representa el número de moléculas, Ω está relacionado al tamaño del sistema y tiene unidades de volumen entre mol, hacer este tipo

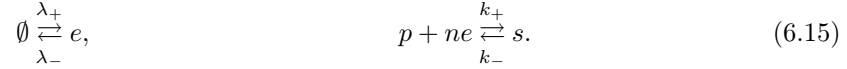
de transición no es del todo correcto [32] ya que podrían no capturar correctamente la dinámica del sistema, por lo que a continuación haremos una derivación de una función de Hill que es estocástica y una que ayudará a cuantificar las fluctuaciones al momento de tomar el límite determinista.

Notamos que en el entorno estocástico no es lo mismo derivar la función de Hill con solo los procesos que presentamos al principio de la sección anterior, que una red que involucre en más procesos, aunque nos interesa esto último ya que en los MTT la función de Hill nos dice como la promotora sintetiza al ARm, analizaremos los dos casos.

6.2.1. Función de Hill en el límite estacionario

Nuestro objetivo es derivar la función de Hill a partir de un proceso estocástico, si recordamos que para la derivación determinista nos fijamos cuando nuestro sistema se encontraba en el límite estacionario, por lo que aquí también haremos una elección similar, aunque para este caso nos fijaremos en la distribución estacionaria, usaremos el método presentado en el capítulo 3 para encontrarla, además, la distribución estacionaria nos permitirá encontrar las fluctuaciones del sistema.

Suponemos que tenemos una proteína p a la cual se le pueden unir n enzimas e formando $s = pe_n$, además pedimos que la encima e llegue o se vaya al sistema y también este puede llegar a degradarse, por lo que nuestro sistema esta descrito por los siguientes procesos reversibles



Notar que las derivaciones que haremos en esta subsección son válidas solo cuando en un sistema se encuentren solo las dos reacciones anteriores. Como ya hemos explicado, la ecuación maestra de una red química se puede obtener a partir del formalismo desarrollado en el capítulo anterior (procesos de vida y muerte multivariable), aquí optaremos por este enfoque, para ello primero definimos a nuestras matrices de coeficientes estequiométricos y la matriz estequiométrica

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ n & 1 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Gamma_{ij} = \begin{pmatrix} 1 & -n \\ 0 & -1 \\ 0 & 1 \end{pmatrix}, \quad (6.16)$$

con ayuda de estas calculamos las tasas de propensión

$$t_i^+(S_j) = k_i^+ \prod_j \frac{S_j!}{\Omega^{\alpha_{ij}} (S_j - \alpha_{ij})!}, \quad t_i^-(S_j) = k_i^- \prod_j \frac{S_j!}{\Omega^{\beta_{ij}} (S_j - \beta_{ij})!}.$$

Por lo que obtenemos

$$\begin{aligned} t_1^+ &= \lambda_+, & t_1^- &= \lambda_- e \frac{1}{\Omega}, \\ t_2^+ &= k_+ p \frac{e!}{(e-n)!} \frac{1}{\Omega^{n+1}}, & t_2^- &= k_- s \frac{1}{\Omega}. \end{aligned} \quad (6.17)$$

(El subíndice 1 corresponde a la reacción de la izquierda de (6.15), mientras que 2 al de la derecha.) Para encontrar la ecuación maestra del sistema se tiene que usar la siguiente relación (tal como hemos explicado anteriormente)

$$\partial_t P(S_j, t) = \Omega \sum_i \left(t_i^-(S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+(S_j) P(S_j, t) + t_i^+(S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^-(S_j) P(S_j, t) \right),$$

sustituyendo todas las cantidades que hemos calculado, finalmente obtenemos nuestra ecuación maestra

$$\begin{aligned} \dot{P}(e, p, s, t) = & k_-(s+1)P(e-n, p-1, s+1, t) - k_+p \frac{e!}{(e-n)!} \frac{1}{\Omega^n} P(e, p, s, t) \\ & + k_+(p+1) \frac{(e+n)!}{e!} \frac{1}{\Omega^n} P(e+n, p+1, s-1, t) - k_-sP(e, p, s, t) \\ & + \lambda_+\Omega P(e-1, p, s, t) - \lambda_-eP(e, p, s, t) + \lambda_-(e+1)P(e+1, p, s, t) - \lambda_+\Omega P(e, p, s, t), \end{aligned} \quad (6.18)$$

si consideramos que nuestro sistema llega muy rápido al equilibrio, entonces es suficiente fijarse en la distribución estacionaria, por lo que la relación anterior será cero, por lo que ahora necesitamos encontrar la distribución estacionaria, dado que tenemos una cantidad conservada, el número total de proteínas activas y desactivas es constante, es decir $N_0 = p + s$ (para hacer contacto con la parte determinista es conveniente definir $N_0 = N_{0d}\Omega$, siendo N_{0d} la condición inicial determinista), si sustituimos obtendremos

$$\begin{aligned} 0 = & k_-(s+1)P(e-n, s+1) - k_+(N_0-s) \frac{e!}{(e-n)!} \frac{1}{\Omega^n} P(e, s) \\ & + k_+(N_0-s+1) \frac{(e+n)!}{e!} \frac{1}{\Omega^n} P(e+n, s-1) - k_-sP(e, s) \\ & + \lambda_+\Omega P(e-1, s) - \lambda_-eP(e, s) + \lambda_-(e+1)P(e+1, s) - \lambda_+\Omega P(e, s). \end{aligned} \quad (6.19)$$

Siguiendo con el método que presentamos anteriormente obtendremos las siguientes dos relaciones

$$\begin{aligned} k_+(N_0-s+1) \left\langle \frac{e!}{(e-n)!} \right\rangle_e \frac{1}{\Omega^n} P_s(s-1) - k_-sP_s(s) &= 0, \\ \lambda_+\Omega P_e(e-1) - \lambda_-eP_e(e) &= 0. \end{aligned} \quad (6.20)$$

Resolviendo cada una obtenemos

$$\begin{aligned} P_s(s) &= \left(\frac{1}{\Omega^n} \frac{k_+}{k_-} \left\langle \frac{e!}{(e-n)!} \right\rangle_e \right)^s \frac{N_0!}{s!(N_0-s)!} P_s(0), \\ P_e(e) &= \left(\Omega \frac{\lambda_+}{\lambda_-} \right)^e \frac{1}{e!} P_e(0). \end{aligned} \quad (6.21)$$

Combinando estas dos expresiones, encontrando las constantes de normalización y el valor exacto de $\left\langle \frac{e!}{(e-n)!} \right\rangle_e$ (promedio, pero solo con respecto a e) finalmente tendremos la distribución estacionaria

$$P(e, p, s) = \frac{1}{\left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^{N_0}} \left(\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^s \frac{N_0!}{s!(N_0-s)!} \frac{1}{e!} \left(\Omega \frac{\lambda_+}{\lambda_-}\right)^e e^{-\Omega \frac{\lambda_+}{\lambda_-}} \delta_{p, N_0-s}. \quad (6.22)$$

Esta distribución es la multiplicación de una de Poisson $\left(\frac{1}{e!} \left(\Omega \frac{\lambda_+}{\lambda_-}\right)^e e^{-\Omega \frac{\lambda_+}{\lambda_-}}\right)$ y una variante de una binomial $\left(\frac{1}{\left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^{N_0}} \left(\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^s \frac{N_0!}{s!(N_0-s)!}\right)$, el proceso de Poisson corresponde a la llegada de la enzima al sistema, mientras que la variante de la binomial nos habla el cómo se van uniendo y separando las proteínas y enzimas. Con esta última ecuación calculamos el promedio de la variable s y e ,

$$\langle s \rangle = N_0 \frac{\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n}{1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n}, \quad \langle e \rangle = \Omega \frac{\lambda_+}{\lambda_-}. \quad (6.23)$$

Si recordamos que la función de Hill también puede ser calculada como $H = \frac{\langle s \rangle}{\langle s+p \rangle} = \frac{\langle s \rangle}{N_0}$, apoyándonos de las dos relaciones anteriores tenemos

$$H = \frac{\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n}{1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n} = \frac{\frac{1}{\Omega^n} \langle e \rangle^n}{\frac{k_-}{k_+} + \frac{1}{\Omega^n} \langle e \rangle^n} = \frac{\hat{e}^n}{K^n + \hat{e}^n}, \quad (6.24)$$

con esto recuperamos la función de Hill determinista, con $K^n = \frac{k_-}{k_+}$ y $\hat{e} = \frac{\langle e \rangle}{\Omega}$, con este resultado nos damos cuenta que la función de Hill determinista también puede llegar a ser derivada desde el formalismo estocástico, más aún, también podemos calcular como fluctúa nuestra función el estado estacionario, para esto calculamos las fluctuaciones de $\frac{s}{N_0}$, ya que esta variable está estrechamente relacionada con la función de Hill. Primero calculamos el promedio de $\left(\frac{s}{N_0} \right)^2$ con ayuda de la distribución estacionaria (6.22), con lo que obtenemos

$$\left\langle \left(\frac{s}{N_0} \right)^2 \right\rangle = \frac{N_0 \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n + \left(N_0 \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n \right)^2}{N_0^2 \left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n \right)^2}, \quad (6.25)$$

con ayuda de esto calculamos la varianza, obteniendo

$$\sigma^2 \left(\frac{s}{N_0} \right) = \left\langle \left(\frac{s}{N_0} \right)^2 \right\rangle - \left\langle \frac{s}{N_0} \right\rangle^2 = \frac{N_0 \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n}{N_0^2 \left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n \right)^2}, \quad (6.26)$$

las fluctuaciones de $\frac{s}{N_0}$ serán:

$$\eta \left(\frac{s}{N_0} \right) = \frac{1}{\Omega} \sqrt{\sigma^2 \left(\frac{s}{N_0} \right)} = \frac{1}{\Omega} \sqrt{\frac{H(1-H)}{\Omega N_0^d}}, \quad (6.27)$$

por lo que las fluctuaciones de la función de Hill serán:

$$\eta(H) = \frac{1}{\Omega} \sqrt{\sigma^2(H)} = \frac{1}{\Omega} \sqrt{\frac{H(1-H)}{\Omega N_0^d}}, \quad (6.28)$$

hemos usado $N_0 = N_{0d}\Omega$, en esta expresión se ve que cuando el sistema se vuelve lo suficientemente grande las fluctuaciones desaparecen, debido a que aparece Ω en el denominador, también cuando H se acerca al valor de 0 o 1 las fluctuaciones se vuelven pequeñas, es decir que debido a que hay pocas enzimas e (el valor de H se acerca a 0) por lo que pocas se pueden unir a la proteína creando una pequeña fluctuación del número de proteínas activas, mientras que cuando hay muchas enzimas e (el valor de H se acerca a 1) el sistema tiene tantas proteínas activas pero también muchas proteínas disponibles para unirse a las proteínas por lo que las fluctuaciones se vuelven pequeñas.

A continuación realizamos una simulación del proceso del cual partimos, para ello empleamos el algoritmo de Gillespie [22]. En la figura 6.6 se ve claramente la estructura de una función de Hill, en comparación con la figura 6.1, pero esta tiene fluctuaciones durante toda la evolución temporal, una característica muy importante es el hecho que esta función jamás supera el valor de 1, incluso las fluctuaciones no lo hacen, en esta figura variamos el tamaño del sistema, sin embargo podemos hacer múltiples simulaciones empleando solo un tamaño, otras de las consideraciones que hemos escogido es que el número de enzimas al iniciar el proceso es cero, se pueden usar otras cantidades iniciales pero esto no muestra de manera clara la forma de la función de Hill, aunque no hay que perder de vista que lo que importa es el estado estacionario.

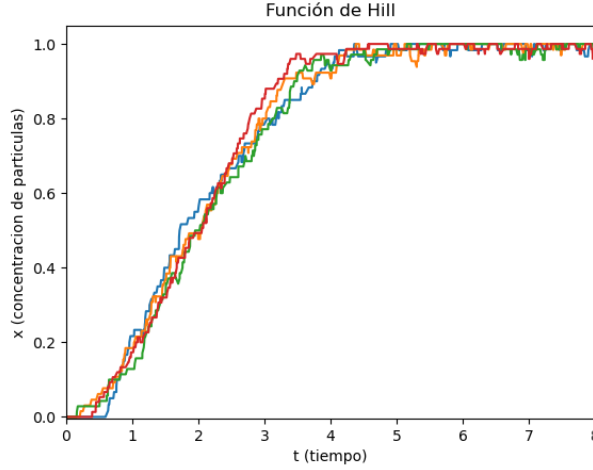


Figura 6.6: Función de Hill estocástica. Se realizó la simulación estocástica de las reacciones de donde se obtiene la función de Hill para diferentes tamaños de sistema $\Omega = \{60, 80\}$ y con los siguientes parámetros: $N_0 = 2\Omega$, $k_+ = k_- = \lambda_- = \lambda_+ = 1$, $n = 2$.

Recordar que esta es la función de Hill que hemos derivado es para un activador, por lo que faltaría calcularla para un represor, para un represor podría ser $D = 1 - H$, decir que para el represor se calcula de esta manera es un tanto incorrecto si no se le da la interpretación correcta, ya que uno debería de realizar los cálculos de manera explícita, sin embargo, si uno los realiza uno llegaría al mismo resulta que el realizar $D = 1 - H$, la distribución estacionaria que se obtendría para el caso de un represor será

$$P(e, p, s) = \frac{1}{\left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^{N_0}} \left(\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^s \frac{N_0!}{s!(N_0 - s)!} \frac{1}{e!} \left(\Omega \frac{\lambda_+}{\lambda_-}\right)^e e^{-\Omega \frac{\lambda_+}{\lambda_-}} \delta_{p, N_0 - s}. \quad (6.29)$$

Esta distribución es la multiplicación de una de Poisson y una variante de una binomial, muy parecida al caso de un activador, en este sistema p representaría a las proteínas activas. Con esta última ecuación calculamos el promedio de la variable p y e ,

$$\langle p \rangle = N_0 \frac{1}{1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n}, \quad \langle e \rangle = \Omega \frac{\lambda_+}{\lambda_-}. \quad (6.30)$$

Si recordamos que la función de Hill también puede ser calculada como $D = \frac{\langle p \rangle}{\langle s + p \rangle} = \frac{\langle p \rangle}{N_0}$, apoyándonos de las dos relaciones anteriores tenemos

$$D = \frac{1}{1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n} = \frac{\frac{k_-}{k_+}}{\frac{k_-}{k_+} + \frac{1}{\Omega^n} \langle e \rangle^n} = \frac{K^n}{K^n + \hat{e}^n}, \quad (6.31)$$

con esto recuperamos la función de Hill determinista para un represor, con $K^n = \frac{k_-}{k_+}$ y $\hat{e} = \frac{\langle e \rangle}{\Omega}$. También podemos calcular como fluctúa nuestra función de Hill en el estado estacionario, para esto calculamos las fluctuaciones de $\frac{s}{N_0}$, ya que esta variable esta estrechamente relacionada con la función de Hill:

$$\left\langle \left(\frac{p}{N_0}\right)^2 \right\rangle = \frac{N_0 \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n + N_0^2}{N_0^2 \left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^2}, \quad (6.32)$$

con ayuda de esto calculamos la varianza, obteniendo

$$\sigma^2 \left(\frac{p}{N_0} \right) = \left\langle \left(\frac{p}{N_0} \right)^2 \right\rangle - \left\langle \frac{p}{N_0} \right\rangle^2 = \frac{N_0 \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n}{N_0^2 \left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-} \right)^n \right)^2}, \quad (6.33)$$

las fluctuaciones de $\frac{p}{N_0}$ serán:

$$\eta \left(\frac{p}{N_0} \right) = \frac{1}{\Omega} \sqrt{\sigma^2 \left(\frac{p}{N_0} \right)} = \frac{1}{\Omega} \sqrt{\sigma^2 \left(\frac{p}{N_0} \right)} = \frac{1}{\Omega} \sqrt{\frac{D(1-D)}{\Omega N_0^d}}, \quad (6.34)$$

por lo que las fluctuaciones de la función de Hill serán:

$$\eta(D) = \frac{1}{\Omega} \sqrt{\sigma^2(D)} = \frac{1}{\Omega} \sqrt{\frac{D(1-D)}{\Omega N_0^d}}, \quad (6.35)$$

hemos usado $N_0 = N_0 d \Omega$. Obtenemos justo el mismo valor de la fluctuación en los dos casos, tanto para el activador como para el represor, notar que esta fluctuación que hemos calculado solo vale para el caso en el que el sistema esté en el estado estacionario, una característica importante es que en el caso en el que $e = 0$ las fluctuaciones se vuelven cero, de igual manera cuando $\lim_{e \rightarrow \infty} e$, en estos puntos las fluctuaciones son más estables, cuando $D = H = \frac{1}{2}$ las fluctuaciones alcanzan su valor máximo.

6.2.2. Estados estacionarios a partir de estados coherentes

A veces el formalismo que hemos usado en la subsección anterior resulta ser un poco larga, por lo que es deseable tener uno más corto, por eso ahora resolveremos las reacciones que aparece en (6.15) en el estado estacionario usando el formalismo de operadores que ya hemos mencionado en el capítulo anterior, este formalismo combinado con lo que se hace en óptica cuántica [33] hace que los cálculos sean más simples. Nos apoyaremos de la ecuación (3.71),

$$\frac{\partial}{\partial t} |G(t)\rangle = \Omega \sum_i \left(\prod_j (a_j^+)^{\beta_{ij}} - \prod_j (a_j^+)^{\alpha_{ij}} \right) \left(k_i^+ \prod_j \left(\frac{a_j^-}{\Omega} \right)^{\alpha_{ij}} - k_i^- \prod_j \left(\frac{a_j^-}{\Omega} \right)^{\beta_{ij}} \right) |G(t)\rangle,$$

recordar que las reacciones son



las matrices de coeficientes estequiométricos son

$$\alpha_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ n & 1 & 0 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

apoyándonos de estas matrices al evaluar obtendremos la siguiente ecuación

$$\frac{\partial}{\partial t} |G(t)\rangle = \Omega \left((a_1^+ - 1) \left(\lambda_+ - \lambda_- \frac{a_1^-}{\Omega} \right) + (a_3^+ - a_2^+ (a_1^+)^n) \left(k_+ \frac{a_2^-}{\Omega} \left(\frac{a_1^-}{\Omega} \right)^n - k_- \frac{a_3^-}{\Omega} \right) \right) |G(t)\rangle,$$

simplificando

$$\frac{\partial}{\partial t} |G(t)\rangle = \left(-\lambda_- (a_1^+ - 1) \left(a_1^- - \frac{\lambda_+}{\lambda_-} \Omega \right) - k_- (a_3^+ - a_2^+ (a_1^+)^n) \left(a_3^- - \frac{k_+}{k_-} a_2^- \left(\frac{a_1^-}{\Omega} \right)^n \right) \right) |G(t)\rangle, \quad (6.36)$$

el corchete de Poisson de estos operadores es

$$[a_i^-, a_j^+] = \delta_{ij} \quad (6.37)$$

los índices corren de 1-3, la ecuación (6.36) tiene una estructura muy similar a los hamiltonianos que aparecen en óptica cuántica, se asemeja mucho a los operadores "squeze" [33], de hecho el hamiltoniano se parece mucho a los usados en los estados coherentes, estos estados coherentes coincidirán con el estado estacionario de nuestro sistema, por lo que una para evitar calcular los promedios de las variables con la distribución estacionaria nos apoyaremos de los estados coherentes, similar a lo que se hace en [33], aplicando este formalismo obtenemos

$$\langle e \rangle_s = \frac{\lambda_+}{\lambda_-} \Omega, \quad \langle s \rangle_s = \frac{1}{\Omega^n} \frac{k_+}{k_-} \langle e \rangle_s^n \langle p \rangle_s, \quad (6.38)$$

definimos $K^n = \frac{k_-}{k_+}$ y además tenemos la siguiente relación $\langle p \rangle_s + \langle s \rangle_s = N_0$, por lo que resolviendo la última relación obtenemos

$$\langle s \rangle_s = \frac{N_0 \left(\frac{\langle e \rangle_s}{\Omega} \right)^n}{K^n + \left(\frac{\langle e \rangle_s}{\Omega} \right)^n}, \quad (6.39)$$

recordando la definición de la función de Hill encontramos la función de Hill para un activador con $\hat{e} = \frac{\langle e \rangle_s}{\Omega}$

$$H = \frac{\hat{e}^n}{K^n + \hat{e}^n}, \quad (6.40)$$

de esta manera recuperamos nuestra función de Hill sin tener que realizar muchos cálculos, solo a partir de los estados coherentes que coinciden con los estados estacionarios de nuestro sistema, lo cual resulta ser muy conveniente y más directo, todos los cálculos que hemos realizado con anterioridad se quedaron ocultos detrás de la metodología usada para estados coherentes.

6.2.3. Variantes de la función de Hill

Aunque ya hemos calculado la función de Hill determinista a partir de un proceso estocástico, siempre esta le pregunta del porque aparece o se introduce la función de Hill determinista en un sistema. Para esto suponemos que existe un sistema en el cual ocurren varias reacciones bioquímicas y además aparecen la siguiente reacción (o una similar)



Se supone que estas reacciones por lo general son rápidas por lo que de aquí se encontraría la función de Hill, sin embargo, cuando en el sistema existen otras reacciones además de (6.41) y el análisis se realiza desde una perspectiva estocástica se tendría que proceder de otra manera. Entonces pensamos en un sistema que contiene a la reacción (6.41) (además de otras) , y suponiendo que nos encontramos en el estado estacionario de la reacción (6.41), tendremos las siguientes dos relaciones (omitimos escribir la ecuación maestra completa ya que esto último dependerá del sistema en particular, aunque más adelante presentaremos un ejemplo para clarificar las ideas que presentaremos)

$$\begin{aligned} k_-(s+1)P(e-n, p-1, s+1) - k_+p \frac{e!}{(e-n)!} \frac{1}{\Omega^n} P(e, p, s) &= 0, \\ k_+(p+1) \frac{(e+n)!}{e!} \frac{1}{\Omega^n} P(e+n, p+1, s-1) - k_-sP(e, p, s) &= 0, \end{aligned} \quad (6.42)$$

igualamos cada una de estas relaciones a cero, ya que esta es una condición que se deriva del hecho de estar en equilibrio, podemos tratar a las relaciones anteriores como ecuaciones algebraicas, de tal manera que bajo unos criterios que describiremos más adelante obtendremos el valor más probable. Lo primero que hacemos es dividir a cada una de estas expresiones por $P(e, p, s, t)$ y reescribimos

$$\begin{aligned} s + 1 &= p \frac{1}{K^n} \frac{e!}{(e-n)!} \frac{1}{\Omega^n} \frac{P(e, p, s)}{P(e-n, p-1, s+1)}, \\ s &= (p+1) \frac{1}{K^n} \frac{(e+n)!}{e!} \frac{1}{\Omega^n} \frac{P(e+n, p+1, s-1)}{P(e, p, s)}, \end{aligned} \quad (6.43)$$

hemos definido $\frac{k_-}{k_+} = K^n$. Para encontrar la función de Hill se pueden hacer varias asunciones sobre las dos relaciones anteriores, estas asunciones se pueden hacer sobre el número de moléculas involucradas y/o la forma de la distribución, las esquematizamos en la siguiente tabla.

Asunciones		
Tamaño del sistema	Distribución	Tipo de la función de Hill
X	X	Estocástico
X	✓	Semi-estocástico
✓	X	Semi-determinista
✓	✓	Determinista

Tabla 6.2: Tabla de asunciones. Las palomitas señalan que tipo de asunción se hizo, a partir de estas asunciones se nombra el tipo de función de Hill.

En esta tabla tenemos 4 tipos de función de Hill, su tipo depende que tipo de asunciones hacemos, en el caso de que no hagamos ninguna tendremos una estocástica, cuando se hacen consideraciones en la distribución de probabilidad tenemos el caso semi-estocástico y así sucesivamente. Hemos elegido a estos nombres de esta manera debido al grado de relevancia de las asunciones en cada formalismo, por ejemplo, en el caso estocástico uno se fija principalmente en la ecuación maestra, la cual es una ecuación diferencial de las distribuciones.

Caso semi-determinista

Primero analizaremos el caso semi-determinista, para ello suponemos que el sistema es lo suficientemente grande, entonces la distribución de probabilidad se pueden aproximar como $\frac{P(e,p,s)}{P(e\pm n, p\pm 1, s\mp 1)} \approx 1$ (esto nos dice que la distribución tiene un pico muy alto), ahora consideraremos dos regiones (relacionados al número de moléculas s) debido a que tenemos dos expresiones en (6.43), además recordar que $p + s = N_0$ y además la función de Hill la definimos como $H = \frac{s}{N_0} = \frac{s}{s+p}$:

- a) Para la primera región suponemos s y N_0 son muy grandes, entonces usamos la primera ecuación de (6.43), en base a esto realizamos las siguientes aproximaciones: $s + 1 \approx s$ y $\frac{P(e,p,s)}{P(e-n, p-1, s+1)} \approx 1$, por lo que tendríamos

$$s = p \frac{1}{K^n} \frac{e!}{(e-n)!} \frac{1}{\Omega^n}, \quad (6.44)$$

sustituyendo en la definición de la función de Hill $H = \frac{s}{s+p}$ obtenemos,

$$H_1 = \frac{e!}{\Omega^n K^n (e-n)! + e!}. \quad (6.45)$$

b) Para la otra región suponemos que p y N_0 son muy grandes, entonces usamos la segunda ecuación de (6.43), a partir de esto realizamos las siguientes aproximaciones: $p+1 \approx p$ (dicho de otra manera s tiene que ser muy pequeño debido a que tenemos la relación $N_0 = s+p$) y $\frac{P(e+n, p+1, s-1)}{P(e, p, s)} \approx 1$, por lo que tendríamos

$$s = p \frac{1}{K^n} \frac{(e+n)!}{e!} \frac{1}{\Omega^n}, \quad (6.46)$$

sustituyendo en la definición de la función de Hill obtenemos,

$$H_2 = \frac{(e+n)!}{\Omega^n K^n e! + (e+n)!}. \quad (6.47)$$

De esta manera hemos obtenido nuestra función de Hill para dos regiones, por lo que podemos decir que una es válida una cuando s es pequeña y la otra cuando es grande, ahora definimos una función de Hill válida (al menos de forma aproximada) para todos los posibles valores de s

$$H_{sd} = \frac{e!}{\Omega^n K^n (e-n)! + e!} + \left(\frac{(e+n)!}{\Omega^n K^n e! + (e+n)!} - \frac{e!}{\Omega^n K^n (e-n)! + e!} \right) \theta \left(\frac{1}{2} - \frac{s}{N_0} \right), \quad (6.48)$$

Se ha introducido una función de Heavinside que está definida como

$$\theta(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}, \quad (6.49)$$

esta nueva función de Hill se apoya de H_1 o H_2 según el valor de s , y así cubrir todos los posibles valores de s , de esta manera hemos obtenido nuestra función de Hill semi-determinista, semi-determinista en el sentido que las consideraciones sobre las probabilidades son unas condiciones casi deterministas. Además, cuando e y/o Ω se vuelven muy grandes

$$H_{sd} \approx H_d = \frac{\frac{e^n}{\Omega^n}}{K^n + \frac{e^n}{\Omega^n}}, \quad (6.50)$$

recuperamos la función de Hill determinista. El uso de la función de Hill determinista dentro de un proceso estocástico se realiza bajo una determinada aproximación tal como lo hemos mostrado, por lo que bajo otras asunciones se debe de considerar alguna de las otras que vamos a presentar.

Caso estocástico

En el caso que en verdad queramos usar una función de Hill estocástica, podemos construir una completamente estocástica, para ello volvemos a las relaciones que se encuentran en (6.43) pero ahora sin hacer ninguna asunción, operamos de una manera similar al caso anterior para abarcar los diferentes valores de s , obtenemos finalmente nuestra función de Hill estocástica

$$H_s = \frac{e!}{\Omega^n K^n (e-n)! \frac{P(e-n, s+1)}{P(e, s)} + e!} + \left(\frac{(e+n)!}{\Omega^n K^n e! \frac{P(e, s)}{P(e+n, s-1)} + (e+n)!} - \frac{e!}{\Omega^n K^n (e-n)! \frac{P(e-n, s+1)}{P(e, s)} + e!} \right) \theta \left(\frac{1}{2} - \frac{s}{N_0} \right). \quad (6.51)$$

Esta es la función de Hill estocástica, se observa que esta depende de la distribución de probabilidades del sistema, usar esta función tal y como la hemos definido dentro de la ecuación maestra resulta ser un poco más complejo.

Caso semi-estocástico

Debido a que es más complicado usar la función de Hill estocástica en un proceso estocástica, vamos a realizar una aproximación sobre las probabilidades. Entonces suponemos lo siguiente

$$\frac{P(e + e_1, s + s_1)}{P(e, s)} \approx e^{-\left(\frac{e_1+s_1}{\Omega}\right)}, \quad (6.52)$$

hemos usado esta aproximación debido a que cuando el sistema es grande, la distribución binomial que aparece en estado estacionario tiende a una de Poisson, además cuando Ω es lo suficientemente grande encontramos la asunción determinista trabajada en la función de Hill semi-determinista. Sustituyendo estas relaciones en la ecuación (6.51) obtenemos nuestra función de Hill semi-estocástica,

$$H_{ss} = \frac{e!}{e^{\left(\frac{n-1}{\Omega}\right)} \Omega^n K^n (e-n)! + e!} + \left(\frac{(e+n)!}{e^{\left(\frac{n-1}{\Omega}\right)} \Omega^n K^n e! + (e+n)!} - \frac{e!}{e^{\left(\frac{n-1}{\Omega}\right)} \Omega^n K^n (e-n)! + e!} \right) \theta \left(\frac{1}{2} - \frac{s}{N_0} \right). \quad (6.53)$$

A esta función la nombramos función de Hill semi-estocástica, debido a que hicimos una consideración solo sobre la distribución de probabilidades, esta función coincide con la semi-determinista cuando $n = 1$ o cuando el valor de Ω es muy grande,

$$\lim_{\Omega \rightarrow \infty} H_{ss} = \frac{\frac{e^n}{\Omega^n}}{K^n + \frac{e^n}{\Omega^n}} = H_d, \quad (6.54)$$

por lo que para realizar las simulaciones de algún determinado sistema estocástico y cuando el tamaño del sistema es pequeño recomendamos usar la función de Hill semi-estocástica, ya que representa de manera más fiel el sistema. Esta función es mucho más sencilla de emplear al momento de realizar simulaciones estocásticas del sistema, debido a que hemos aproximado nuestra distribución de probabilidad a una conocida.

Para ver un poco el comportamiento de las funciones de Hill que hemos calculado H_d , H_{sd} y H_{ss} , hemos realizado unas gráficas que se pueden ver en la figura 6.7, de esta imagen nos fijamos que cuando n está cerca de 1 las tres funciones de Hill se parecen más, lo mismo sucede cuando el valor de K es más pequeño, por lo que en el régimen cuando K es grande estos llegan a diferir de manera significativa, cuando $n = 1$ se observa que las tres funciones son prácticamente iguales, además cuando el valor de e aumenta el valor de las tres son casi idénticas, mientras que a valores pequeños hay una gran diferencia como se puede apreciar en las figuras. Otra característica que comparten es que todas se vuelven prácticamente una función escalón cuando n es grande, aunque el lugar donde está el escalón varía debido a que aparecen factoriales en la semi-determinista y la semi-estocástica.

Más adelante veremos un ejemplo en concreto para entender el cómo aparecen estas funciones de Hill y la elección de cual usar depende de las necesidades del sistema. Las consideraciones que hemos realizado para que aparezcan estas funciones de Hill podrían no llegar a cumplirse totalmente (aún más cuando el tamaño del sistema es pequeño), por lo que si uno quiere tener un proceso lo más realista posible es mejor usar las primeras dos relaciones que presentamos al principio de esta subsección, aunque es posible que surjan otras complicaciones si no se conocen a detalle todos las constantes involucradas, por eso a continuación derivaremos la función de Hill desde otro enfoque.

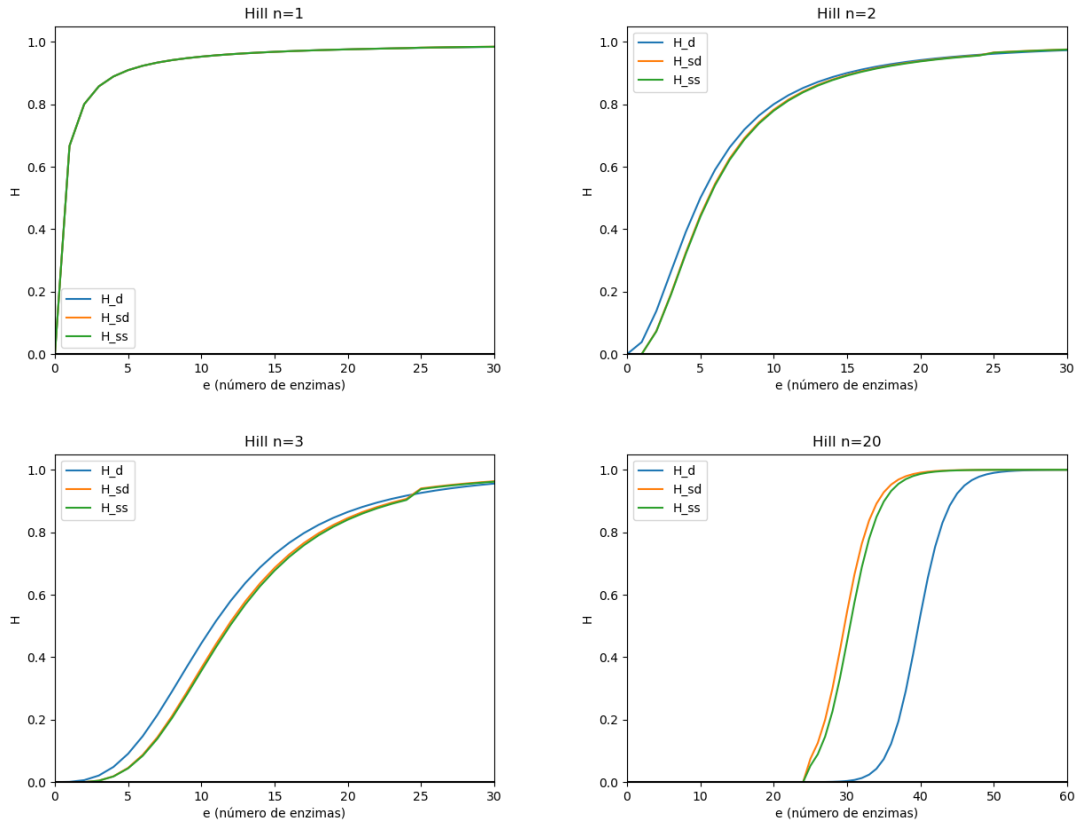


Figura 6.7: Funciones de Hill. En esta figura se gráfica la función de Hill determinista, semi-determinista y semi-estocástica para diferentes valores de n , con $K = 0,01$ y $\Omega = 50$.

6.2.4. Proceso estocástico con una función de Hill

En las subsecciones anterior hemos explicado cómo se derivaría una función de Hill, en donde se considera solo el proceso de unión de enzimas a proteínas, o algún sistema en el cual aparecen esas reacciones, sin embargo, todas las derivaciones se hicieron en pro de poder realizar la simulación estocástica del sistema, pero también es necesaria una metodología que explique cómo introducir adecuadamente una función de Hill en cualquier reacción bioquímica estocástica. Nos apoyaremos de las ideas ya presentadas hasta este punto y del artículo [34], aquí se hace una separación entre variables lentas y rápidas. Para nuestra derivación primero suponemos que la distribución de probabilidad de nuestro sistema se compone de la multiplicación de una parte estacionaria y otra dinámica, de tal manera que tendremos algo de la siguiente manera

$$P(x_l, t) = P_s(x_1, \dots, x_s)P(x_{s+1}, \dots, x_l, t), \quad (6.55)$$

en [34] se usan probabilidades condicionales, pero como nosotros consideramos directamente que una parte del proceso es estacionario se pudo hacer una separación de la parte estacionaria y dinámica, en este caso podemos suponer que es estacionaria porque las concentraciones de las variables asociadas a esta tienden muy rápido y de manera asintótica al valor estacionario (el comportamiento de la función de Hill). De esta manera la ecuación maestra del sistema se dividirá en un parte dinámica y estacionaria que será como a continuación

$$\begin{aligned}
P_s(S_1, \dots, S_s) \frac{dP(S_{s+1}, \dots, S_j, t)}{dt} &= \Omega \sum_i \left(t_i^-(S_j + \Gamma_{ji}) P_s(S_1 + \Gamma_{ii}, \dots, S_s + \Gamma_{si}) P(S_{s+1} + \Gamma_{s+1,i}, \dots, S_j + \Gamma_{j,i}, t) \right. \\
&\quad - t_i^+(S_j) P_s(S_1, \dots, S_s) P(S_{s+1}, \dots, S_j, t) \\
&\quad + t_i^+(S_j - \Gamma_{ji}) P_s(S_1 - \Gamma_{ii}, \dots, S_s - \Gamma_{si}) P(S_{s+1} - \Gamma_{s+1,i}, \dots, S_j \\
&\quad \left. - \Gamma_{j,i}, t) - t_i^-(S_j) P_s(S_1, \dots, S_s) P(S_{s+1}, \dots, S_j, t) \right), \\
P(S_{s+1}, \dots, S_j, t) \frac{dP_s(S_1, \dots, S_s)}{dt} &= 0 = \Omega \sum_{i=rap} \left(t_i^-(S_j + \Gamma_{ji}) P_s(S_1 + \Gamma_{ii}, \dots, S_s + \Gamma_{si}) P(S_{s+1} + \Gamma_{s+1,i}, \dots, S_j + \Gamma_{j,i}, t) \right. \\
&\quad - t_i^+(S_j) P_s(S_1, \dots, S_s) P(S_{s+1}, \dots, S_j, t) \\
&\quad + t_i^+(S_j - \Gamma_{ji}) P_s(S_1 - \Gamma_{ii}, \dots, S_s - \Gamma_{si}) P(S_{s+1} - \Gamma_{s+1,i}, \dots, S_j \\
&\quad \left. - \Gamma_{j,i}, t) - t_i^-(S_j) P_s(S_1, \dots, S_s) P(S_{s+1}, \dots, S_j, t) \right), \tag{6.56}
\end{aligned}$$

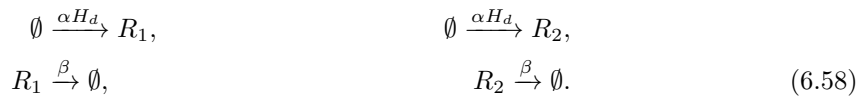
(notar que en la segunda ecuación solo se consideran las reacciones rápidas) en la primer ecuación sacamos el promedio con respecto de la parte estacionaria, mientras que con la segunda ecuación encontramos la forma de la distribución estacionaria, para eso usamos el método que hemos explicado en el capítulo 2, en esta segunda solo entran las reacciones rápidas por lo que se consideran estacionarias las variables asociadas a estas, definimos al promedio con respecto de la parte estacionaria como $\langle t^\pm \rangle_S$ de esta manera nuestra primera ecuación se convierte en

$$\begin{aligned}
\frac{dP(S_{s+1}, \dots, S_j, t)}{dt} &= \Omega \sum_i \left(\langle t_i^-(S_j + \Gamma_{ji}) \rangle_S P(S_{s+1} + \Gamma_{s+1,i}, \dots, S_j + \Gamma_{j,i}, t) - \langle t_i^+(S_j) \rangle_S P(S_{s+1}, \dots, S_j, t) \right. \\
&\quad \left. + \langle t_i^+(S_j - \Gamma_{ji}) \rangle_S P(S_{s+1} - \Gamma_{s+1,i}, \dots, S_j - \Gamma_{j,i}, t) - \langle t_i^-(S_j) \rangle_S P(S_{s+1}, \dots, S_j, t) \right), \tag{6.57}
\end{aligned}$$

es justo esta ecuación la que se usara para modelar la dinámica de este tipo de sistemas solo que ahora las tasas de reacción están promediadas con respecto de la parte estacionaria, para más detalles sobre los desarrollos se puede revisar [34]. Se espera que en los promedios con respecto a la parte estacionaria aparezcan las funciones de Hill, para clarificar el cómo aparecerán analizaremos un ejemplo.

6.2.5. Toggle switch

Ahora analizaremos el “toggle switch” y explicaremos la forma en la que aparece la función de Hill dentro de la ecuación maestra, además, se puede usar la función de Hill semi determinista o alguna otra según las consideraciones que nosotros realicemos. Para simplificar los cálculos analizamos el “toggle switch” simplificado, el “toggle switch” es de hecho un circuito genético, pero omitiremos y/o simplificaremos el modelo al solo fijarnos en las proteínas, donde las proteínas actúan como represores al entrar en contacto el uno con el otro, r_1 y r_2 , si r_1 aumenta, entonces la síntesis de r_2 disminuye, e inversamente, si r_2 aumenta, r_1 disminuye. Una simple red química que describe este sistema son las siguientes 4 reacciones.



De estas reacciones obtenemos las siguientes tasas de propensión

$$\begin{aligned}
t_1^+ &= \alpha H_d (r_2 / \Omega), & t_2^+ &= \alpha H_d (r_1 / \Omega) \\
t_1^- &= \beta \frac{r_1}{\Omega}, & t_2^- &= \beta \frac{r_2}{\Omega},
\end{aligned} \tag{6.59}$$

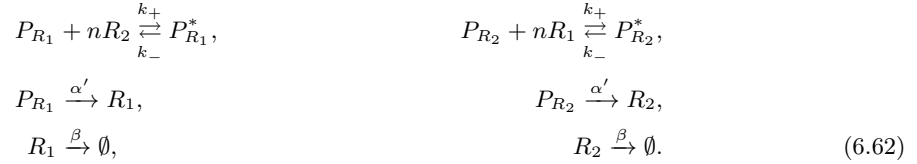
aquí se ha introducido de lleno la ecuación de Hill determinista H_d , definida por

$$H_d(x) = \frac{\left(\frac{x}{K_R}\right)^n}{1 + \left(\frac{x}{K_R}\right)^n}, \quad (6.60)$$

La ecuación maestra de este sistema esta dado por

$$\begin{aligned} \frac{d}{dt} P(r_1, r_2, t) = & \alpha \Omega H_d\left(\frac{r_2}{\Omega}\right) P(r_1 - 1, r_2, t) - \beta r_1 P(r_1, r_2, t) \\ & + \beta(r_1 + 1) P(r_1 + 1, r_2, t) - \alpha \Omega H_d\left(\frac{r_2}{\Omega}\right) P(r_1, r_2, t) \\ & + \alpha \Omega H_d\left(\frac{r_1}{\Omega}\right) P(r_1, r_2 - 1, t) - \beta r_2 P(r_1, r_2, t) \\ & + \beta(r_2 + 1) P(r_1, r_2 + 1, t) - \alpha \Omega H_d\left(\frac{r_1}{\Omega}\right) P(r_1, r_2, t), \end{aligned} \quad (6.61)$$

Tal y como esta este sistema es la forma usual en la que esta llega a plantearse en la literatura, donde se introduce la función de Hill determinista sin una correcta justificación, algo que puede llegar a dar una visión sesgada del problema que se esté analizando y cuando el sistema se vuelve más pequeño el usar la versión determinista puede resultar incluso contraproducente, por lo que para determinar y ver como aparecen las funciones de Hill en el sistema nos fijaremos en las siguientes reacciones químicas, las cuales corresponden al "Toggle Switch" simplificado



Hemos separado a las reacciones químicas en dos partes, a los de la derecha corresponden a cambios en la proteína R_1 y a la izquierda a R_2 , lo que explicaremos a continuación vale tanto para la derecha como para la izquierda, recordar que cuando una proteína llega a un MTT (módulo de transcripción/traducción) esta primero entra en contacto con la región promotora la cual precede al síntesis del ARm, por lo que las proteínas actúan como represoras o activadoras en esta región, pero como estamos analizando el toggle switch estas son represoras, es decir van desactivando a las promotoras, y es justo lo que escribimos en la primera línea (considerando que n proteínas participan para inhibir a una promotora). Dado que estamos optando por un modelo simplificado decimos que la promotora se traduce directamente a la proteína (aunque esto no es del todo correcto), esto lo capturamos en la segunda línea. La última nos hablan sobre como las proteínas se van degradando. Empleamos el formalismo desarrollado en el proceso de vida y muerte multivariable para analizar este sistema, primero calculamos las matrices de coeficientes estequiométricos y la matriz estequiométrica

$$\alpha_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & n \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & n & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \beta_{ij} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Gamma_{ij} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & -n \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -n & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad (6.63)$$

con ayuda de estas, calculamos las tasas de propensión dadas por las siguientes expresiones

$$t_i^+(S_j) = k_i^+ \prod_j \frac{S_j!}{(S_j - \alpha_{ij})!}, \quad t_i^-(S_j) = k_i^- \prod_j \frac{S_j!}{(S_j - \beta_{ij})!}.$$

Por lo que obtenemos

$$\begin{aligned}
 t_1^+ &= k_+ p_1 \frac{r_2!}{(r_2 - n)!} \frac{1}{\Omega^{n+1}}, & t_1^- &= k_- \frac{p_1^*}{\Omega}, \\
 t_2^+ &= \alpha' \frac{p_1}{\Omega}, & t_2^- &= 0, \\
 t_3^+ &= \beta \frac{r_1}{\Omega}, & t_3^- &= 0, \\
 t_4^+ &= k_+ p_2 \frac{r_1!}{(r_1 - n)!} \frac{1}{\Omega^{n+1}}, & t_4^- &= k_- \frac{p_2^*}{\Omega}, \\
 t_5^+ &= \alpha' \frac{p_2}{\Omega}, & t_5^- &= 0, \\
 t_6^+ &= \beta \frac{r_2}{\Omega}, & t_6^- &= 0,
 \end{aligned} \tag{6.64}$$

en estas ecuaciones hemos denotado a r_1 como el número de elementos de la especie química R_1 , p_1 a la polimerasa activa, p_1^* a la polimerasa desactiva, hacemos lo mismo con las otras variables. Para poder calcular nuestra ecuación maestra nos apoyamos de la siguiente relación

$$\partial_t P(S_j, t) = \Omega \sum_i \left(t_i^- (S_j + \Gamma_{ji}) P(S_j + \Gamma_{ji}, t) - t_i^+ (S_j) P(S_j, t) + t_i^+ (S_j - \Gamma_{ji}) P(S_j - \Gamma_{ji}, t) - t_i^- (S_j) P(S_j, t) \right),$$

con todo lo anterior finalmente tenemos dos opciones; la primera es determinar la función de Hill a partir de consideraciones deterministas (semi-deterministas o semi-estocásticas), o encontrar una a partir de la identificación de las reacciones rápidas en la ecuación maestra y apoyarnos de su distribución de probabilidad, empezaremos por tratar el primer caso.

Opción 1

En este primer caso obtendremos la siguiente ecuación maestra

$$\begin{aligned}
 \partial_t P(r_1, r_2, t) &\approx \alpha' (p_2 + 1) P(r_1 - 1, r_2, t) - \alpha' p_1 P(r_1, r_2, t) \\
 &\quad + \alpha' (p_1 + 1) P(r_1, r_2 - 1, t) - \alpha' p_2 P(r_1, r_2, t) \\
 &\quad + \beta (r_1 + 1) P(r_1 + 1, r_2, t) - \beta r_1 P(r_1, r_2, t) \\
 &\quad + \beta (r_2 + 1) P(r_1, r_2 + 1, t) - \beta r_2 P(r_1, r_2, t),
 \end{aligned} \tag{6.65}$$

la primera y la cuarta reacción (6.62) son los procesos relacionados a la región promotora y ocurren muy rápidamente (comparado con la síntesis y degradación de la proteína) entonces se consideran en equilibrio, por lo que se han omitido en la ecuación maestra, dado que estos procesos se encuentren en equilibrio se encuentran las siguientes relaciones

$$\begin{aligned}
 \Omega^n k_- (p_1^* + 1) P(r_1, r_2 - n, p_1 - 1, p_2, p_1^* + 1, p_2^*) - k_+ p_1 \frac{r_2!}{(r_2 - n)!} P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 k_+ (p_1 + 1) \frac{(r_2 + n)!}{r_2!} P(r_1, r_2 + n, p_1 + 1, p_2, p_1^*, p_2^* - 1) - \Omega^n k_- p_1^* P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 \Omega^n k_- (p_2^* + 1) P(r_1 - n, r_2, p_1, p_2 - 1, p_1^*, p_2^* + 1) - k_+ p_2 \frac{r_1!}{(r_1 - n)!} P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 k_+ (p_2 + 1) \frac{(r_1 + n)!}{r_1!} P(r_1 + n, r_2, p_1, p_2 + 1, p_1^*, p_2^* - 1) - \Omega^n k_- p_2^* P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0,
 \end{aligned} \tag{6.66}$$

para ello se podría calcular la función de Hill tal y como lo hemos hecho en la sección de **variantes de la función de Hill**, lo que en realidad se hace aquí es sustituir a p_1 y p_2 en t_3^+ y t_4^+ por su correspondiente función de Hill, debido a que p_1 y p_2 son proporcionales a la función de Hill entonces usamos estas funciones en vez de las variables, concretamente se tendría algo de la forma

$p_1 = \frac{\alpha\Omega}{\alpha'} D(r_2, \Omega)$ (muy similar con p_2), de esta manera finalmente obtenemos

$$\begin{aligned} \partial_t P(r_1, r_2, t) = & \alpha\Omega D(r_1, \Omega) P(r_1, r_2 - 1, t) - \alpha\Omega D(r_2, \Omega) P(r_1, r_2, t) \\ & + \alpha\Omega D(r_2, \Omega) P(r_1 - 1, r_2, t) - \alpha\Omega D(r_1, \Omega) P(r_1, r_2, t) \\ & + \beta(r_1 + 1) P(r_1 + 1, r_2, t) - \beta r_1 P(r_1, r_2, t) \\ & + \beta(r_2 + 1) P(r_1, r_2 + 1, t) - \beta r_2 P(r_1, r_2, t), \end{aligned}$$

este resultado coincide con el que presentamos al principio de la sección, D es una función de Hill para un represor que depende tanto del tamaño del sistema como de alguna variable (r_1, r_2) , para ello nosotros usaremos una función de Hill que actúa como represor en su versión semi-determinista (de igual manera se puede usar la semi-estocástica o determinista)

$$D(e, \Omega) = \frac{\Omega^n K^n (e - n)!}{\Omega^n K^n (e - n)! + e!} - \left(\frac{\Omega^n K^n e!}{\Omega^n K^n e! + (e + n)!} - \frac{\Omega^n K^n (e - n)!}{\Omega^n K^n (e - n)! + e!} \right) \theta \left(\frac{1}{2} - \frac{s}{N_0} \right).$$

Esta función de Hill coincide con la determinista cuando e y/o Ω se vuelven lo suficientemente grandes. A continuación, haremos una simulación del sistema que hemos presentado en esta sección, para ello usaremos la función de Hill semi-determinista, estos resultados se ven en la figura 6.8.

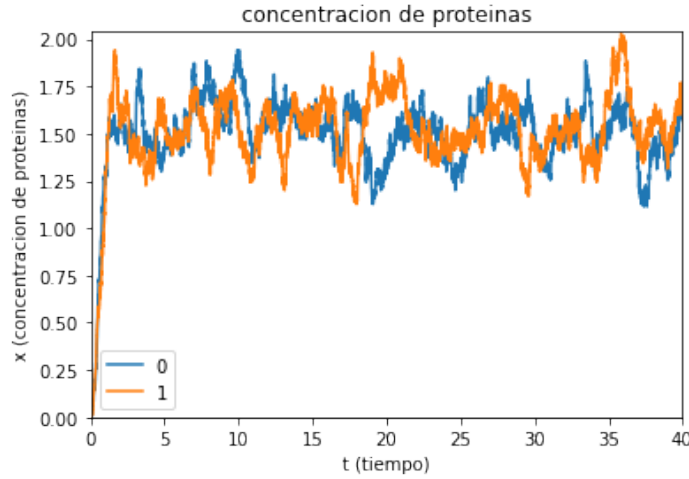


Figura 6.8: Toggle switch. Se realizó la simulación estocástica del toggle switch usando la función de Hill semi-determinista, con $K = 1$, $r_0(0) = r_1(0) = 0$ y $\Omega = 80$.

Opción 2

Lo que se hace aquí es aplicar un promedio sobre la parte dinámica, este promedio se realiza con respecto de la parte estacionaria (notar que solo hemos señalado las variables al que el promedio va a modificar), obteniendo lo siguiente

$$\begin{aligned} \partial_t P(r_1, r_2, t) = & \alpha' \langle p_2 \rangle_s P(r_1 - 1, r_2, t) - \alpha' \langle p_1 \rangle_s P(r_1, r_2, t) \\ & + \alpha' \langle p_1 \rangle_s P(r_1, r_2 - 1, t) - \alpha' \langle p_2 \rangle_s P(r_1, r_2, t) \\ & + \beta(r_1 + 1) P(r_1 + 1, r_2, t) - \beta r_1 P(r_1, r_2, t) \\ & + \beta(r_2 + 1) P(r_1, r_2 + 1, t) - \beta r_2 P(r_1, r_2, t), \end{aligned} \quad (6.67)$$

la primera y la cuarta reacción de (6.62) son las relacionadas a la región promotora y ocurren muy rápido (comparado con la síntesis y degradación de la proteína) entonces se consideran en

equilibrio, por lo que se han omitido en la ecuación maestra, a este tipo de reacciones se les nombra reacciones rápidas, mientras que las reacciones que ocurren más lentamente se les conoce como reacciones lentas, y son estas las que aparecen en la ecuación anterior. De la primera y cuarta reacción de (6.62) se encuentran las siguientes relaciones

$$\begin{aligned}
 \Omega^n k_- (p_1^* + 1) P(r_1, r_2 - n, p_1 - 1, p_2, p_1^* + 1, p_2^*) - k_+ p_1 \frac{r_2!}{(r_2 - n)!} P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 k_+ (p_1 + 1) \frac{(r_2 + n)!}{r_2!} P(r_1, r_2 + n, p_1 + 1, p_2, p_1^*, p_2^* - 1) - \Omega^n k_- p_1^* P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 \Omega^n k_- (p_2^* + 1) P(r_1 - n, r_2, p_1, p_2 - 1, p_1^*, p_2^* + 1) - k_+ p_2 \frac{r_1!}{(r_1 - n)!} P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \\
 k_+ (p_2 + 1) \frac{(r_1 + n)!}{r_1!} P(r_1 + n, r_2, p_1, p_2 + 1, p_1^*, p_2^* - 1) - \Omega^n k_- p_2^* P(r_1, r_2, p_1, p_2, p_1^*, p_2^*) &= 0, \quad (6.68)
 \end{aligned}$$

de estas relaciones se obtiene la distribución de probabilidad estacionaria, para esto usamos el método anteriormente presentado, las distribuciones de probabilidad son las mismas que presentamos anteriormente por lo que omitimos los calcularla, con ayuda de esta distribución estacionaria encontramos los promedios, los cuales resultan ser proporcionales a la función de Hill, por lo que sustituimos directamente este resultado obteniendo lo siguiente

$$\begin{aligned}
 \partial_t P(r_1, r_2, t) &= \alpha \Omega D_1(r_1, \Omega) P(r_1, r_2 - 1, t) - \alpha \Omega D_1(r_2, \Omega) P(r_1, r_2, t) \\
 &+ \alpha \Omega D_1(r_2, \Omega) P(r_1 - 1, r_2, t) - \alpha \Omega D_1(r_1, \Omega) P(r_1, r_2, t) \\
 &+ \beta (r_1 + 1) P(r_1 + 1, r_2, t) - \beta r_1 P(r_1, r_2, t) \\
 &+ \beta (r_2 + 1) P(r_1, r_2 + 1, t) - \beta r_2 P(r_1, r_2, t),
 \end{aligned}$$

con $\alpha' \langle p_2 \rangle_s = \alpha \Omega D_1(r_1, \Omega)$ y $\alpha' \langle p_1 \rangle_s = \alpha \Omega D_1(r_2, \Omega)$, muy similar al caso 1, con la función de Hill definida como

$$D_1(x, \Omega) = \frac{K^n}{K^n + \frac{\langle \frac{x!}{(x-n)!} \rangle}{\Omega^n}}. \quad (6.69)$$

Esta función de Hill es exacta, coincide con la determinista cuando Ω se vuelven lo suficientemente grande, usando una aproximación similar a la presentada en el capítulo 3, particularmente de la expresión

$$\langle f(X_l) \rangle \approx f(\langle X_l \rangle) + \sum_i \sum_j \frac{\sigma^2(X_i, X_j)}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle}, \quad (6.70)$$

por lo que al tomar solo el termino de primer orden del promedio obtenemos

$$D_1(x, \Omega) = \frac{K^n}{K^n + \frac{\langle \frac{x!}{(x-n)!} \rangle}{\Omega^n}} \approx \frac{K^n}{K^n + \frac{\langle x \rangle^n}{\Omega^n}}. \quad (6.71)$$

Con este ejemplo queda más claro el cómo se introduce la función de Hill dentro de los MTT, tal como lo hemos deducido, no siempre es adecuado usar directamente la versión determinista ya que puede llegar a sesgar el cómo se tratan aquellas reacciones en equilibrio, por lo que recomendamos usar directamente las ideas detrás de esta última aproximación.

Hay que recordar que cuando se toma el límite determinista del modelo estocástico se toma el promedio de todas las tasas de propensión, por lo que al tomar el promedio de la función de Hill (6.71) (ocupa el lugar de una tasa de propensión) esta coincidiría consigo mismo, es decir

$$\langle D_1(x, \Omega) \rangle = D_1(x, \Omega), \quad (6.72)$$

comparando este resultado con la expansión en serie (6.70) podemos observar que se cumple

$$\frac{\partial^k}{\partial \langle x \rangle^k} D_1(x, \Omega) = 0, \quad (6.73)$$

($k > 0$) esta condición es de suma importancia Ya que uno de los objetivos de este trabajo es cuantificar las fluctuaciones de las redes de regulación genética, para lograrlo se necesita del teorema de fluctuación-disipación, donde aparecen derivadas. Estas condiciones son sobre una función de Hill supresora, un desarrollo y condiciones similares serán válidas para un activador. En [34, 35] se ha realizado una derivación similar de la función de Hill, sin embargo, ellos no consideran correcciones estocásticas y esta última propiedad que hemos derivado la cual será fundamental cuando se cuantifiquen las fluctuaciones del sistema.

La deducción que realizamos en esta opción 2 es la más exacta, y es como se debería tratar aquellos sistemas estocásticos que tengan reacciones que estén en el estado estacionario, debido a que todos los desarrollos se realizaron sobre la distribución de probabilidad. Cuando se modela un proceso estocástico producido de esta ecuación maestra resulta ser complicada de tratar debido a que es necesario conocer el promedio, sin embargo, cuando se toma el límite determinista de está ecuación resulta ser bastante útil y nos sale directamente la función de Hill determinista, incluso podrían introducirse dentro de estas algunas correcciones estocásticas. Cuando estemos analizando las redes genéticas se hará uso de todas las derivaciones hechas hasta este momento, solo que también para funciones de Hill que actúen como activadoras, no escribimos un ejemplo para estas ya que la derivación sería similar al caso del represor.

6.3. Función de Hill determinista con correcciones

En la parte final de la sección anterior derivamos la función de Hill con correcciones estocásticas a partir de un proceso estocástico, sin embargo, esto resulta ser una tarea más larga. Por lo que ahora intentaremos derivarla desde la aproximación determinista, y le daremos algunas constricciones. Si recordamos la manera en la que se derivó la función de Hill estocástica para un activador, uno puede apoyarse de las tasas de reacción (6.17) y tomar su límite determinista (como hemos explicado anteriormente) por lo que obtendremos

$$\begin{aligned} R_1^{D+} &= \lambda_+, & R_1^{D-} &= \lambda_- \hat{e}, \\ R_2^{D+} &= k^+ \hat{p} \hat{e}^n, & R_2^{D-} &= k^- \hat{s}. \end{aligned} \quad (6.74)$$

En el límite estacionario los términos que aparecen en el lado derecho de la ecuación determinista de las reacciones serían cero, por lo que encontraríamos una relación similar a la siguiente

$$1 = \frac{R_i^+(s_j)}{R_i^-(s_j)}, \quad (6.75)$$

si consideramos que el sistema no es lo suficientemente grande podemos apoyarnos de algunos términos para realizar correcciones, la expresión que usaremos está dada por (6.70),

$$1 = \frac{k_i^+ \left(R_i^{D+}(s_j) + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(S_{l_1}, S_{l_2})}{2\Omega^2} \left[\frac{\partial^2}{\partial s_{l_1} \partial s_{l_2}} (R_i^{D+}(s_j)) \right] \right)}{k_i^- \left(R_i^{D-}(s_j) + \sum_{l_1} \sum_{l_2} \frac{\sigma^2(s_{l_1}, s_{l_2})}{2\Omega^2} \left[\frac{\partial^2}{\partial s_{l_1} \partial s_{l_2}} (R_i^{D-}(s_j)) \right]_{S_j = \Omega s_j} \right)},$$

por lo que al sustituir dentro de esta expresión los términos dados por (6.74) obtenemos

$$1 = \frac{k^+ \left(\hat{p} \hat{e}^n + \frac{\sigma^2(p,e)}{\Omega^2} n \hat{e}^{n-1} + \frac{\sigma^2(e,e)}{\Omega^2} n(n-1) \hat{p} \hat{e}^{n-2} \right)}{k^- \hat{s}}, \quad (6.76)$$

despejando

$$\hat{s} = \frac{k^+}{k^-} \left(\hat{p}\hat{e}^n + \frac{\sigma^2(p, e)}{\Omega^2} n\hat{e}^{n-1} + \frac{\sigma^2(e, e)}{\Omega^2} n(n-1)\hat{p}\hat{e}^{n-2} \right), \quad (6.77)$$

por ahora solo faltaría encontrar el valor específico de las covarianzas, y así determinar de manera adecuada la nueva función de Hill. Hay que recordar que con la ayuda de la distribución de probabilidad del sistema (no se ha considerado aun si solo el proceso de la función de Hill es la que se está analizando) podemos determinar que las covarianzas cumplirán las siguientes condiciones

$$\frac{\sigma^2(p, e)}{\Omega^2} = 0, \quad \frac{\sigma^2(e, e)}{\Omega^2} \neq 0, \quad (6.78)$$

por lo que ahora

$$\hat{s} = \frac{k^+}{k^-} \left(\hat{p}\hat{e}^n + \frac{\sigma^2(e, e)}{\Omega^2} n(n-1)\hat{p}\hat{e}^{n-2} \right), \quad (6.79)$$

recordando la definición de la función de Hill, $H = \frac{s}{s+p}$, obtenemos nuestra función de Hill con correcciones estocásticas ($K^n = \frac{k^-}{k^+}$)

$$H_{ce} = \frac{\hat{e}^n + \frac{\sigma^2(e, e)}{\Omega^2} n(n-1)\hat{e}^{n-2}}{K^n + \hat{e}^n + \frac{\sigma^2(e, e)}{\Omega^2} n(n-1)\hat{e}^{n-2}} \quad (6.80)$$

Esta función de Hill es válida para cualquier sistema en el cual aparezcan las reacciones en el cual se unen enzimas a proteínas y su dinámica llegue rápidamente al estacionario, la varianza se puede inferir directamente del sistema, para ello uno puede emplear el teorema de fluctuación disipación que derivamos en el capítulo anterior. Siguiendo una derivación similar encontraríamos una función de Hill para un represor

$$D_{ce} = \frac{K^n}{K^n + \hat{e}^n + \frac{\sigma^2(e, e)}{\Omega^2} n(n-1)\hat{e}^{n-2}} \quad (6.81)$$

La función de Hill que encontramos al final de la sección anterior coincide con esta última relación, por lo que se puede usar este si se consideran correcciones estocásticas, para ello se tiene que considerar que las derivadas de esta función son cero

$$\partial_e^k D_{ap}(e, \Omega) = 0, \quad (6.82)$$

esta última propiedad se considera en concordancia a (6.72). En el caso en el que se tenga un sistema que está compuesto solo de las siguientes reacciones podremos inferir el valor de la varianza,



ya hemos derivado la forma explícita de la distribución de probabilidad del estado estacionario, la cual es

$$P(e, p, s) = \frac{1}{\left(1 + \frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^{N_0}} \left(\frac{k_+}{k_-} \left(\frac{\lambda_+}{\lambda_-}\right)^n\right)^s \frac{N_0!}{s!(N_0 - s)!} \frac{1}{e!} \left(\Omega \frac{\lambda_+}{\lambda_-}\right)^e e^{-\Omega \frac{\lambda_+}{\lambda_-}} \delta(p, N_0 - s).$$

con ayuda de esta expresión determinamos la covarianza, obteniendo $\frac{\sigma^2(e, e)}{\Omega^2} = \frac{\hat{e}}{\Omega}$, finalmente tendremos la siguiente función de Hill

$$H_{ap} = \frac{\hat{e}^n + \frac{1}{\Omega} n(n-1)\hat{e}^{n-1}}{K^n + \hat{e}^n + \frac{1}{\Omega} n(n-1)\hat{e}^{n-1}}. \quad (6.83)$$

Siguiendo un desarrollo similar para el caso de un represor, obtendremos lo siguiente

$$D_{ap} = \frac{K^n}{K^n + \hat{e}^n + \frac{1}{\Omega}n(n-1)\hat{e}^{n-1}}. \quad (6.84)$$

Esta es una nueva función de Hill contiene términos correctivos que dependen del inverso de Ω y además solo cuando el coeficiente de Hill es diferente de uno se puede observar su contribución, por lo que en sistemas que son lo suficientemente grandes los términos correctivos desaparecen y recuperamos la función de Hill usual muy bien conocida.

$$\lim_{\Omega \rightarrow \infty} \frac{\hat{e}^n + \frac{1}{\Omega}n(n-1)\hat{e}^{n-1}}{K^n + \hat{e}^n + \frac{1}{\Omega}n(n-1)\hat{e}^{n-1}} = \frac{\hat{e}^n}{K^n + \hat{e}^n} = H_d \quad (6.85)$$

Podemos realizar incluso una aproximación de tercer o cuarto orden en las tasas de reacción, y después introducirlas en la función de Hill y de esta manera obtener una nueva función, sin embargo, esto último no lo haremos debido a que dependerán del inverso de Ω^2 o términos de orden superior y estos ya serían muy pequeños, entonces su contribución sería prácticamente despreciable, estos términos correctivos se deben a las fluctuaciones de la unión de la enzima y la proteína.

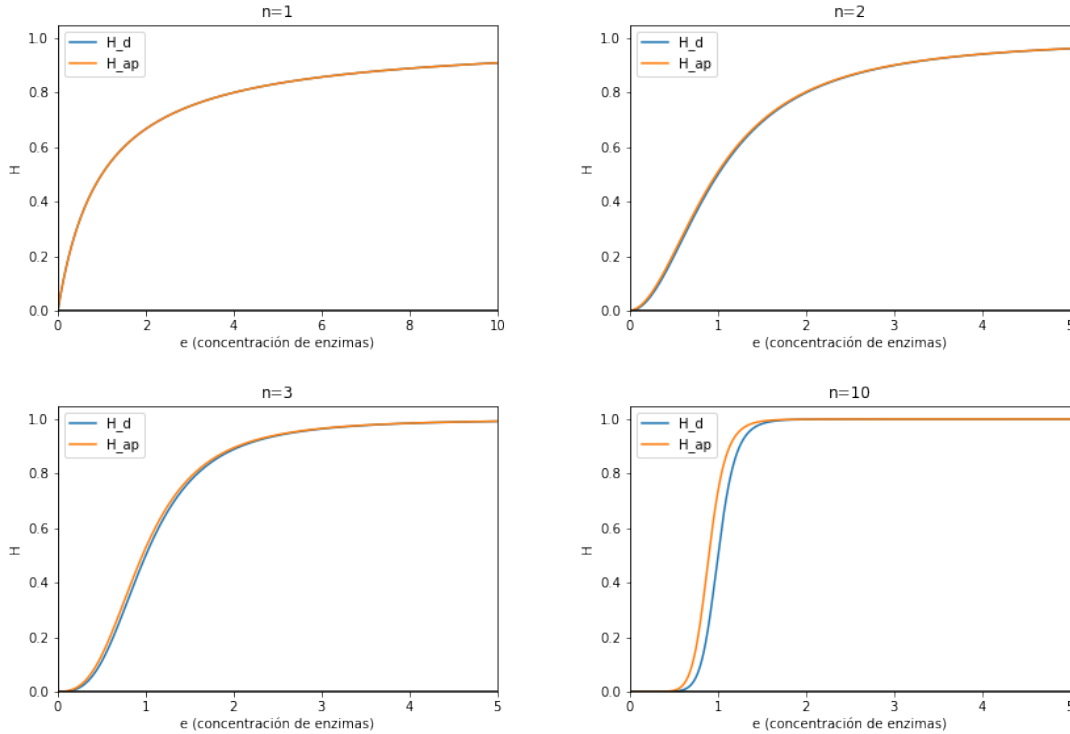


Figura 6.9: Gráfica de la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$, $\Omega = 50$.

En la figura 6.9 veremos cómo es el comportamiento de esta nueva función de Hill (para un activador) con correcciones estocásticas y la comparamos con la función de Hill determinista. Lo primero que vemos es que en el caso en el cual $n = 1$ la función de Hill con correcciones coincide con la determinista, mientras que con coeficientes de Hill mayores que 1 se aprecian las diferencias, lo cual podría ser significativo para sistemas pequeños, en todas las figuras se observa que hay un rango de valores en el cual si existe una diferencia entre estas dos funciones mientras n más se

aleja de 1.

En alguno de los modelos que analizaremos en los próximos capítulos emplearemos esta función de Hill para construir y modelar algunas redes de regulación genética.

6.4. Coeficiente de Hill decimal

Existen algunos aspectos de la función de Hill que a veces no pueden llegar a ser claras, uno de estos es el coeficiente de Hill (n), por lo que a continuación analizaremos esta cuestión, lo desarrollado en esta sección puede saltarse, pero aun así recomendamos leerla para entender mejor la función de Hill.

Originalmente Hill derivó su ecuación para coeficientes de Hill n enteros, sin embargo, los experimentos arrojaron que este coeficiente toma valores con términos decimales, como en el caso de la hemoglobina, que tiene por sí sola un coeficiente de Hill $n = 4$, sin embargo en los resultados experimentales se encontró que $n = 2,8 - 3$, con lo cual la perspectiva que originalmente dio Hill (y las mismas ideas que hemos optado hasta este punto) podrían ser insuficientes, por lo que la teoría de Hill podría resultar ser más de carácter descriptivo que explicativo [36], por lo que ahora intentaremos dar una generalización de la función de Hill para seguir con el carácter explicativo de esta teoría, se consideraran dos maneras más detalladas de como unen las enzimas a las proteínas, por ejemplo colocando procesos intermedios, además, considerando la unión secuencial o independiente [37], así como definir de manera adecuada las tasas de reacción, en la figura 6.10 se puede apreciar alguno de estos dos tipos de procesos.

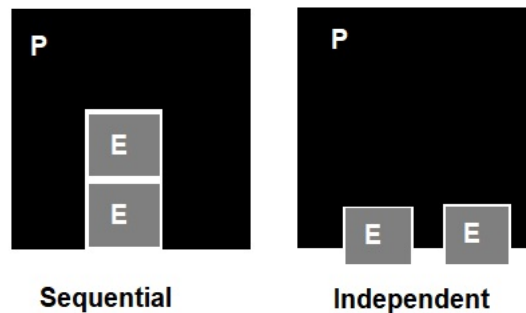


Figura 6.10: Esquemización de procesos de unión de proteína-enzima tipo secuencial e independiente.

Aunque en la figura 6.10 solo muestra una representación de cómo sería la unión de enzimas y proteínas resulta ser bastante ilustrativa al momento de pensar el cómo es la estructura de los lugares en los cuales se unen las enzimas. Algunos ejemplos donde estos se pueden aplicar son los siguientes:

1. Secuencial: Transporte de iones.
2. Independiente: Proteínas poliméricas.

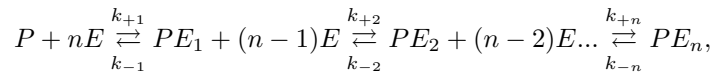
A pesar de que solo trabajaremos con estos dos modelos, es posible usar otras generalizaciones [37], como es el caso de que las proteínas estén activas o desactivas, incluso se puede insertar la temperatura a través de la ecuación de Artenius.

6.4.1. Caso determinista

Ahora derivaremos las funciones de Hill con esta generalización apoyándonos de la ley de acción de masas y de las reacciones que aparecen en [37], se supone que existen otros mecanismos para la unión de la enzima a la proteína, para ello uno supone que existen procesos intermedios, estos procesos intermedios se pueden dividir en secuenciales e independientes como en la figura 6.10, primero empezaremos por analizar el caso secuencial.

Caso secuencial

Para el caso secuencial tenemos las siguientes reacciones



en el estado estacionario las reacciones químicas cumplen la siguiente relación

$$1 = \frac{R_i^+(x)}{R_i^-(x)}, \quad (6.86)$$

para las reacciones que tenemos anteriormente tendríamos las tasas de reacción de la siguiente manera

$$\begin{aligned} R_i^+(x) &= k_i^+ e^{n+1-i} s_{i-1}, \\ R_i^-(x) &= k_i^- e^{n-i} s_i, \end{aligned} \quad (6.87)$$

($i=1, \dots, n$) donde hemos definido a s_0 como la concentración de P , e a la concentración de E y s_n como la concentración de PE_n , con esto la ecuación (6.86), quedara como

$$1 = e \frac{s_{i-1}}{s_i K_i}, \quad (6.88)$$

($K_i = \frac{k_{-i}}{k_{+i}}$) despejando

$$s_i = e s_{i-1} \frac{1}{K_i}, \quad (6.89)$$

iterando llegamos a la siguiente relación

$$s_i = s_0 e^i \left(\prod_{l=1}^i K_l \right)^{-1}, \quad (6.90)$$

nuestra función de Hill está definida como

$$H_d = \frac{s_n}{s_0 + \sum_{i=1}^n s_i}, \quad (6.91)$$

sustituyendo el valor de s_i en esta última expresión

$$H_d = \frac{e^n \left(\prod_{l=1}^n K_l \right)^{-1}}{1 + \sum_{i=1}^n e^i \left(\prod_{l=1}^i K_l \right)^{-1}}, \quad (6.92)$$

simplificando, finalmente obtenemos

$$H_d = \frac{e^n}{\prod_{l=1}^n K_l + \sum_{i=1}^n e^i \left(\prod_{l=i+1}^n K_l \right)}, \quad (6.93)$$

para el caso en el cual $K_i = K$, esta expresión se reducirá a

$$H_d = \frac{e^n}{K^n + \sum_{i=1}^n e^i K^{n-i}} = \frac{e^n(K - e)}{K^{n+1} - e^{n+1}}. \quad (6.94)$$

para ver cómo es la forma de esta expresión tomaremos como ejemplo el caso en el cual $n = 4$, por lo que nuestra función de Hill tendrá la siguiente forma

$$H_d = \frac{e^4}{K^4 + K^3e + K^2e^2 + Ke^3 + e^4}, \quad (6.95)$$

A continuación, mostramos a través de unas gráficas cómo se comporta la expresión (6.94) y se puede ver en la figura 6.11, aquí podemos observar que el caso en el cual $n = 1$ coincide con la función de Hill usual, mientras que para diferentes valores de la n mientras la concentración de enzimas sea pequeña la sigmoide está más acostada.

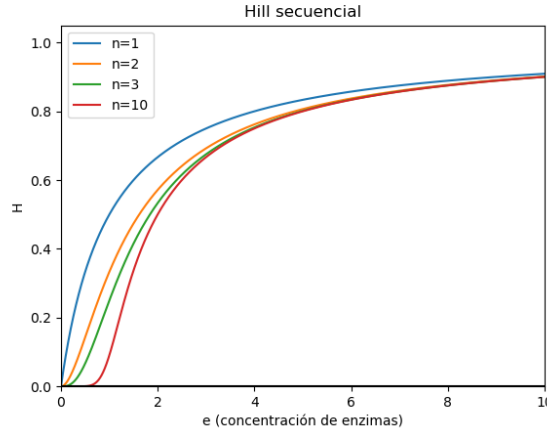
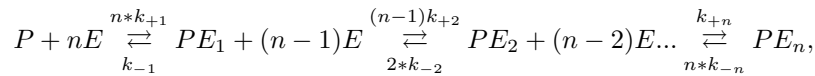


Figura 6.11: Función de Hill con unión de proteína-enzima secuencial, se evalúa la función de Hill para distintos valores de n y con $K = 1$.

Caso independiente

Para el caso en el cual la unión de enzimas y proteínas se realiza de manera independiente tendremos las siguientes reacciones



realizaremos un análisis similar al del caso secuencial, por lo que estas reacciones en el límite estacionario también cumplen (6.86), pero ahora sus tasas de reacción serán las siguientes

$$\begin{aligned} R_i^+(x) &= k_i^+(n-i)e^{n+1-i}s_{i-1}, \\ R_i^-(x) &= k_i^-(i)e^{n-i}s_i, \end{aligned} \quad (6.96)$$

($i = 1, \dots, n$) donde hemos definido a s_0 como la concentración de P , e a la concentración de E y s_n como la concentración de PE_n , con esto la ecuación (6.86) quedara como

$$1 = e \frac{s_{i-1}}{s_i K_i} \frac{(n-i)}{i}, \quad (6.97)$$

($K_i = \frac{k_{-i}}{k_{+i}}$) despejando

$$s_i = e s_{i-1} \frac{1}{K_i} \frac{(n-i)}{i}, \quad (6.98)$$

iterando obtenemos

$$s_i = s_0 e^i \left(\prod_{l=1}^i K_l \right)^{-1} \frac{n!}{i!(n-i)!}, \quad (6.99)$$

nuestra función de Hill está definida como

$$H_d = \frac{s_n}{s_0 + \sum_{i=1}^n s_i}, \quad (6.100)$$

sustituyendo el valor de s_i en esta última expresión

$$H_d = \frac{e^n \left(\prod_{l=1}^n K_l \right)^{-1}}{1 + \sum_{i=1}^n e^i \frac{n!}{i!(n-i)!} \left(\prod_{l=1}^i K_l \right)^{-1}}, \quad (6.101)$$

simplificando,

$$H_d = \frac{e^n}{\prod_{l=1}^n K_l + \sum_{i=1}^n e^i \frac{n!}{i!(n-i)!} \left(\prod_{l=i+1}^n K_l \right)}. \quad (6.102)$$

para el caso en el cual $K_i = K$, esta expresión se reducirá a

$$H_d = \frac{e^n}{K^n + \sum_{i=1}^n e^i \frac{n!}{i!(n-i)!} K^{n-i}} = \frac{e^n}{(K + e)^n}. \quad (6.103)$$

Para ver una representación de esto vemos el caso en el cual $n = 4$, por lo que nuestra función de Hill tendrá la siguiente forma

$$H_d = \frac{e^4}{K^4 + 4K^3e + 6K^2e^2 + 4Ke^3 + e^4}. \quad (6.104)$$

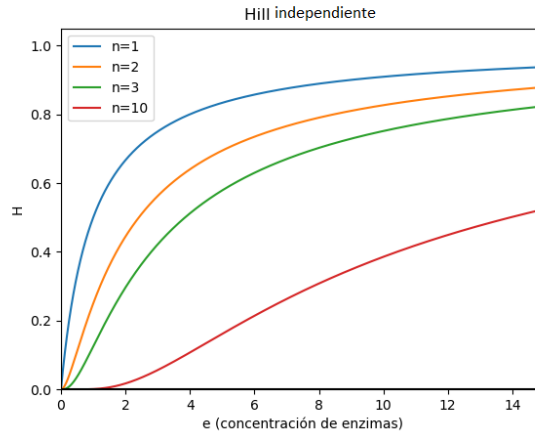


Figura 6.12: Función de Hill con unión de proteína-enzima independiente, se evalúa la función de Hill para distintos valores de n y con $K = 1$.

En la figura 6.12 se ve cómo se comporta la expresión (6.103). En las figuras que hemos realizado para los dos casos podemos observar que cuando $n = 1$ coincide el valor de las funciones de Hill usual, secuencial e independiente, mientras que cuando estamos con un n mayor se ve que la función tarda más en acercarse a uno cuando es el caso secuencial, mientras que cuando es independiente esta característica se remarca aún más.

Nos damos cuenta de que en realidad si hay una gran diferencia de la forma de estas funciones, además, en la mayoría de los procesos biológicos el valor de la n no sobrepasa 4 por lo que casi no se emplean este tipo de funciones, también el valor de la K^n resulta ser muy pequeño, en algunos casos es del orden de 10^{-4} . (6.94) y (6.103) se reducirían a la función de Hill usual (6.4) si eliminamos todos los procesos intermedios. A un sistema que tenga $n > 1$ se le dice que tiene cooperatividad positiva, cuando $n < 1$ cooperatividad negativa, y finalmente cuando $n = 1$ no hay cooperatividad. Algunos resultados importantes y que vale la pena mencionar de estas expresiones son las siguientes [37]:

- 1.- El coeficiente de Hill no es un indicador confiable del número de sitios de unión de ligandos funcionalmente importantes en el receptor, pero proporciona solo un valor mínimo.
- 2.- La única condición bajo la cual el coeficiente de Hill estima con precisión el número de sitios de unión para un esquema de unión secuencial o independiente simple es cuando está presente una cooperatividad positiva marcada.
- 3.- Los cambios en la cooperatividad generalmente también afectan el coeficiente K_i así como el coeficiente de Hill.
- 4.- Es posible que un sistema que considere procesos intermedios reproduzca diferentes grados de cooperatividad (n con valores decimales).

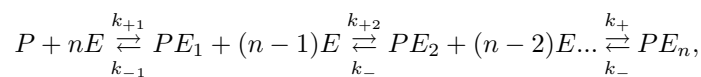
La función de Hill presentada en las secciones anteriores en realidad no es una descripción realista de la unión enzimática, sin embargo, puede llegar a describir como es este proceso, el coeficiente de Hill puede ser mejor entendido como un coeficiente que describe el grado de interacción entre las enzimas y los receptores, como se puede observar en el artículo [37], muchos sistemas que tengan diferentes parámetros de Hill pueden llegar a comportarse de manera similar dependiendo del grado de cooperatividad que existe en el sistema, para más detalles revisar el artículo antes mencionado.

6.4.2. Caso estocástico

Tal como hemos mencionado, es posible que la forma en la que las enzimas se van uniendo a la proteína no sea a través de un solo proceso, e incluso el número de elementos en nuestro sistema es pequeño, entonces es posible que existan fluctuaciones, por lo que nos preguntamos cual sería la versión estocástica de las ecuaciones (6.94) y (6.103), por lo que para ello aplicaremos el formalismo del proceso de vida y muerte multivariable, primero nos fijaremos en una función de Hill semi-determinista, después introduciremos correcciones estocásticas.

Función de Hill semi-determinista

Solo nos fijaremos en el caso en el cual la unión de las enzimas a la proteína se hace de manera secuencial. Dado que la derivación de la correspondiente función de Hill estocástica y semi-estocástica representa una tarea sumamente compleja nos conformaremos con la función de Hill semi-determinista, para ello nos fijamos en el siguiente proceso



omitiremos escribir la ecuación maestra, para ello solo nos centraremos en los coeficientes de transición, desarrollando se obtiene la i -énesima tasa de transición de este proceso que contiene n reacciones están dados por

$$t_i^+ = \frac{1}{\Omega^{n+2-i}} k_{+i} s_{i-1} \frac{e!}{(e-1+i-n)!}, \quad t_i^- = \frac{1}{\Omega^{n+1-i}} k_{-i} s_i \frac{e!}{(e-n+i)!}, \quad (6.105)$$

($i = 1, 2, \dots, n$) con e , s_0 y s_i el número de moléculas E , P y PE_i respectivamente, se tendrían que realizar un muchas de asunciones para este proceso si se quisiese calcular la función de Hill estocástica y semi-estocástica, pero debido a que solo calcularemos la semi-determinista, se puede llegar a la siguiente relación, el cual se deriva del hecho que en el límite estacionario estas dos últimas relaciones son iguales,

$$\frac{1}{\Omega} k_{+i} s_{i-1} \frac{E!}{(E-1+i-n)!} = k_{-i} s_i \frac{E!}{(E+n-i)!},$$

de aquí despejamos y encontramos ($K_i = \frac{k_{-i}}{k_{+i}}$)

$$s_i = s_{i-1} \frac{1}{\Omega K_i} (e+n-i), \quad (6.106)$$

iterando obtenemos (con $s_0 = p$)

$$s_i = s_0 \frac{(E+n-1)!}{(E+n-i-1)!} \prod_{l=1}^i \frac{1}{\Omega K_l}, \quad (6.107)$$

para este caso la función de Hill será

$$H_{se} = \frac{s_n}{s_0 + \sum_{i=1}^n s_i}. \quad (6.108)$$

Con ayuda de la relación para s_i sustituimos en esta última expresión, obteniendo finalmente nuestra función de Hill semi-determinista

$$H_{se} = \frac{\frac{(E+n-1)!}{(E-1)!} \prod_{l=1}^n \frac{1}{\Omega K_l}}{1 + \sum_{i=1}^n \frac{(E+n-1)!}{(E+n-i-1)!} \prod_{l=1}^i \frac{1}{\Omega K_l}}, \quad (6.109)$$

en el caso en el que todas las constantes sean la misma $K_i = K$, la relación anterior se reduce a

$$H_{se} = \frac{\frac{(E+n-1)!}{(E-1)!}}{\sum_{i=0}^n \frac{(E+n-1)!}{(E+n-i-1)!} (\Omega K)^{n-i}}, \quad (6.110)$$

calculamos el sumatorio del denominador, finalmente obtenemos nuestra función de Hill

$$H_{se} = \frac{\exp(-K\Omega) (K\Omega)^{E-1} (n+E-1)!}{(E-1)((E-1)!\Gamma(n+E, K\Omega) - (n+E-1)!\Gamma(E-1, K\Omega))}. \quad (6.111)$$

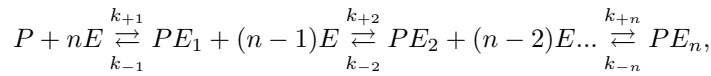
Esta función de Hill se ha calculado por primera vez, aunque es de hecho muy complicada de usar, podemos remarcar que incluso puede llegar a funcionar tan solo en algunas secciones debido a que la asunción sobre la distribución de probabilidades, debido a la gran cantidad de elementos presentes en el sistema. El caso que hemos calculado pertenece al secuencial, podemos aplicarlo cuando la unión de enzimas a la proteína se hace de manera independiente, seguiríamos un desarrollo similar a lo que aquí presentamos, aunque esto último lo dejaremos para otra ocasión.

Función de Hill con procesos intermedios y correcciones estocásticas

Ahora lo que sigue será hacer una derivación de estas funciones de Hill con procesos intermedios pero ahora considerando correcciones estocásticas, haremos una derivación similar hecha en la sección de la función de Hill con correcciones estocásticas, recordar que en el estado estacionario se cumplen las siguientes condiciones

$$\frac{\sigma^2(p, e)}{\Omega^2} = 0, \quad \frac{\sigma^2(e, e)}{\Omega^2} \neq 0, \quad 1 = \frac{R_i^+(s_j)}{R_i^-(s_j)}. \quad (6.112)$$

Consideraremos primero el caso secuencial, es decir tendremos un conjunto de reacciones de la siguiente forma



para las reacciones que tenemos anteriormente tendríamos unas tasas de reacción de la siguiente manera

$$\begin{aligned} R_i^+(x) &= k_i^+(e^{n+1-i} + \sigma^2(e, e)(n+1-i)(n-i)e^{n-i-1})s_{i-1}, \\ R_i^-(x) &= k_i^-(e^{n-i} + \sigma^2(e, e)(n-i)(n-i-1)e^{n-i-2})s_i, \end{aligned} \quad (6.113)$$

($i=1, \dots, n$) donde hemos definido a s_0 como la concentración de P , e la concentración de E y s_n como la concentración de PE_n , estas tasas de reacción incluyen correcciones estocásticas. Considerando a la tercera condición de (6.112) y un poco de álgebra obtenemos

$$s_i = s_{i-1} \frac{1}{K_i} \left(\frac{e + \sigma^2(e, e)(n+1-i)(n-i)e^{-1}}{1 + \sigma^2(e, e)(n-i)(n-i-1)e^{-2}} \right), \quad (6.114)$$

($K_i = \frac{k_i^-}{k_i^+}$) podemos ver que la primera diferencia con respecto de la función de Hill determinista con procesos intermedios es la aparición de un factor que modifica a e , resolviendo esta ecuación encontramos que se cumple

$$s_j = s_0 \left(\prod_{i=1}^j K_i \right)^{-1} \prod_{i=0}^j \left(\frac{e + \sigma^2(e, e)(n+1-i)(n-i)e^{-1}}{1 + \sigma^2(e, e)(n-i)(n-i-1)e^{-2}} \right), \quad (6.115)$$

evaluando el último productorio obtenemos

$$s_j = s_0 \left(\prod_{i=1}^j K_i \right)^{-1} e^j \left(\frac{e^2 + \sigma^2(e, e)n(n-1)}{e^2 + \sigma^2(e, e)(n-j)(n-1-j)} \right), \quad (6.116)$$

la función de Hill está definida como

$$H_{ap} = \frac{s_n}{s_0 + \sum_{j=1}^n s_j}, \quad (6.117)$$

evaluado s_j obtenemos

$$H_{ap} = \frac{\left(\prod_{i=1}^n K_i \right)^{-1} e^n \left(\frac{e^2 + \sigma^2(e, e)n(n-1)}{e^2} \right)}{1 + \sum_{j=1}^n \left(\prod_{i=1}^j K_i \right)^{-1} e^j \left(\frac{e^2 + \sigma^2(e, e)n(n-1)}{e^2 + \sigma^2(e, e)(n-j)(n-1-j)} \right)}, \quad (6.118)$$

simplificando

$$H_{ap} = \frac{e^{n-2} (e^2 + \sigma^2(e, e)n(n-1))}{\left(\prod_{i=1}^n K_i \right) + \sum_{j=1}^n \left(\prod_{i=j+1}^n K_i \right) e^j \left(\frac{e^2 + \sigma^2(e, e)n(n-1)}{e^2 + \sigma^2(e, e)(n-j)(n-1-j)} \right)}. \quad (6.119)$$

En el caso en el que $K_i = K$ y $\sigma^2(e, e) = \frac{e}{\Omega}$ (se toma de esta manera porque en limite estacionario la mayoría de las distribuciones da esta varianza) nuestra expresión se simplifica a

$$H_{ap} = \frac{e^{n-1} \left(e + \frac{n(n-1)}{\Omega} \right)}{K^n + \sum_{j=1}^n K^{n-j} e^j \left(\frac{e + \frac{n(n-1)}{\Omega}}{e + \frac{(n-j)(n-1-j)}{\Omega}} \right)}, \quad (6.120)$$

en el caso en el que $n = 4$ obtendremos

$$H_{ap} = \frac{e^3 \left(e + \frac{12}{\Omega} \right)}{K^4 + K^3 e \left(\frac{e + \frac{12}{\Omega}}{e + \frac{6}{\Omega}} \right) + K^2 e^2 \left(\frac{e + \frac{12}{\Omega}}{e + \frac{2}{\Omega}} \right) + K e^3 \left(\frac{e + \frac{12}{\Omega}}{e} \right) + e^3 \left(e + \frac{12}{\Omega} \right)}, \quad (6.121)$$

de esta manera claramente como cambia la estructura de esta nueva función de Hill y cuando Ω se vuelve muy grande se recupera la contra parte determinista. También graficaremos (6.120) para diferentes valores de n .

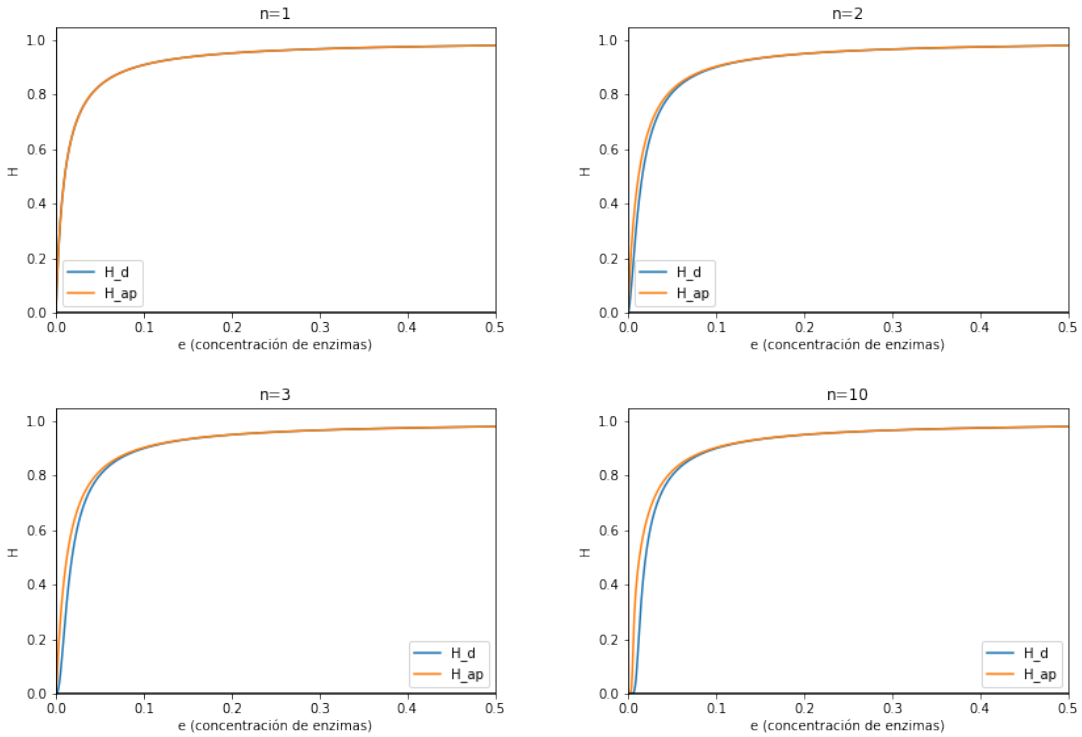


Figura 6.13: Función de Hill con unión de proteína-enzima secuencial. Se grafica la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$.

En la figura 6.13 comparamos la función de Hill determinista con procesos intermedio H_d con la que contiene correcciones estocásticas H_{ap} , podemos observar que el caso en el que $n = 1$ las dos coinciden, mientras que cuando crece n empiezan a haber diferencias entre una y otra, principalmente cuando la concentración de e es baja.

Podemos también derivar una función de Hill con correcciones estocásticas para el caso independiente, omitiremos el desarrollo algebraico ya que sería muy similar al caso anterior, el resultado

que obtendremos será

$$H_{ap} = \frac{e^{n-2} (e^2 + \sigma^2(e, e)n(n-1))}{\prod_{l=1}^n K_l + \sum_{i=1}^n \frac{n!}{i!(n-i)!} (\prod_{l=i+1}^n K_l) e^i \left(\frac{e^2 + \sigma^2(e, e)n(n-1)}{e^2 + \sigma^2(e, e)(n-i)(n-1-i)} \right)}. \quad (6.122)$$

Para el caso en el cual $K_i = K$ y $\sigma^2(e, e) = \frac{e}{\Omega}$, esta expresión se reducirá a

$$H_{ap} = \frac{e^{n-1} \left(e + \frac{n(n-1)}{\Omega} \right)}{K^n + \sum_{i=1}^n \frac{n!}{i!(n-i)!} K^{n-i} e^i \left(\frac{e + \frac{n(n-1)}{\Omega}}{e + \frac{(n-i)(n-1-i)}{\Omega}} \right)}. \quad (6.123)$$

Para ver una representación de esto vemos el caso en el cual $n = 4$, por lo que nuestra función de Hill tendrá la siguiente forma

$$H_{ap} = \frac{e^3 \left(e + \frac{12}{\Omega} \right)}{K^4 + 4K^3 e \left(\frac{e + \frac{12}{\Omega}}{e + \frac{6}{\Omega}} \right) + 6K^2 e^2 \left(\frac{e + \frac{12}{\Omega}}{e + \frac{2}{\Omega}} \right) + 4K e^3 \left(\frac{e + \frac{12}{\Omega}}{e} \right) + e^3 \left(e + \frac{12}{\Omega} \right)}. \quad (6.124)$$

de esta manera claramente como cambia la estructura de esta nueva función de Hill y cuando Ω se vuelve muy grande se recupera la contra parte determinista. También graficaremos (6.123) para diferentes valores de n .

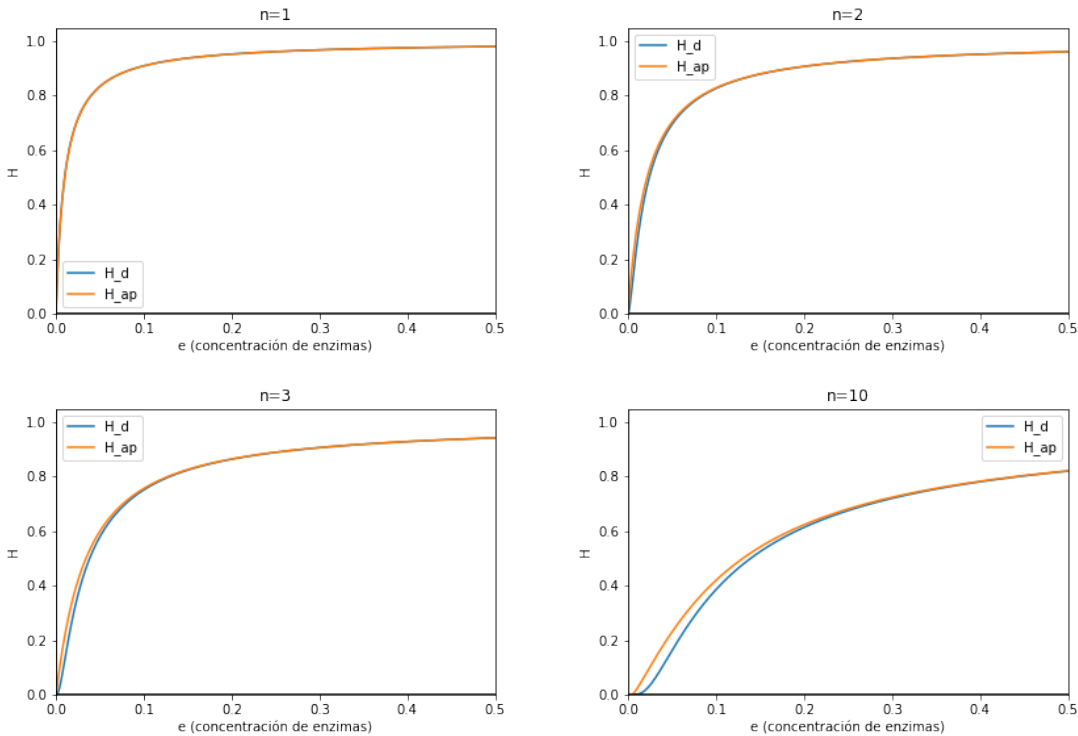


Figura 6.14: Función de Hill con unión de proteína-enzima independiente. Se grafica la función de Hill determinista H_d y función de Hill con correcciones H_{ap} para diferentes valores de n , con $K = 0,01$.

En la figura 6.13 comparamos la función de Hill determinista con procesos intermedio H_d con la que contiene correcciones estocásticas H_{ap} , podemos observar que el caso en el que $n = 1$

las dos coinciden, mientras que cuando crece n empiezan a haber diferencias entre una y otra, principalmente cuando la concentración de e es baja.

De esta manera hemos mostrado la función de Hill que considera procesos intermedios y correcciones estocásticas, esta función de Hill puede llegar a ser usada en redes de regulación genética o para describir la unión enzimática, aunque estas funciones de Hill no las usaremos directamente en este trabajo. Otra propiedad que estas funciones que hemos derivado son

$$\frac{\partial^k}{\partial x^k} H_{ap}(x) = 0, \tag{6.125}$$

($k > 0$) las derivadas parciales de cualquier orden valen cero, esto nos ayudara a cuantificar mejor las fluctuaciones del sistema, como ya hemos mencionado.

Capítulo 7

Redes estocásticas de regulación genética

Ya hemos descrito las redes genéticas en el capítulo de biología sintética, incluso presentamos algunos ejemplos simplificados, ahora con todas las herramientas y conceptos desarrollados hasta este punto construiremos y describiremos algunos modelos estocásticos de las redes de regulación genéticas, esto con el objetivo de conocer la dinámica del sistema cuando están presentes las fluctuaciones y además cuantificarlas durante toda la evolución del sistema, para esto describiremos como se construye la ecuación maestra de estas, además elaboraremos un programa en Python para poder hacer simulaciones del sistema usando el algoritmo de Gillespie así como el uso de las ecuaciones diferenciales del teorema de fluctuación disipación. Recordar que los códigos que usamos para simular se encuentran en la carpeta de Drive (https://drive.google.com/drive/folders/1DcDw7D0B-hq-9gxt-8z4P3_K63m1eU01?usp=sharing).

7.1. Introducción

La expresión génica es un proceso fundamentalmente estocástico, con aleatoriedad en la transcripción y traducción que conduce a variaciones de célula a célula en los niveles de ARNm y proteínas. Esta variación aparece en organismos que van desde microbios hasta metazoos, y sus características dependen tanto de los parámetros biofísicos que gobiernan la expresión génica como de la estructura de la red de genes. La expresión génica estocástica tiene importantes consecuencias para la función celular, siendo beneficiosa en algunos contextos y dañina en otros. Estas situaciones incluyen la respuesta al estrés, el metabolismo, el desarrollo, el ciclo celular, los ritmos circadianos y el envejecimiento [27].

Existen varios modelos estocásticos de expresión genética, algunos de los modelos se toman las mismas consideraciones que en el caso determinista, estos modelos por lo general son válidos solo para procariotas, sin embargo, para las eucariotas estos modelos no reproducen los datos experimentales, por lo que se han propuesto dos hipótesis [27]:

i) **Burst.**

Se ha observado que existe una variabilidad en la expresión genética, una posible explicación de esto se debe a que la producción de proteínas no se realiza de manera continua, por lo que se asume que el gen transita entre estados de activación y desactivación.

ii) **Chromatin remodeling.**

Cuando la cromatina circundante está en un estado acetilado abierto, el gen puede transcribir

con relativa libertad, mientras que cuando la cromatina está en un estado condensado, la transcripción se reprime.

En este trabajo no asumiremos tales hipótesis, a continuación, describiremos con más detalle los módulos de transcripción/traducción (MTT) enfocándonos en modelos que se parecen al caso determinista, exactamente en aquellos que al tomar promedios a partir de la ecuación maestra se recupera el caso determinista.

7.2. Red estocástica de regulación genética I

Ya hemos presentado en uno de los capítulos anteriores un modelo determinista de un MTT, se puede revisar el capítulo 2 para una explicación más detallada de los procesos que ocurren en un MTT. Ahora nos daremos a la tarea de encontrar uno estocástico esto con el fin de llegar a cuantificar las fluctuaciones del sistema durante toda la evolución dinámica, para lograrlo nos apoyaremos del algoritmo de Gillespie y el TFD, en este último caso usaremos la función de Hill con correcciones estocásticas.

Para describir a estas redes nos apoyamos de la figura 7.1, esta figura representa un modelo simple y simplificado de lo que ocurre en un MTT, aquí el sistema empieza a funcionar por la función de entrada, por lo regular esta función de entrada es una función de Hill (repressor o activador según sea el caso), esta función de entrada depende de la síntesis basal y de factores de transcripción (ver apéndice B), en este primer caso que analizaremos suponemos que una sola proteína es la que actúa sobre el MTT.

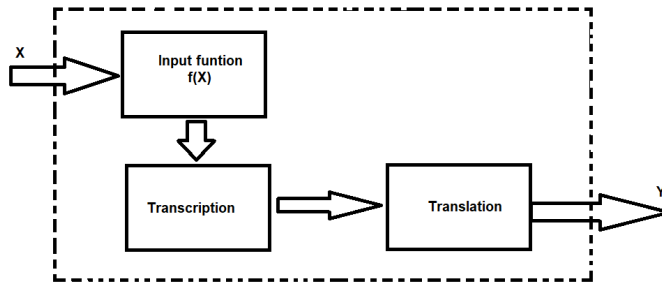


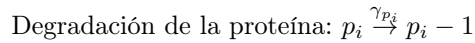
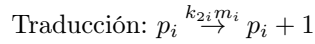
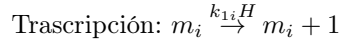
Figura 7.1: Módulo de transcripción/traducción [8].

Para una mejor perspectiva de lo que pasa, se pueden usar las siguientes reacciones, estas son completamente deterministas, pero describen mejor el funcionamiento el proceso de transcripción y traducción.



En las tasas de reacción que activan al gen, la H es la función de Hill (notar que aún no hemos especificado que tipo de función de Hill, esta puede ser una determinista, semi-determinista o con correcciones estocásticas), esta puede ser activadora o supresora según el sistema en particular, en el apéndice B mostramos porqué se usa en las redes de regulación genética. A partir de este punto supondremos que tenemos N módulos, con $N \geq 2$, muy similar a la figura 7.2. Para el caso en el que tenemos un solo modulo, presentaremos un modelo más detallado en la siguiente sección, ahora

nombraremos con más detalle las reacciones que hemos presentado, por lo que para el i -ésimo modulo ocurren las siguientes reacciones



En estas relaciones las funciones de activación tienen asociadas unas funciones de entrada que sería un represor si la proteína suprime la activación genética, en caso contrario sería un activador. Ahora nos pondremos a derivar la ecuación maestra de la red de regulación genética en el cual una sola proteína actúa como supresora o activadora sobre un MTT, pero antes nos preguntamos como saber que proteína actúa sobre que modulo o módulos, para ello nos inspiramos en la teoría de grafos [38], particularmente en la matriz de adyacencia A , por lo que si la proteína p_l actúa como activador sobre el MTT i entonces la entrada i, l de la matriz será uno, es decir $A_{i,l} = 1$, cuando sea un represor $A_{i,l} = -1$ (notar que $A_{i,l} \neq A_{l,i}$). Un ejemplo de una red genética se puede ver en la figura 7.2, el cual corresponde a la bacteria E. Coli, una bacteria muy usada a nivel experimental [5].

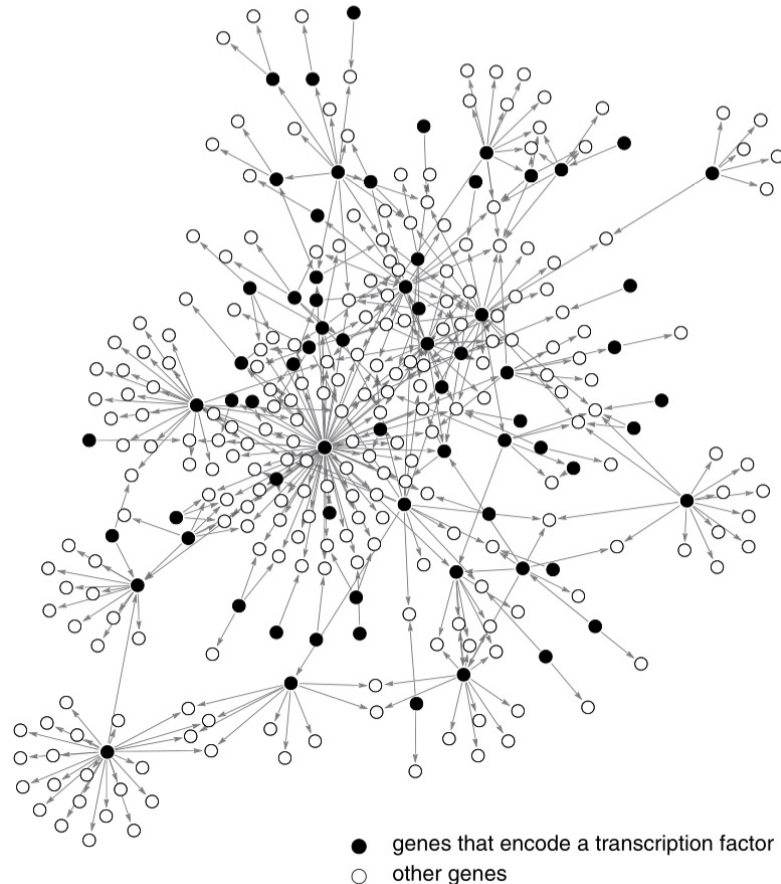


Figura 7.2: Red genética de la E. Coli [4].

Capturaremos la actividad de la proteína p_l sobre algún modulo i a través de una función de Hill, pero como estamos particularmente interesados en construir un modelo estocástico, no se puede introducir directamente la función de Hill en las tasas de propensión, pero esta cuestión ya la abordamos en el capítulo anterior en donde se expone el cómo estas aparecen en un proceso estocástico, por lo que ahora tenemos la libertad de usarlas (recordar que hay dos opciones de como aparece la función de Hill, por lo que hay que elegir de acuerdo a las necesidades y objetivos, hacer simulaciones con el algoritmo de Gillespie o encontrar el límite determinista). Suponemos que la proteína p_l producida por el MTT l actúa solo como represor (D) o activador (H) sobre otro MTT i pero no las dos al mismo tiempo y además es un solo tipo de proteína es la que actúa sobre nuestro modulo, para el caso en el que actúen varios presentaremos un modelo más adelante. Con ayuda de la matriz de conexiones A_{ij} definiremos una función de Hill determinista,

$$H_i = \frac{\sum_j |A_{ij}| \left(\frac{p_j}{\Omega K_j} \right)^{n_{ij} * A_{ij}}}{1 + \sum_j |A_{ij}| \left(\frac{p_j}{\Omega K_j} \right)^{n_{ij} * A_{ij}}}, \quad (7.2)$$

con los valores de A_{ij} recuperamos las funciones de Hill para una supresora y activadora, el termino n_{ij} corresponde al coeficiente de Hill debido al módulo j , las K_i son las constantes de Michelson-Menten y son debido al módulo j . Si se consideran correcciones estocásticas se introducirá una función similar a

$$H_i^e = \frac{\sum_j |A_{ij}| \left(\frac{\langle p_j \rangle^{n_{ij}} + \sigma^2(p_j, p_j) n_{ij} (n_{ij} - 1) \langle p_j \rangle^{(n_{ij} - 2)}}{(\Omega K_j)^{n_{ij}}} \right)^{A_{ij}}}{1 + \sum_j |A_{ij}| \left(\frac{\langle p_j \rangle^{n_{ij}} + \sigma^2(p_j, p_j) n_{ij} (n_{ij} - 1) \langle p_j \rangle^{(n_{ij} - 2)}}{(\Omega K_j)^{n_{ij}}} \right)^{A_{ij}}}. \quad (7.3)$$

Esta función de Hill con correcciones estocásticas la derivamos en el capítulo anterior, notar que esta depende de unos promedios por lo que es complicada de usar al momento de emplear el algoritmo de Gillespie, la parte realmente relevante de esta función viene cuando se llega a tomar el límite determinista.

Usando los elementos y las ideas que hemos desarrollado hasta este punto vamos a construir nuestra ecuación maestra, con ayuda las reacciones que están al principio de esta sección y asumiendo que el proceso es Markoviano, se desprende la siguiente ecuación maestra, donde $\tau_{1i} = \frac{1}{\gamma_{m_i}}$ y $\tau_{2i} = \frac{1}{\gamma_{p_i}}$ son los tiempos de degradación del ARm y la proteína respectivamente,

$$\begin{aligned} \frac{\partial P(m_i, p_i, t)}{\partial t} = & \sum_i (\Omega k_{1i} (H_i(p_l, \Omega) P(m_i - 1, p_i, t) - H_i(p_l, \Omega) P(m_i, p_i, t))) \\ & + \frac{m_i + 1}{\tau_{1i}} P(m_i + 1, p_i, t) - \frac{m_i}{\tau_{1i}} P(m_i, p_i, t) \\ & + k_{2i} m_i P(m_i, p_i - 1, t) - k_{2i} m_i P(m_i, p_i, t) \\ & + \frac{p_i + 1}{\tau_{2i}} P(m_i, p_i + 1, t) - \frac{p_i}{\tau_{2i}} P(m_i, p_i, t). \end{aligned} \quad (7.4)$$

Notar que la suma se realiza con respecto a cada módulo. Una observación muy importante que queremos recalcar es que la expresión (7.4) nunca ha sido derivada con el enfoque que hemos optado, en la literatura se acostumbra a usar operadores para suplir la función de Hill [39, 40] o usar directamente la función de Hill determinista [24], sin embargo, nosotros tenemos la opción de usar varias, particularmente la que contiene correcciones estocásticas que nos conecta de manera más clara con el límite determinista, lo cual resulta muy útil al momento de comparar los dos formalismos, el determinista y el estocástico.

Una cantidad muy importante y que nos permitirá continuar analizando este tipo de redes es la distribución estacionaria, para ello igualamos la ecuación (7.4) a cero y posteriormente aplicamos

el método que hemos desarrollado en el capítulo de análisis estocástico, es importante notar que las reacciones (7.1) no tienen deficiencia cero, pero siguiendo al artículo [20] se pueden transformar de tal manera que si la tengan (ver apéndice C). La distribución estacionaria que encontramos es la siguiente

$$P(m_i, p_i) = \prod_i \frac{(k_{1i}\tau_{1i}\Omega \langle H_i(p_l, \Omega) \rangle)^{m_i} e^{-k_{1i}\tau_{1i}\Omega \langle H_i(p_l, \Omega) \rangle}}{m_i!} \frac{(k_{2i}\tau_{2i} \langle m_i \rangle)^{p_i} e^{-k_{2i}\tau_{2i} \langle m_i \rangle}}{p_i!}. \quad (7.5)$$

Podemos fijarnos que la distribución es la multiplicación de dos de tipo Poisson para cada módulo, notar también que la multiplicación que aparece es sobre todos los módulos. Tener esta distribución nos permitirá encontrar los promedios, las covarianzas y otros elementos estadísticos. El promedio de las variables del módulo i del sistema será

$$\langle m_i \rangle = k_{1i}\tau_{1i}\Omega \langle H_i(p_l, \Omega) \rangle, \quad \langle p_i \rangle = k_{2i}\tau_{2i} \langle m_i \rangle. \quad (7.6)$$

Notar que en el caso en el que se usó la función de Hill determinista tendremos que usar la siguiente aproximación

$$\langle f(X_l) \rangle \approx f(\langle X_l \rangle) + \sum_i \sum_j \frac{(\sigma^2(X_i, X_j))}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle},$$

por lo que si se usa la función de Hill determinista tendremos

$$\langle H_i(p_l, \Omega) \rangle \approx H_i(\langle p_l \rangle, \Omega) + \sum_i \sum_j \frac{(\sigma^2(p_i, p_j))}{2} \frac{\partial^2 H_i(p_l, \Omega)}{\partial p_i \partial p_j} \Big|_{p_l = \langle p_l \rangle}, \quad (7.7)$$

mientras que si se usa una con correcciones estocásticas se tendrá

$$\langle H_i^e(p_l, \Omega) \rangle = H_i^e(p_l, \Omega). \quad (7.8)$$

También encontramos las fluctuaciones en el estado estacionario, para ello usaremos la distribución estacionaria, calculando las varianzas y luego dividiendo entre Ω^2 , obtenemos

$$\eta_{m_i}^2 = \frac{\langle m_i \rangle}{\Omega^2}, \quad \eta_{p_i}^2 = \frac{\langle p_i \rangle}{\Omega^2}. \quad (7.9)$$

Algo interesante que se encuentra de este tipo de sistemas es que cada una de las fluctuaciones es contribución de su propia variable, una de las razones por la que el sistema tiene esta característica se debe a que el sistema es descrito puramente por una distribución de Poisson (para cada variable). Debido a que los promedios son proporcionales a las variables deterministas $\langle x_i \rangle = \Omega x_i$ entonces las fluctuaciones solo dependerán del inverso de Ω , entonces cuando el tamaño del sistema sea lo suficientemente grande las fluctuaciones se volverán muy pequeñas, lo que hace que en los sistemas deterministas no se tomen en cuenta, sin embargo, cuando el sistema es pequeño estas fluctuaciones son relevantes. Dado que hemos considerado que las fluctuaciones intrínsecas son todas aquellas debido a los elementos que conforman nuestro sistema, podemos decir que las fluctuaciones que hemos calculado son intrínsecas. Nos gustaría recalcar que las fluctuaciones que hemos calculado son válidas para cualquier tipo de redes que cumpla las condiciones que mencionamos. Las fluctuaciones en el estado estacionario nos darán una idea de que modelo es el más correcto, es decir, que tipo de función de Hill es la que mejor coincide con las fluctuaciones del estado estacionario, esto lo veremos más adelante con un ejemplo en particular.

7.2.1. Elementos para la simulación

Podemos realizar las simulaciones de nuestro sistema a través de dos métodos, el primero emplear el algoritmo de Gillespie (por lo que tendremos que usar a fuerza cualquier función de Hill

excepto la que contiene correcciones estocásticas) y de ahí inferir los promedios y las fluctuaciones de las variables del sistema, lo otro es emplear la aproximación determinista y al teorema de fluctuación disipación (TFD) (permitiéndonos usar cualquier función de Hill, siendo la más relevante la que contiene correcciones estocásticas), a continuación, mostraremos cada uno de estos métodos.

Algoritmo de Gillespie

Con todo lo anterior ya tenemos el desarrollo teórico para poder escribir el programa que realice las simulaciones estocásticas del sistema a través del algoritmo de Gillespie, de hecho ya tenemos una muy buena parte ya hecha en el capítulo 3, en donde hablamos del algoritmo de Gillespie, para hacer simulaciones con este algoritmo es necesario la matriz estequiométrica y el vector de propensión del sistema, para ello construiremos cada uno de estos, primero nos fijamos en las reacciones que suceden en cada uno de los nodos, para cada uno de los nodos se tendría la siguiente matriz estequiométrica y su correspondiente vector de propensión

$$S_n = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad \nu_n = \{\Omega k_{1n} H_i(p_l, \Omega), \gamma_{m_n} m_n, k_{2n} m_n, \gamma_{p_n} p_n\}, \quad (7.10)$$

(en el caso que un nodo actué como represor se modificaría la función de Hill) por lo que ahora el vector de propensión y la matriz estequiométrica de todo el sistema será

$$S = \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & S_n \end{pmatrix}, \quad \nu = \{\nu_1, \nu_2, \dots, \nu_n\}, \quad (7.11)$$

con estos elementos podemos implementar el algoritmo de Gillespie para redes genéticas, para saber cómo sería la dinámica de estos sistemas analizaremos uno de estos a continuación.

Aproximación determinista

La otra forma de modelar a las redes de regulación genética es a través de la aproximación determinista y el TFD, aunque se seguirá usando a S que aparece en (7.11), mientras que el vector de propensión se le tomará el límite determinista para encontrar las tasas de reacción, de tal manera que obtendremos algo similar a

$$t_n = \{k_{1n} H_i(\hat{p}_l, \Omega), \gamma_{m_n} \hat{m}_n, k_{3n} \hat{m}_n, \gamma_{p_n} \hat{p}_n\} \quad (7.12)$$

aquí las variables \hat{g}_n , \hat{m}_n y \hat{p}_n denotan las concentraciones deterministas, la función de Hill sería la determinista o la que considera correcciones estocásticas. En general usaremos los siguientes dos elementos para este tipo de redes

$$S = \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & S_n \end{pmatrix}, \quad t_+ = \{t_1, t_2, \dots, t_n\}, \quad (7.13)$$

A continuación, analizaremos un sistema en particular para ver qué resultados se obtienen.

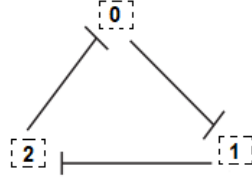


Figura 7.3: Represilador.

7.2.2. Represilador

El sistema que analizaremos será el Represilador, para ello primero simularemos el sistema usando el algoritmo de Gillespie para ello usaremos la función de Hill determinista, posteriormente también modelaremos el sistema con la aproximación determinista y el TFD, donde usaremos la función de Hill determinista y la que contiene correcciones estocásticas. Ahora aplicaremos este formalismo para construir la ecuación maestra del represilador, una representación gráfica de esto puede verse en la figura 7.3, posteriormente haremos un programa en Python (*redgefturepresilador.ipynb*) para hacer una simulación estocástica del sistema, este programa promediara las concentraciones del ARm y las proteínas sobre un determinado ensamble. Una de las razones por las que hemos elegido a este sistema es porque es relativamente simple de implementar y se ha construido de manera experimental [5]. Con ayuda de la figura 7.3 construimos nuestra matriz de conexiones A

$$A_{ij} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}, \quad (7.14)$$

lo que nos dice la primera fila de la matriz A es que una proteína sintetizada en el módulo C actúa como supresora sobre el módulo A, y de manera similar para los otros módulos. Dado que ya hemos derivado una formula general para determinar la ecuación maestra de una red genética en el que una proteína actúa sobre un MTT, los únicos elementos que necesitamos son los parámetros del sistema y la matriz de conexiones A , asumimos que $k_{1i}^- = k_1^-$, por lo que ahora la ecuación maestra de nuestro sistema será

$$\begin{aligned} \frac{\partial P(m_i, p_i, t)}{\partial t} = & \sum_{i=1}^3 ((k_1^- \Omega(D_i(p_l, \Omega)P(m_i - 1, p_i, p_l, t) - D_i(p_l, \Omega)P(m_i, p_i, p_l, t))) \\ & + \frac{m_i + 1}{\tau_{2i}} P(m_i + 1, p_i, t) - \frac{m_i}{\tau_{2i}} P(m_i, p_i, t) \\ & + k_{3i} m_i P(m_i, p_i - 1, t) - k_{3i} m_i P(m_i, p_i, t) \\ & + \frac{p_i + 1}{\tau_{3i}} P(m_i, p_i + 1, t) - \frac{p_i}{\tau_{3i}} P(m_i, p_i, t)). \end{aligned} \quad (7.15)$$

A partir de esta ecuación es relativamente más sencillo obtener las ecuaciones deterministas del represilador, esto se logra a través del límite determinista que hemos explicado anteriormente, incluso se pueden introducir correcciones estocásticas, para obtener el límite determinista multiplicamos a toda la expresión anterior por alguna de las variables, dividimos entre Ω y luego sacamos el promedio a toda la expresión. Podemos incluso explorar cómo se comporta el sistema, para ello vamos a realizar algunas simulaciones del sistema, entonces nuestra matriz estequiométrica y el

vector de propensión serán

$$S = \begin{pmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{pmatrix}, \quad \nu = \{\nu_1, \nu_2, \nu_3\}, \quad (7.16)$$

(los elementos que la componen son los mismos que están en la subsección anterior) en el lugar que va la función de Hill hemos usado la que corresponde al caso de un represilador (una supresora). Como simularemos el sistema usando el algoritmo de Gillespie escribimos un programa de Python, el programa que escribimos se encuentra en la carpeta de Drive (https://drive.google.com/drive/folders/1DcDw7D0B-hq-9gxt-8z4P3_K63m1eU01?usp=sharing) con el nombre de *redgflurepresilador.ipynb*.

Lo que se obtiene de este programa son las figuras 7.4, en este caso hemos usado la función de Hill determinista, el tamaño del sistema es 200, además solo hemos graficado la concentración de proteínas y ARm, debido a que estas son las cantidades que se miden al momento de realizar los experimentos, además hemos usado una función de Hill determinista. En estas figuras claramente se nota el comportamiento oscilatorio en la concentración de las proteínas, debido a que el sistema es pequeño las fluctuaciones están presentes.

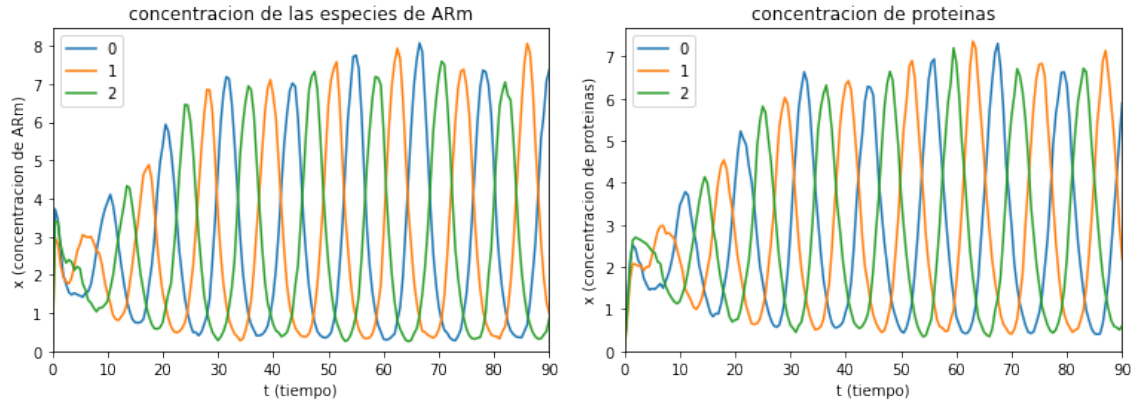


Figura 7.4: Represilador. Simulación estocástica usando al algoritmo de Gillespie y la función de Hill determinista, en la primera figura se observa la concentración de ARm, en la segunda la concentración de las proteínas. Se han usado los siguientes parámetros: $n = 2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$.

Podemos analizar las fluctuaciones de este sistema, para ello agregamos a nuestro código la parte que hace la inferencia estadística en el sistema (para los detalles del código revisar la carpeta de Drive), lo que se obtiene está en la figura 7.5, en las figuras de la parte superior se observa a la concentración del ARm y a en las de abajo a las de la proteína, la línea verde muestra el promedio de las variables sobre el ensamble, mientras que las líneas punteadas nos dicen que en este intervalo se encontraran la mayoría de las fluctuaciones del sistema, podemos ver que a pesar de que estas bandas son muy grandes, nuestro sistema en general sigue una tendencia oscilatoria en todas las especies bioquímicas involucradas, una de las razones por la que elegimos este tamaño de sistema se debe a que se usó este tamaño de sistema para el experimento de Elowitz [5]. Con estos resultados hemos cuantificado las fluctuaciones intrínsecas del sistema (aún estamos usando la función de Hill determinista).

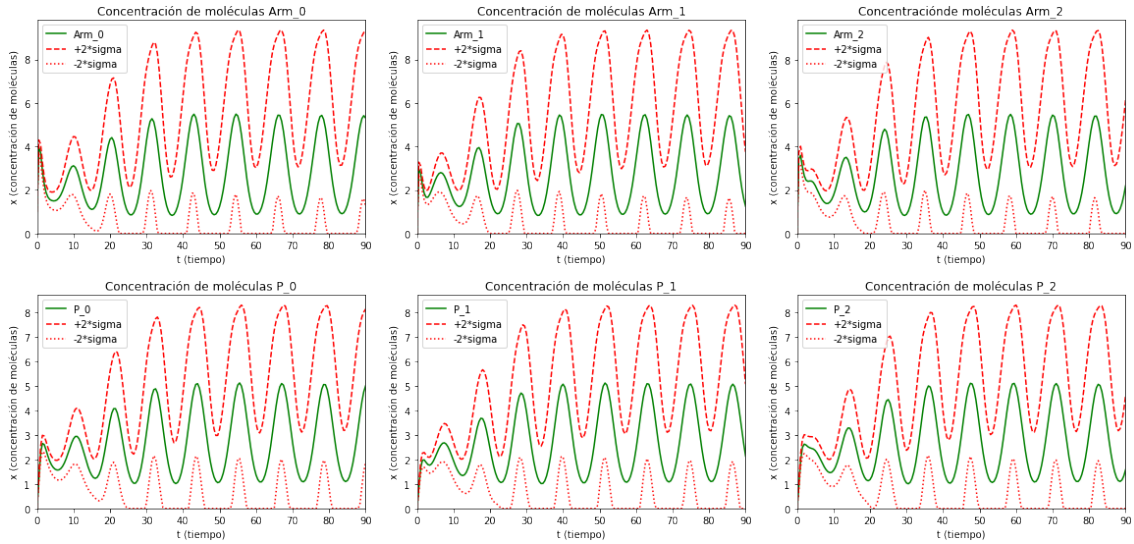


Figura 7.5: Represilador. Dinámica de las concentraciones deterministas y las fluctuaciones a partir de simulaciones estocásticas con tamaño de ensamble= 10000, en las figuras de la parte superior se observa la dinámica de los ARM´s, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas en donde se encuentran la banda de las fluctuaciones. Se han usado los siguientes parámetros: $Arm_2(0) = r_1(0) = r_2(0) = 0$, $Arm_1(0) = \Omega$, $n=2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$.

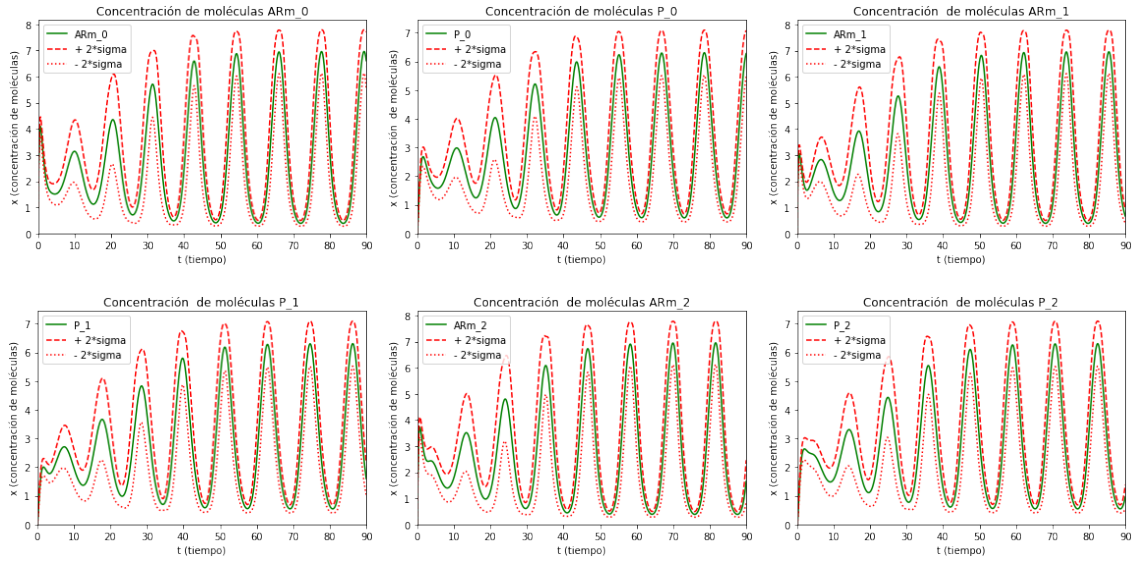
Ahora usaremos la aproximación determinista y el TFD para modelar este sistema, para ello tomamos los promedios del vector de propensión y luego dividimos entre Ω para obtener las tasas de reacción. Escribimos un programa de Python, el programa que escribimos se encuentra en la carpeta de Drive con el nombre de *arepresilador.ipynb*.

Lo que obtenemos cuando se usa la función de Hill determinista se puede ver en la figura 7.6 (a). Es importante mencionar que el tiempo de computación del sistema se redujo dramáticamente al usar la aproximación determinista y el TFD en comparación con el algoritmo de Gillespie. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas crean una banda en donde se encontrarían todas las fluctuaciones del sistema. Podemos ver que se sigue observando la tendencia oscilatoria en todas las especies bioquímicas del sistema, aunque los promedios de las concentraciones varían con respecto a lo que obtuvimos con el algoritmo de Gillespie, también la banda que representa a las fluctuaciones es diferente, hay tres posibles razones detrás de estas diferencias, el primero es el ensamble que usamos no es lo suficientemente grande para capturar toda la información del sistema, el segundo es que el teorema de fluctuación-disipación que usamos es solo una aproximación por lo que no podría capturar toda la información del sistema.

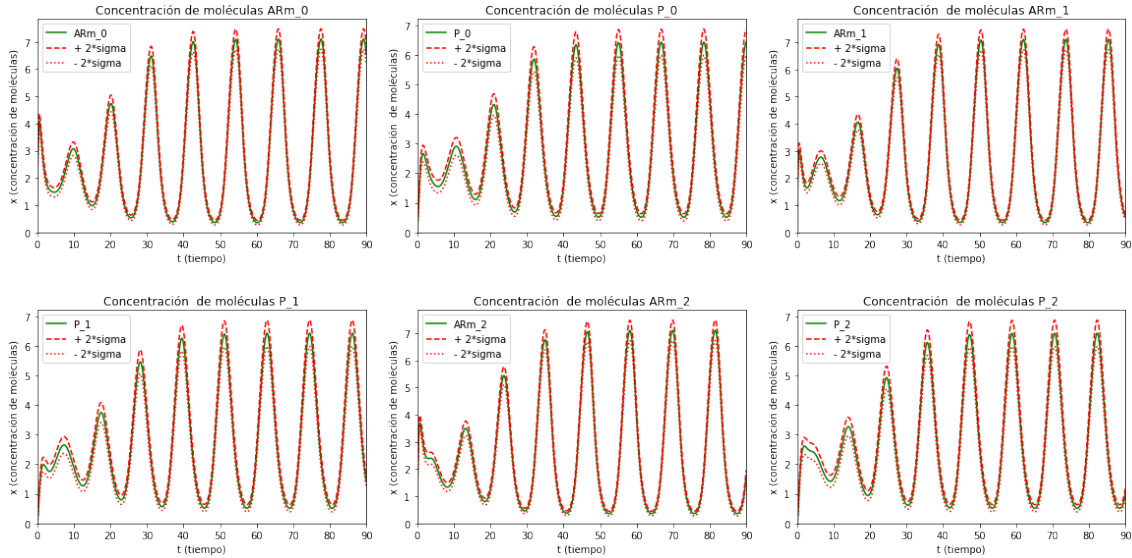
Lo que obtenemos cuando se usa la función de Hill con correcciones estocásticas se puede ver en la figura 7.6 (b). Lo primero que observamos es que ahora el tamaño de las fluctuaciones se ha reducido considerablemente con respecto a las figuras 7.5 y 7.6 (a), una de las razones se debe al tipo de función de Hill que hemos usado, en esta última hemos usado a la función de Hill con correcciones estocásticas, por lo que al momento de usar el TFD se han omitido algunos términos. De hecho, este último resultado que hemos obtenido se acerca más a lo obtenido con las ecuaciones deterministas. Como hemos venido mencionando el usar la función de Hill con correcciones estocásticas es la más correcta si se quiere considerar un sistema estocástico, esto también se puede ver porque en el estado estacionario o cerca del estado estacionario las fluctuaciones deberían de

Redes estocásticas de regulación genética

7.2 Red estocástica de regulación genética I



(a) Función de Hill determinista.



(b) Función de Hill con correcciones estocásticas.

Figura 7.6: Represilador. Se ha modela la dinámica y las fluctuaciones del sistema usando la aproximación determinista y el TFD, en la figura (a) se ha usado la función de Hill determinista, en (b) la función de Hill con correcciones estocásticas. Tanto para (a) y (b) las figuras de la parte superior se observa la dinámica de los ARm's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $Arm_2(0) = r_1(0) = r_2(0) = 0$, $Arm_1(0) = 1$, $n=2$, $\Omega = 200$, $\tau_{2i} = \tau_{3i} = 1$, $k_{2i} = k_{3i} = 10$.

dependen como $\sqrt{\frac{\hat{x}}{\Omega}}$, y el modelo que más se acerca a esto es en el que se usó la función de Hill con correcciones estocásticas, ya que si se usa la función de Hill determinista existe una sobreestimación.

De esta manera hemos cuantificado las fluctuaciones del represilador durante toda la dinámica del sistema, ahora podemos analizar cualquier red de regulación genética en el que una sola proteína actúe como factor de transcripción en un MTT y cuantificar de manera simultánea las fluctuaciones de sus variables, como hemos visto en estos ejemplos el introducir la función de Hill determinista directamente en nuestro modelo estocástico cambia considerablemente el tamaño de las fluctuaciones en comparación de la función de Hill que contiene correcciones estocásticas, usar esta función con correcciones es el enfoque más correcto ya que detrás de esta hay una derivación sistemática. El análisis de otro tipo de redes lo dejaremos para trabajos futuros.

7.3. Red estocástica de regulación genética general

En una de las secciones anteriores ya hemos presentado un modelo estocástico de una red genética cuando un solo tipo de proteína actúa sobre otro MTT, sin embargo, existen modelos y/o sistemas en los cuales es posible que simultáneamente actúen diferentes proteínas y de diferente manera, esta asunción es lo único que cambia con respecto a la red estocástica de regulación genética I, lo que se traduciría en usar una función de Hill diferente como veremos más adelante, por lo que seguimos prácticamente un desarrollo similar al de la sección anterior.

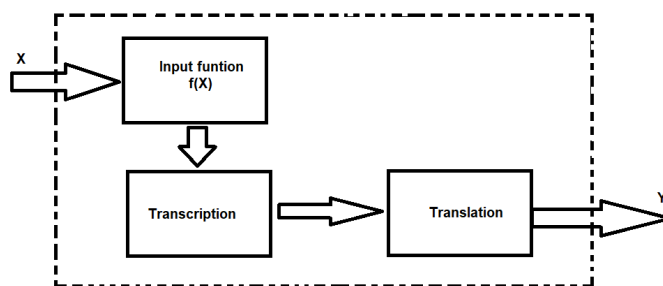
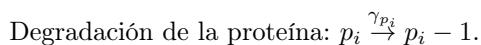
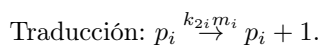
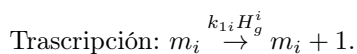


Figura 7.7: Módulo de transcripción/traducción [8].

Primero nos apoyaremos de nuevo de la matriz A que ya hemos definido en una de las secciones anteriores, esta matriz nos ayuda a entender cómo actúan determinadas proteínas sobre otras, donde las entradas de esta matriz serán 1 o 0, cuando $A_{ij} = 1$ actúa como activador, mientras que cuando $A_{ij} = -1$ como supresora, debido a que queremos calcularlo para una red que sea lo suficientemente general, nuestro modelo incluirá los casos de auto activación y auto supresión (recordar que A en general no es simétrica), para ello nos apoyamos de las siguientes reacciones:



Estas reacciones son justo las que corresponden al i -ésimo MTT, hemos usado una nueva función de Hill el cual considera que varias proteínas actúen como factores de transcripción de manera simultánea sobre un MTT (notar que aún no hemos dicho si esta función de Hill es determinista o con correcciones estocásticas), una derivación de esta se puede encontrar en el apéndice B. Las

entradas de esta función son las proteínas y un vector correspondiente a la i -ésima fila de la matriz A , de tal manera que la función de Hill (determinista) para el módulo i está definida como:

$$H_g^i(p_l, \Omega) = \frac{\beta_0 K_0 + \sum_j \theta(A_{ij}) * \beta_{1j} \left(\frac{p_j}{K_j \Omega} \right)^{n_{ij}}}{1 + K_0 + \sum_j \theta(|A_{ij}|) \beta_{2j} \left(\frac{p_j}{K_j \Omega} \right)^{n_{ij}}}, \quad (7.17)$$

el termino n_{ij} corresponde al coeficiente de Hill debido al módulo j , las K_i son las constantes de Michelson-Menten y son debido al módulo j . Las constantes β_{1j} , β_{2j} , β_0 y K_0 se eligen de acuerdo con el sistema que se esté analizando. En este punto nos podemos dar cuenta que nuestro sistema solo varia en la forma de función de Hill, esto con respecto de la ecuación maestra para una red estocástica de regulación genética en el cual una sola proteína actúa sobre un MTT. Podemos incluso usar una función de Hill con correcciones estocásticas, entonces tendremos una función de la forma

$$H_{gue}^i(p_l, \Omega) = \frac{\beta_0 K_0 + \sum_j \theta(A_{ij}) * \beta_{1j} \left(\frac{\langle p_j \rangle^{n_{ij} + \sigma^2(p_j, p_j) n_{ij} (n_{ij} - 1)} \langle p_j \rangle^{n_{ij} - 2}}{(K_j \Omega)^{n_{ij}}} \right)}{1 + K_0 + \sum_j \theta(|A_{ij}|) \beta_{2j} \left(\frac{\langle p_j \rangle^{n_{ij} + \sigma^2(p_j, p_j) n_{ij} (n_{ij} - 1)} \langle p_j \rangle^{n_{ij} - 2}}{(K_j \Omega)^{n_{ij}}} \right)}. \quad (7.18)$$

Usar esta función de Hill para hacer simulaciones estocásticas resulta una tarea compleja, sin embargo, cuando se toma el límite determinista de la ecuación maestra que la usa se verá su relevancia.

Con ayuda de las reacciones que están al principio de esta sección y asumiendo que el proceso es Markoviano, se desprende la siguiente ecuación maestra de una red de regulación genética, donde $\tau_{1i} = \frac{1}{\gamma_{m_i}}$ y $\tau_{2i} = \frac{1}{\gamma_{p_i}}$ son los tiempos de degradación del ARm y la proteína respectivamente,

$$\begin{aligned} \frac{\partial P(m_i, p_i, t)}{\partial t} = & \sum_i \left(\Omega k_{1i} \left(H_g^i(p_l, \Omega) P(m_i - 1, p_i, t) - H_g^i(p_l, \Omega, A_i) P(m_i, p_i, t) \right) \right. \\ & + \frac{m_i + 1}{\tau_{2i}} P(m_i + 1, p_i, t) - \frac{m_i}{\tau_{2i}} P(m_i, p_i, t) \\ & + k_{2i} m_i P(m_i, p_i - 1, t) - k_{3i} m_i P(m_i, p_i, t) \\ & \left. + \frac{p_i + 1}{\tau_{3i}} P(m_i, p_i + 1, t) - \frac{p_i}{\tau_{3i}} P(m_i, p_i, t) \right). \end{aligned} \quad (7.19)$$

Al momento de tomar los promedios de las variables involucradas en este sistema recuperamos la versión determinista de este tipo de redes. De esta manera ya hemos construido nuestra red genética estocástica general, es decir puede ser aplicado a cualquier red tan solo es necesario la matriz A_{ij} , así como las constantes involucradas. Nos gustaría recalcar que esta ecuación maestra aún no ha sido presentada en la literatura ya que por lo general se introduce la función de Hill determinista sin una correcta justificación [24], mientras que nosotros tenemos la opción de usar incluso una que contenga correcciones estocásticas. Las ideas que hemos desarrollado hasta este punto pueden servir para desarrollar modelos similares en aquellos que empleen funciones de Hill.

Una cantidad muy importante y que nos permitirá continuar analizando este tipo de redes es la distribución estacionaria, para ello igualamos la ecuación (7.19) a cero y posteriormente empleamos el método que hemos desarrollado para calcular distribuciones estacionarias, de tal manera que encontramos

$$P_S(m_i, p_i) = \prod_i \frac{(k_{1i} \tau_{1i} \Omega \langle H_i^g(p_l, \Omega) \rangle)^{m_i} e^{-k_{1i} \tau_{1i} \langle H_i^g(p_l, \Omega) \rangle}}{m_i!} \frac{(k_{2i} \tau_{2i} \langle m_i \rangle)^{p_i} e^{-k_{2i} \tau_{2i} \langle m_i \rangle}}{p_i!}. \quad (7.20)$$

Esta distribución es la misma que el de la sección anterior salvo por la función de Hill. Con esta distribución estacionaria podemos calcular diferentes cantidades estadísticas como el promedio y

la varianza. El promedio de las variables del módulo i del sistema será

$$\langle m_i \rangle = k_{1i} \tau_{1i} \langle H_i^g(p_l, \Omega) \rangle, \quad \langle p_i \rangle = k_{2i} \tau_{2i} \langle m_i \rangle. \quad (7.21)$$

Notar que en el caso en el que se usando la función de Hill determinista tendremos que usar la siguiente aproximación

$$\langle f(X_l) \rangle \approx f(\langle X_l \rangle) + \sum_i \sum_j \frac{(\sigma^2(X_i, X_j))}{2} \frac{\partial^2 f(X_l)}{\partial X_i \partial X_j} \Big|_{X_l = \langle X_l \rangle},$$

por lo que si se usa la función de Hill determinista tendremos

$$\langle H_g^i(p_l, \Omega) \rangle \approx H_g^i(\langle p_l \rangle, \Omega) + \sum_i \sum_j \frac{(\sigma^2(p_i, p_j))}{2} \frac{\partial^2 H_g^i(p_l, \Omega)}{\partial p_i \partial p_j} \Big|_{p_l = \langle p_l \rangle}, \quad (7.22)$$

mientras que si se usa una con correcciones estocásticas se tendrá

$$\langle H_{gue}^i(p_l, \Omega) \rangle = H_{gue}^i(\langle p_l \rangle, \Omega). \quad (7.23)$$

También encontramos las fluctuaciones en el estado estacionario, para ello usaremos la distribución estacionaria, calculando las varianzas y luego dividiendo entre Ω^2 , obtenemos

$$\eta_{m_i}^2 = \frac{\langle m_i \rangle}{\Omega^2}, \quad \eta_{p_i}^2 = \frac{\langle p_i \rangle}{\Omega^2}. \quad (7.24)$$

Algo interesante que se encuentra de este tipo de sistemas es que cada una de las fluctuaciones es contribución de su propia variable, una de las razones por la que el sistema tiene esta característica se debe a que el sistema es descrito puramente por una distribución de Poisson (para cada variable). Debido a que los promedios son proporcionales a las variables deterministas $\langle x_i \rangle = \Omega x_i$ entonces las fluctuaciones solo dependerán del inverso de Ω , entonces cuando el tamaño del sistema sea lo suficientemente grande las fluctuaciones se volverán muy pequeñas, lo que hace que en los sistemas deterministas no se tomen en cuenta, sin embargo, cuando el sistema es pequeño estas fluctuaciones son relevantes. Dado que hemos considerado que las fluctuaciones intrínsecas son todas aquellas debido a los elementos que conforman nuestro sistema, podemos decir que las fluctuaciones que hemos calculado son intrínsecas. Nos gustaría recalcar que las fluctuaciones que hemos calculado son válidas para cualquier tipo de redes que cumpla las condiciones que mencionamos. Las fluctuaciones en el estado estacionario nos darán una idea de que modelo es el más correcto, es decir, que tipo de función de Hill es la que mejor coincide con las fluctuaciones del estado estacionario, esto lo veremos más adelante con un ejemplo en particular.

El modelo que desarrollamos en esta sección podría decirse que es el caso más general cuando están conectamos más de dos módulos, incluso contiene al modelo que ya presentamos, podemos ahora analizar cualquier red genética aplicando estas ideas tan solo es necesario conocer los parámetros de nuestra red y la matriz de conexiones. Incluso es posible usar la función de Hill con correcciones estocásticas, por lo que esperamos encontrar resultados más cercanos al determinista

7.3.1. Elementos para simulación

Podemos realizar las simulaciones de nuestro sistema a través de dos métodos, el primero emplear el algoritmo de Gillespie (usando cualquier función de Hill excepto la que contiene correcciones estocásticas) y de ahí inferir los promedios y las fluctuaciones de las variables del sistema, lo otro es emplear a la aproximación determinista y al teorema de fluctuación disipación (TFD) (usando cualquier función de Hill, particularmente la que contiene correcciones estocásticas), a continuación, mostraremos cada uno de estos métodos.

Algoritmo de Gillespie

Con todo lo anterior ya tenemos el desarrollo teórico para poder escribir el programa que realice las simulaciones del sistema, de hecho ya tenemos una muy buena parte ya hecha en el capítulo anterior así como en el primer caso que hemos analizado en este capítulo, para hacer simulaciones con este algoritmo es necesario la matriz estequiométrica y el vector de propensión del sistema, para ello construiremos cada uno de estos, primero nos fijamos en las reacciones que suceden en cada uno de los nodos, para cada uno de los nodos se tendría la siguiente matriz estequiométrica y su correspondiente vector de propensión

$$S_n = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad \nu_n = \{\Omega k_{1n} H_g^n(p_l, \Omega), \gamma_{m_n} m_n, k_{3n} m_n, \gamma_{p_n} p_n\},$$

por lo que ahora el vector de propensión y la matriz estequiométrica de todo el sistema será

$$S = \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & S_n \end{pmatrix}, \quad \nu = \{\nu_1, \nu_2, \dots, \nu_n\}, \quad (7.25)$$

con estos elementos podemos implementar el algoritmo de Gillespie para redes genéticas, para saber cómo sería la dinámica de estos sistemas analizaremos uno de estos a continuación.

Aproximación determinista

La otra forma de modelar a las redes de regulación genética es a través de la aproximación determinista y el TFD, aunque se seguirá usando a S que aparece en (7.25), mientras que el vector de propensión se le tomará el límite determinista para encontrar las tasas de reacción, de tal manera que obtendremos algo similar a

$$t_n = \{k_{1n} H_g^n(\hat{p}_l, \Omega) \gamma_{m_n} \hat{m}_n, k_{3n} \hat{m}_n, \gamma_{p_n} \hat{p}_n\} \quad (7.26)$$

aquí las variables \hat{g}_n , \hat{m}_n y \hat{p}_n denotan las concentraciones deterministas, la función de Hill sería la determinista o la que considera correcciones estocásticas. En general usaremos los siguientes dos elementos para este tipo de redes

$$S = \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & S_n \end{pmatrix}, \quad t_+ = \{t_1, t_2, \dots, t_n\}. \quad (7.27)$$

A continuación, analizaremos un sistema en particular para ver qué resultados se obtienen.

7.3.2. Reloj activador-represor

Un buen ejemplo de una red genética donde podemos aplicar el modelo que hemos desarrollado es el reloj activador-represor, primero analizaremos el sistema usando el algoritmo de Gillespie usando la función de Hill determinista, posteriormente la aproximación determinista y el TFD, usando la función de Hill determinista y la que contiene correcciones estocásticas. Debido a que en

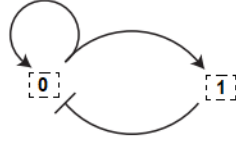


Figura 7.8: Reloj activador-represor.

este modelo uno de los módulos se va autoactivando y entonces participan dos tipos de proteínas sobre uno de los módulos, una representación gráfica de esto puede verse en la figura 7.8.

Con ayuda de la imagen anterior construimos la matriz de conexiones del sistema

$$A_{ij} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \quad (7.28)$$

sobre el módulo A una proteína sintetizada por el módulo B actúa como supresora al mismo tiempo la proteína sintetizada por el módulo A actúa como activadora, mientras que una proteína sintetizada por el módulo A actúa como activadora sobre el módulo B. Con ayuda de la ecuación maestra que hemos derivado en esta sección calcularemos la ecuación maestra del sistema, donde los elementos que necesitamos son los parámetros y la matriz de conexiones. Asumimos que $k_{1i} = k_1$, además $n_{11} = n_{12} = n_{21} = 2$ y $n_{22} = 0$, $\beta_{11} = \beta_{21} = 1$, $\beta_{12} = 0,0004$ por lo que la ecuación maestra será

$$\begin{aligned} \frac{\partial P(m_i, p_i, t)}{\partial t} = & \sum_{i=1}^2 \left(k_{1i} \Omega \left(H_g^i(p_i, \Omega) P(m_i - 1, p_i, t) - H_g^i(p_i, \Omega) P(m_i, p_i, t) \right) \right. \\ & + \frac{m_i + 1}{\tau_{2i}} P(m_i + 1, p_i, t) - \frac{m_i}{\tau_{2i}} P(m_i, p_i, t) \\ & + k_{3i} m_i P(m_i, p_i - 1, t) - k_{3i} m_i P(m_i, p_i, t) \\ & \left. + \frac{p_i + 1}{\tau_{3i}} P(m_i, p_i + 1, t) - \frac{p_i}{\tau_{3i}} P(m_i, p_i, t) \right). \end{aligned} \quad (7.29)$$

Siendo las funciones de Hill de la siguiente manera

$$H_g^1(p_1, p_2, \Omega) = \frac{\beta_1 \left(\frac{p_1}{K_1 \Omega} \right)^{n_{11}} + \alpha_0}{1 + \left(\frac{p_1}{K_1 \Omega} \right)^{n_{11}} + \left(\frac{p_2}{K_2 \Omega} \right)^{n_{12}}}, \quad H_g^2(p_1, \Omega) = \frac{\beta_2 \left(\frac{p_1}{K_1 \Omega} \right)^{n_{21}}}{1 + \left(\frac{p_1}{K_1 \Omega} \right)^{n_{21}}}. \quad (7.30)$$

Estas funciones las encontramos de acuerdo con (7.17). Vamos a usar también la que contiene correcciones estocásticas, aunque omitiremos escribirlas. Lo que falta ahora es simular este sistema, empezaremos por usar el algoritmo de Gillespie y usaremos una función de Hill determinista, para esto volvemos a escribir un programa en Python el cual se es muy similar al empleado en el caso del represilador. Al hacer la simulación estocástica de un solo proceso obtenemos lo que se observa en la figura 7.9, podemos observar claramente la presencia de fluctuaciones en el sistema, debido a que hemos usado un tamaño de sistema relativamente pequeño estas fluctuaciones son más marcadas, además se observa que la concentración de las especies bioquímicas oscila.

También analizaremos y cuantificamos las fluctuaciones de este sistema usando nuevamente herramientas computacionales y el algoritmo de Gillespie, de donde haremos inferencias estadísticas. Lo que se obtiene de la simulación está en 7.10, la línea verde muestra el promedio de las variables sobre el ensamble, mientras que las líneas punteadas nos dicen en que intervalo se encontraran la

mayoría de las fluctuaciones del sistema. Las fluctuaciones de estas cantidades se mantienen estables, aunque como vemos generan una banda muy amplia, lo que está completamente en contra de lo que se obtendría en el estado estacionario, ya que las fluctuaciones dependerían de $\frac{\sqrt{x}}{\Omega}$, además, la dinámica de las concentraciones parece tender a un valor estacionario, cuando se espera que el sistema presente oscilaciones.

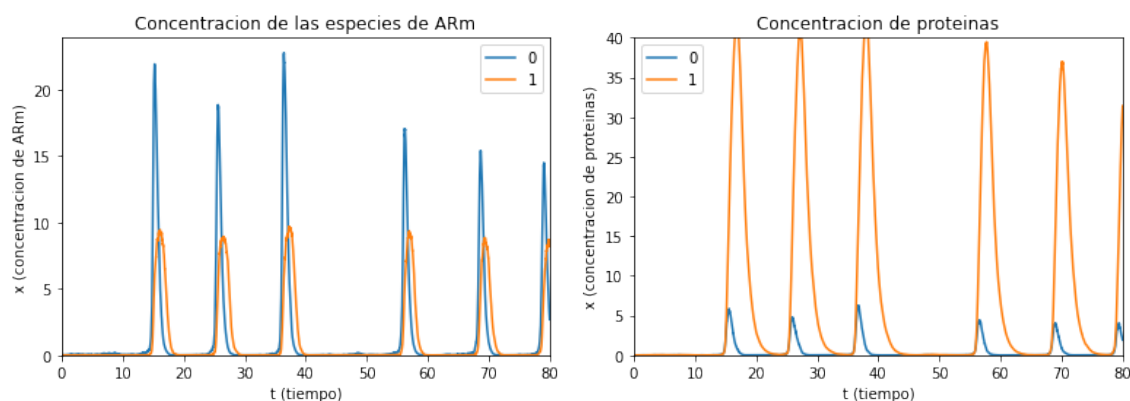


Figura 7.9: Reloj activador-represor. Simulación estocástica usando al algoritmo de Gillespie y la función de Hill determinista, en la primera figura se observa la concentración de ARm, en la segunda la concentración de las proteínas. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$, $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{21} = 3$.

Ahora analizaremos el mismo sistema, pero ahora usando las ecuaciones diferenciales que se derivan de la aproximación determinista y el teorema de fluctuación disipación (TFD), construimos su correspondiente programa en Python (se puede revisar la carpeta de Drive para ver el archivo con nombre *arelojar.ipynb*). Primero empleamos a la función de Hill determinista, los resultados que obtenemos se pueden ver en la figura 7.11 a). Lo primero que nos gustaría remarcar es que al usar el TFD el tiempo de computación se reduce dramáticamente. En esta figura las líneas verdes corresponden a las concentraciones deterministas, mientras que las líneas rojas punteadas crean una banda en donde se podría encontrar las fluctuaciones del sistema. En este caso las concentraciones deterministas y las fluctuaciones difieren mucho a comparación de lo que se obtuvo a través del algoritmo de Gillespie, en este caso las fluctuaciones son más pequeñas. Tenemos una banda de fluctuaciones más grandes asociadas al ARm de la proteína P_1 , una posible razón se debe a que tanto las fluctuaciones de P_1 y P_2 afectan a m_1 . Aunque la dinámica de nuestro sistema ya presenta oscilaciones más marcadas aún está lejos de lo que podría obtenerse realmente si se considera el tamaño de las fluctuaciones en el estado estacionario, por lo que a continuación volveremos a simular el sistema solo que ahora emplearemos una función de Hill que incluya correcciones estocásticas.

Usando nuevamente a la aproximación determinista y al TFD, pero ahora con la función de Hill con correcciones estocásticas, simulamos el sistema, lo que obtenemos se puede observar en la figura 7.11 b). Lo primero que observamos es que ahora el tamaño de las fluctuaciones se ha reducido considerablemente con respecto a las figuras 7.10 y 7.11 b), una de las razones se debe al tipo de función de Hill que hemos usado, en esta última hemos usado a la función de Hill con correcciones estocásticas, por lo que al momento de usar el TFD se han omitido algunos términos. De hecho, este último resultado que hemos obtenido se acerca más a lo obtenido con las ecuaciones deterministas. Como hemos venido mencionando el usar la función de Hill con correcciones estocásticas es la más correcta si se quiere considerar un sistema estocástico, esto también se puede ver porque en el estado estacionario o cerca del estado estacionario las fluctuaciones deberían de depender como $\sqrt{\frac{x}{\Omega}}$, y el modelo que más se acerca a esto es en el que

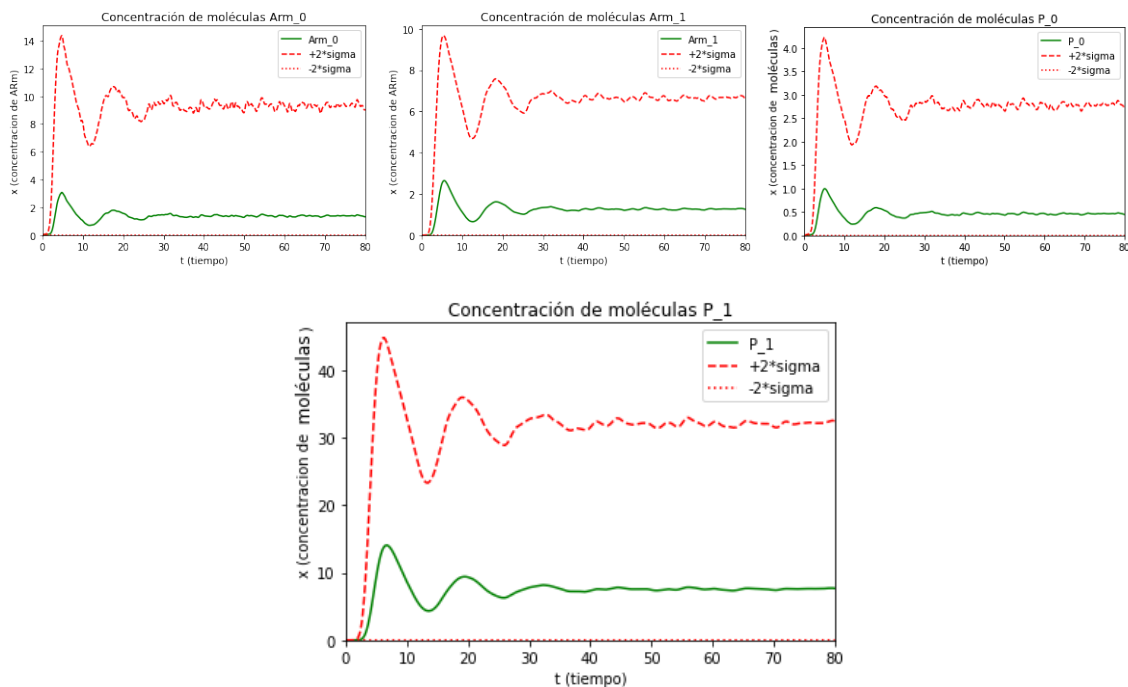


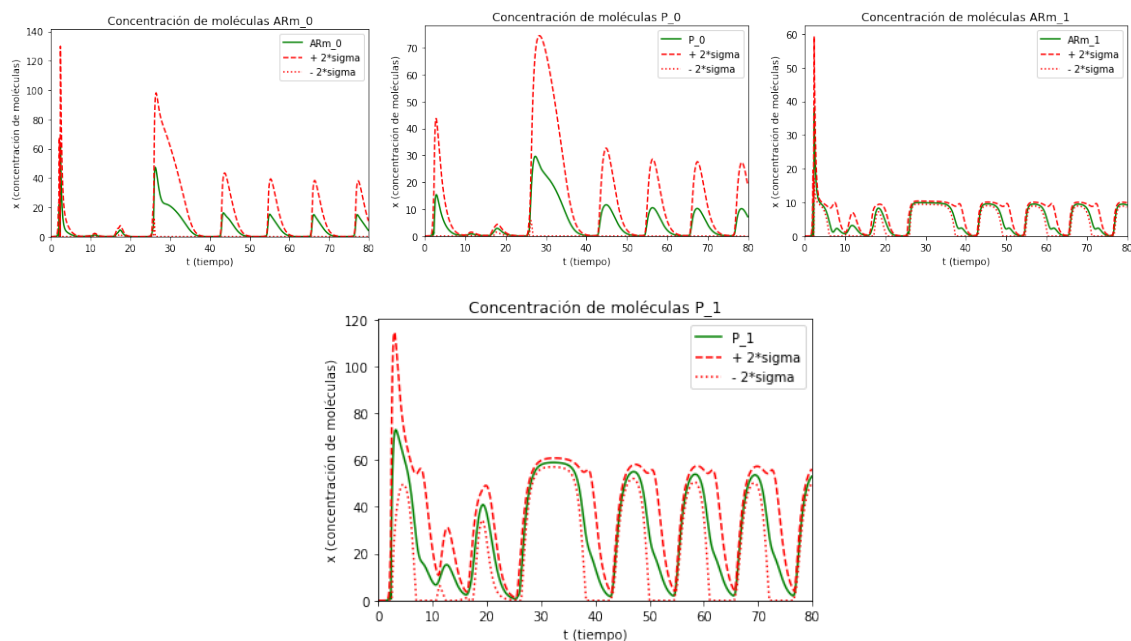
Figura 7.10: Reloj activador-represor. Dinámica de las concentraciones deterministas y las fluctuaciones a partir de simulaciones estocásticas con tamaño de ensamble= 10000, en las figuras de la parte superior se observa la dinámica de los ARm's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$, $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{31} = 3$.

se usó la función de Hill con correcciones estocásticas.

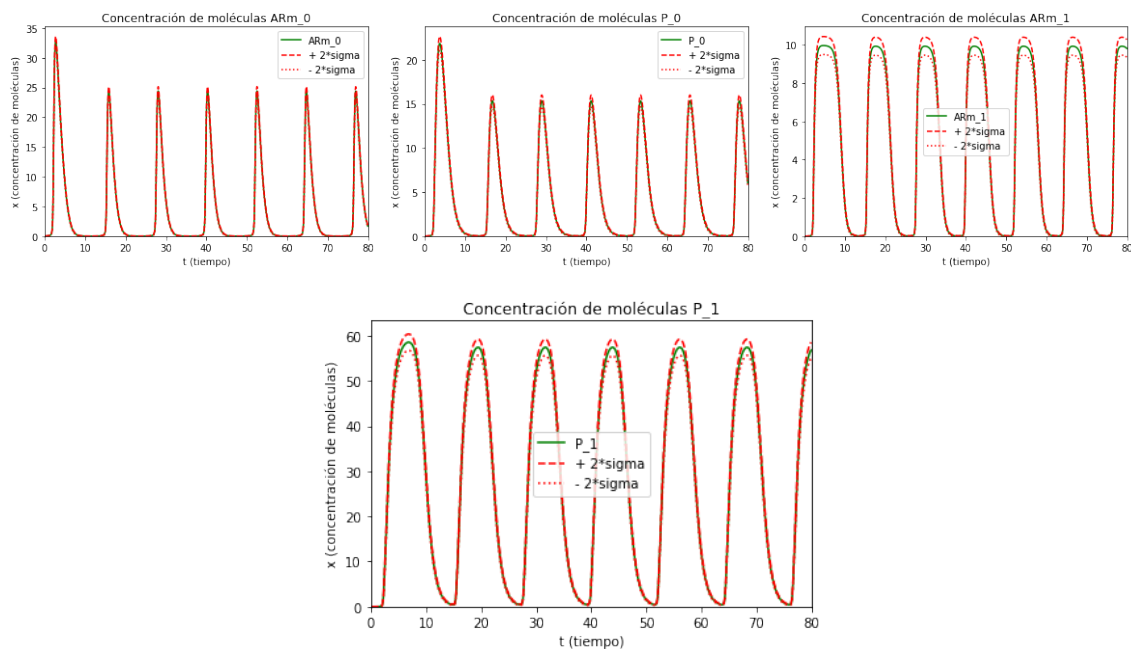
Con estos ejemplos nos damos cuenta que las herramientas y desarrollos que hemos hecho hasta este punto nos permiten modelar la dinámica de los circuitos genéticos estocásticos, promediar sus concentraciones sobre un ensamble si se usa el algoritmo de Gillespie, además de cuantificar las fluctuaciones de las especies bioquímicas involucradas, esto bajo dos formalismos, el algoritmo de Gillespie y el TFD, más aún hemos dado una banda alrededor del promedio en cual se encontrarán la mayoría de las fluctuaciones, esta cuantificación se realiza durante toda la evolución dinámica del sistema. Incluso tenemos la libertad de elegir el tipo de función de Hill ya que hemos hecho una derivación exacta de su aparición como tasa de propensión en los sistemas estocásticos, siendo la función de Hill con correcciones estocásticas las que nos da resultados más exactos, ya que si se usa la determinista existe una sobreestimación de la concentración de las moléculas, así como el tamaño de las fluctuaciones (esto se recalca más en el reloj activador-represor). Las fluctuaciones de nuestros sistemas tienen una banda perceptible debido a que nuestros sistemas son pequeños, para sistemas más grandes esta cantidad se irá atenuando. Con todo el formalismo, los métodos y las ecuaciones desarrolladas en este capítulo nos permiten estudiar redes cuya estructura sea más compleja y/o contengan un número mayor de módulos, aunque esto lo dejamos para trabajos futuros.

Redes estocásticas de regulación genética

7.3 Red estocástica de regulación genética general



(a) Función de Hill determinista.



(b) Función de Hill con correcciones estocásticas.

Figura 7.11: Reloj activador-represor. Se ha modela la dinámica y las fluctuaciones del sistema usando la aproximación determinista y el TFD, en la figura (a) se ha usado la función de Hill determinista, en (b) la función de Hill con correcciones estocásticas. Tanto para (a) y (b) las figuras de la parte superior se observa la dinámica de los ARm's, en las de abajo la de las proteínas. Las líneas verdes representan las concentraciones deterministas, las líneas rojas punteadas la banda en donde se encuentran las fluctuaciones. Se han usado los siguientes parámetros: $n_{11} = n_{12} = n_{21} = 2$, $\Omega = 180$, $k_{11} = 250$, $k_{12} = 30$, $\alpha_0 = 0,0004$ $\tau_{2i} = \frac{1}{3}$, $\tau_{31} = \frac{1}{3}$, $\tau_{32} = 1$, $k_{21} = 1$, $k_{21} = 3$.

Capítulo 8

Conclusiones y resultados

Revisamos las redes de regulación genética y la retroactividad, nuestro análisis fue orientado a cuantificar las fluctuaciones del sistema usando un modelo estocástico, para ello hemos desarrollado algunos métodos, obteniendo resultados importantes.

En el capítulo 3 de análisis estocástico hemos propuesto un algoritmo para encontrar la forma de la distribución estacionaria para un sistema de vida y muerte multivariable, ya que estamos particularmente interesados en las distribuciones estacionarias (por la función de Hill), este algoritmo es válido cuando el sistema es al menos débilmente reversible y tiene deficiencia cero, este método es una alternativa a los escasos que actualmente se muestran en la literatura. También hemos usado una aproximación a segundo alrededor del promedio para derivar las ecuaciones deterministas de los procesos de vida y muerte multivariable, obteniendo las ecuaciones diferenciales derivadas de la ley de acción de masas deterministas, pero ahora con términos que se deben a correcciones estocásticas, estas correcciones se desprecian cuando el sistema se vuelve lo suficientemente grande. Además, de esta aproximación hemos derivado una definición de fluctuación, la definición que hemos propuesto depende únicamente de la varianza y el tamaño del sistema, también derivamos un conjunto de ecuaciones que serían equivalentes al teorema de fluctuación disipación (TFD) (4.20), aunque en nuestras ecuaciones aparece un término correctivo (comparado con el TFD usual) que desaparece cuando el sistema es lineal. La derivación que hemos hecho muestra enfoque alternativo para analizar los sistemas estocásticos y permite una mejor comparación con su contra parte determinista, ya que con estas ecuaciones es posible analizar la dinámica y las fluctuaciones del sistema de manera simultánea.

Con las herramientas desarrolladas analizamos la dinámica y las fluctuaciones de las redes químicas y/o bioquímicas, particularmente a la retroactividad en los sistemas estocásticos. En la segunda mitad del capítulo 2 están algunas derivaciones deterministas y se encuentra una expresión para la retroactividad que es válida para cualquier red de reacciones químicas. En aras de cuantificar la retroactividad en los sistemas estocásticos en la parte final del capítulo 4 mostramos una expresión para la retroactividad que incluye correcciones estocásticas (4.41), para derivar la retroactividad se considera que algunas reacciones se encuentran equilibrio, entonces las contribuciones estocásticas al sistema serán las mismas que cuando el sistema se encuentre aislado. Para ilustrar el efecto de la retroactividad analizamos el proceso de fosforilación/desfosforilación, en donde hemos encontrado cómo se comporta el sistema a nivel estocástico, aún más, hemos cuantificado las fluctuaciones del sistema a través de la aproximación determinista y al TFD, entre los resultados que hemos obtenido se encuentra que las fluctuaciones del sistema que tiene más retroactividad son mayores en comparación con el sistema que tiene menos retroactividad, una posible razón de porque pasa esto podría deberse al hecho que la interacción entre los dos módulos es mayor, por consiguiente las fluctuaciones entre estas se amplifica, este resultado se observa

durante toda la dinámica. También hemos calculado la distribución estacionaria del sistema, entonces podemos determinar de manera exacta las fluctuaciones en el estado estacionario, de esto se puede notar que cuando más retroactividad haya más fluctuaciones habrá.

En el capítulo 6 hemos derivado la versión estocástica de la función de Hill. Primero hemos construido una ecuación maestra a partir de la cual hemos derivado las ecuaciones de Hill usuales (6.24), debido a que tenemos la ecuación maestra de nuestro sistema podemos calcular de manera exacta las fluctuaciones de esta función (6.28), es importante mencionar que hasta donde sabemos, no se habían calculado las fluctuaciones. Además, estos cálculos solo son válidos cuando aparece solo el proceso de Hill (las reacciones de donde se deriva la función de Hill). Es indispensable conocer el cómo aparecen las funciones de Hill en procesos más complejos o en la que participen más reacciones, entonces hemos derivado algunas variantes de la función de Hill a la cual hemos nombrado semi-determinista, estocástica y semi-estocástica, estas derivaciones se hacen en aras de explicar la manera de como se introduce la función de Hill determinista en los sistemas estocásticos. Finalmente mostramos como aparece esta función dentro de un proceso estocástico a partir de consideraciones más sistemáticas, notar que esta última derivación es la más exacta, esta derivación tiene sus limitantes al hacer simulaciones de Gillespie, caso contrario cuando se toma el límite determinista ya que nos permite hacer una derivación más limpia de la función de Hill. De esto último obtenemos una nueva función de Hill con correcciones estocásticas, cuando el tamaño del sistema se vuelve grande se recupera la versión usual, esta nueva función de Hill podrá ayudar a dar una descripción más precisa de la dinámica celular, principalmente en las redes de regulación genética. Esta función que hemos derivado permitirá cuantificar mejor la dinámica de las concentraciones y cuantificar mejor las fluctuaciones de los sistemas bioquímicos que involucren la función de Hill. Con ayuda de [37] describimos una manera de obtener una ecuación de Hill más realista considerando que existen estados intermedios entre la unión de una enzima y una proteína, y además fluctuaciones, por lo que obtenemos las funciones de Hill con correcciones estocásticas, aunque solo hemos considerado el caso en el que existen procesos intermedios a través de la unión secuencial e independiente. Estas funciones nos permitirían estudiar sistemas más pequeños y cuantificar de manera más exacta las fluctuaciones.

En el capítulo 7 hemos construido la ecuación maestra de una red de regulación genética, debido a que ya hemos hecho una derivación sistemática de la función de Hill, tenemos la libertad de usar alguna de la que hemos encontrado, como nos interesa la dinámica y las fluctuaciones del sistema realizamos un análisis con el algoritmo de Gillespie (usando la función de Hill determinista), la aproximación determinista y el TFD (usando la función de Hill determinista y la que contiene correcciones estocásticas). Primero nos enfocamos en un tipo de redes de regulación en la cual una sola proteína actúa como represora o activadora sobre otro módulo. Debido a que tenemos la ecuación maestra se puede calcular la distribución de estos sistemas en el estado estacionario, para ello usamos el algoritmo que hemos desarrollado y obtenemos una distribución que es la multiplicación de varias distribuciones de Poisson, con esta distribución calculamos los promedios deterministas y las fluctuaciones, las fluctuaciones en el estado estacionario nos permitirán determinar que modelo es más exacto. Nos enfocamos en analizar el represilador, realizamos tres diferentes simulaciones, primero usamos al algoritmo de Gillespie, después a la aproximación determinista y al TFD usando la función de Hill usual y la que tiene correcciones estocásticas. En la simulación que usa la función de Hill con correcciones estocásticas es la que presenta resultados más satisfactorios ya que esperamos que cerca de los puntos estacionarios las fluctuaciones dependan como $\frac{\sqrt{\hat{x}}}{\Omega}$ (\hat{x} son las concentraciones deterministas) y es el sistema con la función de Hill con correcciones deterministas la que lo hace. También analizamos una red estocástica de regulación genética en donde ahora varias proteínas actúan como factores de transcripción (represoras y activadoras), la única diferencia que tenemos con el primer sistema que analizamos es la forma de la función de Hill, su ecuación maestra, los promedios y fluctuación tienen una estructura muy similar al primer caso analizado. Para este tipo de sistemas nos

enfocamos en analizar el reloj activador-represor, primero usamos el algoritmo de Gillespie para inferir la dinámica del promedio de las concentraciones de las proteínas y el ARm y las fluctuaciones del sistema, los resultados que obtenemos se alejan completamente de lo esperado en la contra parte determinista, modificando radicalmente la dinámica del sistema y la magnitud de las fluctuaciones. Posteriormente usamos la aproximación determinista y el TFD, empezando con una función de Hill determinista, en este caso obtuvimos que la dinámica de las concentraciones ya presenta oscilaciones, aunque las fluctuaciones aun eran grandes. Finalmente, con la función de Hill con correcciones estocásticas, los resultados que encontramos muestran que el sistema sigue una dinámica oscilatoria y la banda de fluctuaciones es pequeña, lo que va de acuerdo con lo que se obtiene en las fluctuaciones del estado estacionario. De estos resultados podemos decir que el introducir la función de Hill determinista en un sistema estocástico sin una correcta justificación impide cuantificar adecuadamente las fluctuaciones, ya que existe una sobreestimación de estas, como se puede observar de las simulaciones del reloj activador-represor.

Algunos temas para desarrollar en el futuro a partir de este trabajo sería analizar diferentes redes de regulación genética en particular e incorporar fluctuaciones extrínsecas, podemos analizar algunas redes de regulación de algunas plantas o enfermedades, permitiendo diseñar fármacos y terapias que consideren la presencia de fluctuaciones intrínsecas y extrínsecas. También podemos emplear a la aproximación determinista y la variante del TFD para modelar la dinámica y cuantificar las fluctuaciones de cualquier red química, siempre y cuando sigan a la ley de acción de masas, incluso podemos llegar a hacer uso del tercer momento central.

Apéndice A

Teorema de fluctuación disipación

Aquí derivaremos al teorema de fluctuación-disipación de la manera en la que se acostumbra a hacer en la literatura, para ello nos apoyaremos de [18] usando la aproximación de Van Kampen. Recordar que estamos analizando una red de procesos químicos, por lo que se tendrían que usar la ecuación maestra que se obtuvo en los procesos de vida y muerte multivariable

$$\begin{aligned} \partial_t P(S_j, t) = \Omega \sum_i & (t_i^-(S_j + \Gamma_{ji})P(S_j + \Gamma_{ji}, t) - t_i^+(S_j)P(S_j, t) \\ & + t_i^+(S_j - \Gamma_{ji})P(S_j - \Gamma_{ji}, t) - t_i^-(S_j)P(S_j, t)). \end{aligned} \quad (\text{A.1})$$

(Para más detalles de la derivación de esta ecuación y las variables involucradas revisar el capítulo de análisis estocástico). Podemos reescribir a esta ecuación con ayuda de unos operadores \mathbf{E} , estos operadores actúan de la siguiente manera

$$\mathbf{E}^a f(x) = f(x + a), \quad (\text{A.2})$$

nuestra ecuación maestra se reescribirá como

$$\partial_t P(S_j, t) = \Omega \sum_i \left(\left[\prod_j \mathbf{E}^{-\Gamma_{ji}} - 1 \right] t_i^+(S_j)P(S_j, t) + \left[\prod_j \mathbf{E}^{\Gamma_{ji}} - 1 \right] t_i^-(S_j)P(S_j, t) \right). \quad (\text{A.3})$$

Ahora realizaremos una expansión de Van Kampen, o también conocida como aproximación lineal de ruido (LNA) [24], se supone que el sistema es lo suficientemente grande por lo que el número de partículas del sistema se puede aproximar como

$$S_j = \Omega s_j + \sqrt{\Omega} \alpha_j, \quad (\text{A.4})$$

donde s_j corresponde a la concentración determinista, mientras que α_j corresponde a las fluctuaciones del sistema, se elige de esta manera porque esto va de acuerdo con los resultados experimentales. Ahora aplicamos esta consideración sobre cada uno de los elementos de la ecuación maestra.

- 1.- Suponemos que podemos escribir a la distribución de probabilidades en términos de las fluctuaciones, es decir

$$P(S_j, t) = P(\alpha_j, t), \quad (\text{A.5})$$

por lo que entonces la derivada temporal que tenemos en (A.3) se transforma usando la regla de la cadena,

$$\partial_t P(S_j, t) = \partial_t P(\alpha_j, t) - \Omega^{\frac{1}{2}} \dot{s}_j \partial_t P(\alpha_j, t) = \partial_t P(\alpha_j, t) (1 - \Omega^{\frac{1}{2}} \dot{s}_j). \quad (\text{A.6})$$

2.- Las tasas de propensión se pueden aproximar como

$$t_i^\pm(S_j) \approx R_i^{D^\pm}(s_j) + \sum_l \Omega^{-\frac{1}{2}} \alpha_l \frac{\partial}{\partial s_l} R_i^{D^\pm}(s_j), \quad (\text{A.7})$$

siendo $R_j^{D^\pm}(s_j)$ las tasas deterministas.

3.- Nos fijamos que los operadores $\mathbf{E}^{\pm k}$ se pueden aproximar como

$$\mathbf{E}^{\pm k} \approx \left[1 \pm \sum_l \frac{k}{\sqrt{\Omega}} \frac{\partial}{\partial \alpha_l} + \frac{k^2}{2\Omega} \sum_l \sum_s \frac{\partial^2}{\partial \alpha_l \partial \alpha_s} \right]. \quad (\text{A.8})$$

Sustituyendo a (A.5,A.6,A.7, A.8) en (A.3), obtenemos

$$\begin{aligned} \partial_t P(\alpha_j, t) (1 - \Omega^{\frac{1}{2}} \dot{s}_l) = & \\ & \sum_i \left(\left[\sum_l -\Gamma_{li} \sqrt{\Omega} \frac{\partial}{\partial \alpha_l} + \frac{\Gamma_{li} \Gamma_{si}}{2} \sum_l \sum_s \frac{\partial^2}{\partial \alpha_l \partial \alpha_s} \right] \left(R_i^{D^+}(s_j) + \sum_l \Omega^{-\frac{1}{2}} \alpha_l \frac{\partial}{\partial s_l} R_i^{D^+}(s_j) \right) P(\alpha_j, t) \right. \\ & \left. + \left[\sum_l \Gamma_{li} \sqrt{\Omega} \frac{\partial}{\partial \alpha_l} + \frac{\Gamma_{li} \Gamma_{si}}{2} \sum_l \sum_s \frac{\partial^2}{\partial \alpha_l \partial \alpha_s} \right] \left(R_i^{D^-}(s_j) + \sum_l \Omega^{-\frac{1}{2}} \alpha_l \frac{\partial}{\partial s_l} R_i^{D^-}(s_j) \right) P(\alpha_j, t) \right), \end{aligned} \quad (\text{A.9})$$

esta relación se debe de cumplir para cualquier Ω que sea lo suficientemente grande, por lo que los términos se tienen que igualar uno a uno, primero nos fijamos en aquellos que dependen de $\Omega^{\frac{1}{2}}$ y obtenemos

$$\dot{s}_l = \sum_l \Gamma_{li} (R_i^{D^+}(s_j) - R_i^{D^-}(s_j)), \quad (\text{A.10})$$

estas son justo las ecuaciones deterministas. Ahora nos fijamos en los otros términos que no dependen de Ω y despreciamos a los que dependen de $\Omega^{-\frac{1}{2}}$, obteniendo finalmente

$$\begin{aligned} \partial_t P(\alpha_j, t) = & \sum_i \left(\sum_l -A_{il} \frac{\partial}{\partial \alpha_i} [\alpha_l P(\alpha_j, t)] + \frac{1}{2} \sum_l B_{il} \frac{\partial^2}{\partial \alpha_i \partial \alpha_l} [P(\alpha_j, t)] \right), \\ A_{ij}(S_j) = & \sum_i \Gamma_{ji} \frac{\partial}{\partial s_j} [t_i^+(S_j) - t_i^-(S_j)], \\ B_{jl}(S_j) = & \sum_i \Gamma_{ji} \Gamma_{li} [t_i^+(S_j) + t_i^-(S_j)], \end{aligned} \quad (\text{A.11})$$

esta expresión es una distribución de probabilidad para las fluctuaciones del sistema, también es ecuación diferencial parcial de tipo Fokker-Planck, en algunos casos puede resolverse de forma analítica y es mucho más sencillo de obtener su solución numérica [18]. Debido a que las fluctuaciones no se pueden calcular directamente del sistema, se tiene que usar sus covarianzas o varianzas para determinarla, y como ya tenemos una distribución de probabilidad la usaremos. Hay que recordar que la covarianza está definida como

$$\sigma^2(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle, \quad (\text{A.12})$$

entonces con ayuda de la primera expresión de (A.11) calculamos estos términos, en este caso los promedios serían sobre todos los posibles valores de α_j , es decir haremos uso de integrales, por lo que, integrando por partes, finalmente llegaremos a

$$\frac{\partial}{\partial t} \sigma^2 = \mathbf{A} \sigma^2 + \sigma^2 \mathbf{A}^T + \mathbf{B}, \quad (\text{A.13})$$

\mathbf{A} y \mathbf{B} denotan a las matrices que aparecen en (A.11), debido a que esta relación es sobre las fluctuaciones, se le conoce como el Teorema de Fluctuación-Disipación [24].

Apéndice B

Función de Hill general

A continuación, derivaremos una función de Hill que se usara en redes de regulación genética cuando en un módulo de transcripción/traducción participen más de una proteína como factor de transcripción, para eso recordemos que el proceso de transcripción es como en la figura siguiente.

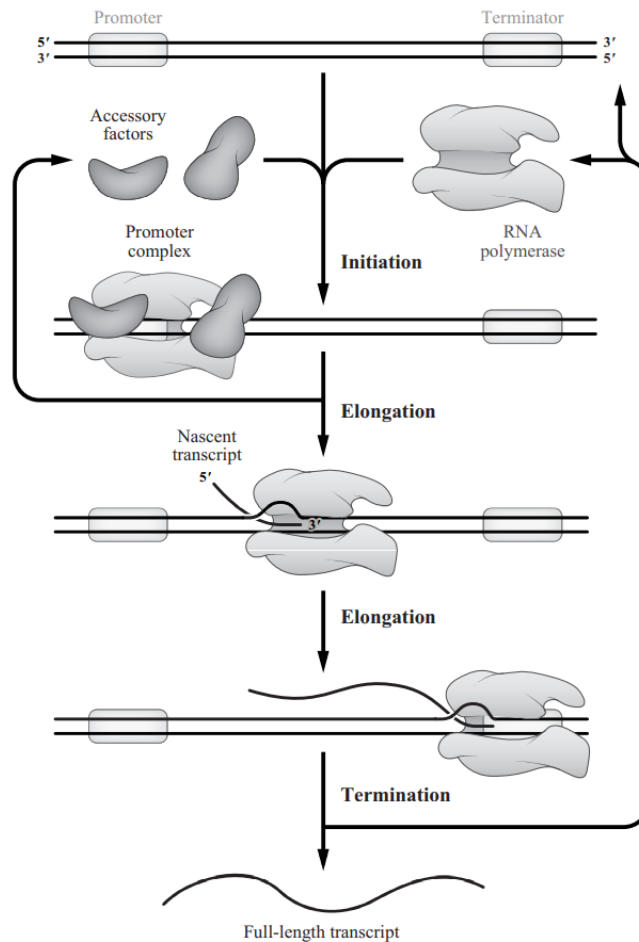


Figura B.1: Proceso de transcripción [15].

Pero ahora la región promotora puede ser afectado por unos factores de transcripción similar a como se observa en la siguiente figura.

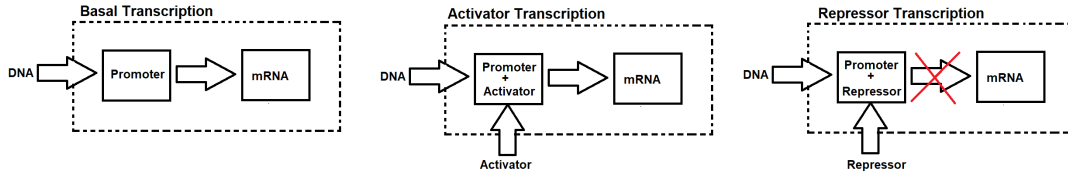
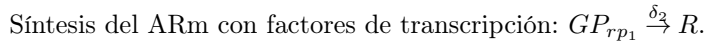
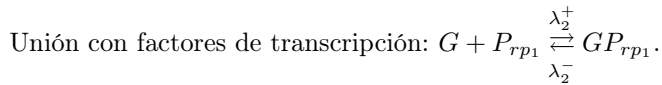
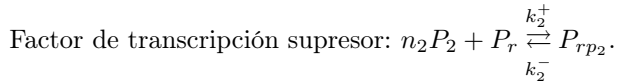
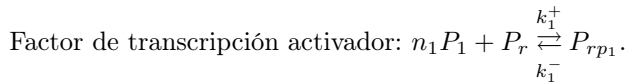
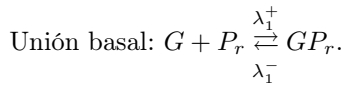


Figura B.2: Regulación genética. En la primera figura se observa la síntesis del ARm sin que participen factores de transcripción. En la segunda figura participan factores de transcripción de activación que propician la síntesis del ARm. En la tercera figura participan factores de transcripción de supresión que inhiben la síntesis del ARm.

A partir de estas dos imágenes derivamos un conjunto de reacciones bioquímicas que nos describen la actividad de la promotora hasta sintetizar a un ARm. Para esto suponemos que dos tipos de proteínas P_1 y P_2 actúan como factores de transcripción como activadoras y supresoras respectivamente.



G representa al gen, P_r a la promotora, R al ARm, GP_r a la unión del gen a la promotora, P_{rp_1} y P_{rp_2} cuando el factor de transcripción se ha unido a la promotora. La primera reacción que tenemos en la parte de arriba es un proceso reversible que nos dice que el gen se une a la promotora estén o no los factores de transcripción. La segunda reacción es también un proceso reversible en el que los factores de transcripción se unen a la promotora de tal manera que crean una nueva molécula que hace que la síntesis del ARm sea más rápida. En la tercera reacción el factor de transcripción crea una nueva molécula de tal manera que esta no se pueda unir al gen. En la cuarta reacción el gen G se une a la molécula P_{rp_1} . Finalmente, las dos últimas reacciones nos dicen como se va sintetizando el ARm a través de un proceso de una sola dirección.

Usamos la ley de acción de masas para construir un conjunto de ecuaciones diferenciales que describan las concentraciones de nuestro sistema, para ello g representa la concentración del gen, p_r a concentración la promotora, r a la concentración ARm, s_1 a la concentración de GP_r , s_2 y s_3 a la concentración de P_{rp_1} y P_{rp_2} respectivamente, s_4 a la concentración de GP_{rp_1} . Las ecuaciones diferenciales que obtenemos son las siguientes,

$$\begin{aligned}
 \frac{ds_1}{dt} &= \lambda_1^+ gp_r - \lambda_1^- s_1, \\
 \frac{ds_2}{dt} &= k_1^+ p_1^{n_1} p_r - k_1^- s_2 - \lambda_2^+ gs_2 + \lambda_1^- s_4, \\
 \frac{ds_3}{dt} &= k_2^+ p_2^{n_2} p_r - k_2^- s_3, \\
 \frac{ds_4}{dt} &= \lambda_2^+ gs_2 - \lambda_1^- s_4, \\
 \frac{dp_r}{dt} &= -\lambda_1^+ gp_r + \lambda_1^- - k_1^+ p_1^{n_1} + k_1^- p_r - k_2^+ p_2^{n_2} + k_2^- p_r, \\
 \frac{dr}{dt} &= \delta_1 s_1 + \delta_2 s_4,
 \end{aligned} \tag{B.1}$$

a partir de estas ecuaciones diferenciales podemos ver que tenemos una cantidad conservada, la concentración inicial de promotoras es constante $p_0 = p_r + s_1 + s_2 + s_3 + s_4$. De las reacciones que tenemos consideramos que las 4 primeras se encuentran prácticamente en equilibrio, por lo que las primeras cuatro ecuaciones diferenciales se consideran en equilibrio, de donde obtenemos las siguientes condiciones

$$\begin{aligned}
 s_1 &= \frac{\lambda_1^+}{\lambda_1^-} gp_r = \lambda_1 g p_r, \\
 s_2 &= \frac{k_1^+}{k_1^-} p_1^{n_1} p_r = \frac{p_1^{n_1}}{K_1^{n_1}} p_r, \\
 s_3 &= \frac{k_2^+}{k_2^-} p_2^{n_2} p_r = \frac{p_2^{n_2}}{K_2^{n_2}} p_r, \\
 s_4 &= \frac{\lambda_2^+}{\lambda_2^-} g s_2 = \lambda_2 g s_2.
 \end{aligned} \tag{B.2}$$

Nuestro objetivo es encontrar una relación que describa a s_1 y s_4 en términos de los factores de transcripción y la síntesis basal, ya que estos están estrechamente relacionados con la síntesis del ARm como se puede ver en (B.1), entonces con ayuda de las relaciones anteriores podemos escribir lo siguiente

$$\begin{aligned}
 H_1 &= \frac{s_1}{p_0} = \frac{s_1}{p_r + s_1 + s_2 + s_3 + s_4}, \\
 H_2 &= \frac{s_4}{p_0} = \frac{s_4}{p_r + s_1 + s_2 + s_3 + s_4},
 \end{aligned} \tag{B.3}$$

al evaluar las cantidades obtenidas en (B.2) en estas relaciones obtenemos

$$\begin{aligned}
 H_1 &= \frac{\lambda_1 g}{1 + \lambda_1 g + \frac{p_1^{n_1}}{K_1^{n_1}} + \frac{p_2^{n_2}}{K_2^{n_2}} + \lambda_2 g \frac{p_1^{n_1}}{K_1^{n_1}}}, \\
 H_2 &= \frac{\lambda_2 g \frac{p_1^{n_1}}{K_1^{n_1}}}{1 + \lambda_1 g + \frac{p_1^{n_1}}{K_1^{n_1}} + \frac{p_2^{n_2}}{K_2^{n_2}} + \lambda_2 g \frac{p_1^{n_1}}{K_1^{n_1}}}.
 \end{aligned} \tag{B.4}$$

Recordar que la síntesis del ARm está dado por el termino $\delta_1 s_1 + \delta_2 s_4$, por lo que ahora tendremos

$$\delta_1 s_1 + \delta_2 s_4 = \delta_1 p_0 \frac{s_1}{p_0} + \delta_2 p_0 \frac{s_4}{p_0} = \delta_1 p_0 H_1 + \delta_2 p_0 H_2 = H, \tag{B.5}$$

donde

$$H = \frac{\delta_1 p_0 \lambda_1 g + \delta_2 p_0 \lambda_2 g \frac{p_1^{n_1}}{K_1^{n_1}}}{1 + \lambda_1 g + \frac{p_1^{n_1}}{K_1^{n_1}} + \frac{p_2^{n_2}}{K_2^{n_2}} + \lambda_2 g \frac{p_1^{n_1}}{K_1^{n_1}}} = \frac{\beta_0 K_0 + \beta_{11} \frac{p_1^{n_1}}{K_1^{n_1}}}{1 + K_0 + \beta_{21} \frac{p_1^{n_1}}{K_1^{n_1}} + \beta_{22} \frac{p_2^{n_2}}{K_2^{n_2}}}, \quad (\text{B.6})$$

($\beta_0 = \delta_1 p_0$, $K_0 = \lambda_1 g$, $\beta_{11} = \delta_2 p_0 \lambda_2 g$, $\beta_{21} = (1 + \lambda_2 g)$, $\beta_{22} = 1$) esta función de H que hemos definido correspondería a la función de Hill de nuestro sistema, nos dirá como se va sintetizando el ARm en términos de los factores de transcripción P_1 y P_2 , además de una constante basal K_0 . A β_0 se le conoce como la velocidad de síntesis basal. Ahora podemos dar una generalización de esta expresión, para ello suponemos que tenemos un sistema con m proteínas que actúa tanto como activadoras o supresora, entonces nuestra función de Hill generalizada será de la forma

$$H_g = \frac{\beta_0 K_0 + \sum_j \beta_{1j} \left(\frac{p_j^{n_j}}{K_j^{n_j}} \right)^q}{1 + K_0 + \sum_j \beta_{2j} \frac{p_j^{n_j}}{K_j^{n_j}}}, \quad (\text{B.7})$$

hemos agregado una q en el numerador, debido a que si la proteína actúa como activadora $q = 1$, mientras que si actúa como supresora $q = 0$. De esta manera hemos realizado una generalización de la función de Hill y que será muy útil para analizar cualquier módulo de transcripción/traducción en el que participen muchas proteínas. Podemos incluso introducir correcciones estocásticas a esta función de Hill, solo será necesario cambiar a $p_j^{n_j}$ de la siguiente manera

$$p_j^{n_j} \rightarrow p_j^{n_j} + \sigma^2 (p_j, p_j) n_j (n_j - 1) p_j^{n_j - 2}, \quad (\text{B.8})$$

este cambio se hace de acuerdo con la forma en la que se obtuvo la función de Hill con correcciones estocásticas.

Apéndice C

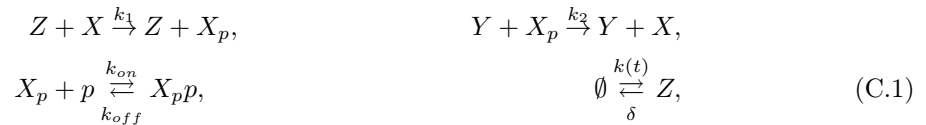
Cálculo de la deficiencia

A continuación, mostraremos como se transforman las redes de procesos bioquímicos para que tengan deficiencia cero (y sean al menos débilmente reversibles) y así el método para encontrar la distribución estacionaria pueda ser aplicado. Hay que recordar que la deficiencia se calcula como $\Delta = |C| - l - s$, donde $|C|$ es el número de reactantes y reactivos, l es el número de clases de enlace del grafo de red, y s es la dimensión del subespacio estequiométrico de la red.

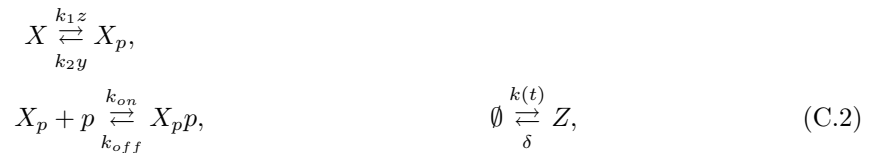
C.1. Fosforilación

C.1.1. Caso I

Primero trataremos al proceso de fosforilación/desfosforilación cuando no existen procesos intermedios, recordar que las reacciones que describen a este proceso son las siguientes



obtenemos los siguientes valores $|C| = 8$, $l = 4$ y $s = 3$, por lo que ahora la deficiencia será $\Delta = 1$, si ahora usamos las ideas de [20] y modificamos nuestras reacciones de la siguiente manera



(siendo z y y las concentraciones de Z y Y respectivamente), ahora volveremos a calcular la deficiencia, con $|C| = 6$, $l = 3$ y $s = 3$, obtenido $\Delta = 0$, por lo que ahora podremos aplicar el método que hemos desarrollado para encontrar la forma de la distribución estacionaria. Algo similar se haría cuando existen procesos intermedios.

C.2. Redes de regulación genética

Los cálculos que realizaremos son válidos para los dos tipos de redes que analizamos. Hay que recordar que las reacciones que describen estos sistemas son



estas reacciones corresponden a un solo modulo de transcripción/traducción (MTT), calculamos su deficiencia con $|C| = 7$, $l = 3$ y $s = 2$, obteniendo $\Delta_1 = 2$, por lo que para una red con n módulos obtendremos $\Delta_n = 2n$, si ahora usamos las ideas de [20] y modificamos nuestras reacciones de la siguiente manera



donde r es la concentración del ARm, ahora tendremos $|C| = 4$, $l = 2$ y $s = 2$, obteniendo $\Delta_1 = 0$, y para una red de regulación genética con n módulos tendremos $\Delta_n = 0$, por lo que ahora ya podremos determinar la distribución estacionaria del sistema.

Apéndice D

Proceso de vida y muerte

Usando el formalismo de operadores resolvemos el problema de vida-muerte de manera analítica, recordar que encontrar una solución a este problema es particularmente complicado. En el problema de vida y muerte se supone que unos individuos (o partículas) son creadas a una tasa fija, mientras que estos mueren en proporción a los que están vivos, es decir, tienen una transición de probabilidad similar a

$$W(x|x', t) = t^+(x)\delta_{x,x'+1} + t^-(x)\delta_{x,x'-1}. \quad (\text{D.1})$$

donde $t^+ = \gamma_1$ y $t^- = \gamma_2 x$, por ahora no introduciremos estas cantidades explícitamente. Por ahora explicaremos de manera general como son los pasos para resolver esta ecuación usando el formalismo de operadores, y sorprendentemente tiene una gran similitud con lo que está presente en los cursos de mecánica cuántica. Nuestro sistema solo tiene 2 procesos, por lo que ahora nuestra ecuación maestra tomara la siguiente forma

$$\partial_t P(x, t|x', t') = t^+(x-1)P(x-1, t|x', t') + t^-(x+1)P(x+1, t|x', t') - [t^+(x) + t^-(x)]P(x, t|x', t'), \quad (\text{D.2})$$

simplificamos un poco la notación tenemos,

$$\partial_t p_x(t) = t^+(x-1)p_{x-1}(t) + t^-(x+1)p_{x+1}(t) - [t^+(x) + t^-(x)]p_x(t). \quad (\text{D.3})$$

En base a la notación usada en el capítulo 3 aplicamos esta ecuación sobre un ket $|x\rangle$ y sumamos sobre esta variable

$$\sum_x \partial_t p_x(t) |x\rangle = \sum_x t^+(x-1)p_{x-1}(t) |x\rangle + \sum_x t^-(x+1)p_{x+1}(t) |x\rangle - \sum_x [t^+(x) + t^-(x)]p_x(t) |x\rangle, \quad (\text{D.4})$$

aunque no hemos definido los límites de esta suma es bastante obvio que la suma empieza de cero, ya que el mínimo número de individuos es cero, simplificando la relación anterior,

$$\partial_t |G(t)\rangle = \sum_{x'} p_{x'}(t) t^+(x') |x'+1\rangle + \sum_{x''} p_{x''}(t) t^-(x'') |x''-1\rangle - \sum_x p_x(t) [t^+(x) + t^-(x)] |x\rangle. \quad (\text{D.5})$$

Ahora suponemos que existen un conjunto de operadores que son de la siguiente forma

$$\begin{aligned} a^- |x\rangle &= t^-(x) |x-1\rangle, & a^+ |x\rangle &= t^+(x) |x+1\rangle, \\ b^- |x\rangle &= t^-(x) |x\rangle, & b^+ |x\rangle &= t^+(x) |x\rangle, \end{aligned} \quad (\text{D.6})$$

se tiene además la propiedad $a^- |0\rangle = 0$ (la cota de la suma que hemos mencionado). Como ya hemos mencionado en la sección anterior el objetivo de definir operadores de esta manera es para hacer aparecer al ket $|G(t)\rangle$. Estos operadores tienen un gran parecido con los operadores de

aniquilación y creación presentes en el problema del oscilador cuántico [21], con estos operadores (D.5) se simplifica

$$\partial_t |G(t)\rangle = (a^+ + a^- - b^+ - b^-) |G(t)\rangle = -L |G(t)\rangle. \quad (\text{D.7})$$

Esta ecuación es muy similar a la que se encuentra en mecánica cuántica [21], más concretamente muy parecida a la ecuación de Schrödinger, en el lado derecho de esta ecuación ningún operador depende del tiempo podemos solucionarlo fácilmente de la siguiente manera: primero suponemos que existe un operador de evolución de tal manera que

$$|G(t)\rangle = U(t) |G(0)\rangle, \quad (\text{D.8})$$

por lo que ahora solo faltaría encontrar este operador de evolución, y este es justo

$$U(t) = \exp(-Lt). \quad (\text{D.9})$$

En el caso en el que el operador L dependa del tiempo, y su conmutador a tiempos diferentes sea diferente de cero, uno podría usar las series de Dyson (aunque se tendría que cambiar ligeramente la notación para hacerla más entendible) para más detalles de cómo solucionar esto se puede ver [21]. Este operador L que hemos usado para simplificar un poco la notación podría englobar un gran número de operadores, además se podría usar incluso cuando están involucradas muchas variables o procesos de varios pasos. Ahora nos enfocamos con más detalle en este problema, primero volvemos a escribir nuestra ecuación maestra

$$\partial_t p_x(t) = \gamma_1 p_{x-1}(t) + \gamma_2 (x+1) p_{x+1}(t) - [\gamma_1 + \gamma_2(x)] p_x(t). \quad (\text{D.10})$$

En base a la notación usada en la sección anterior y las relaciones que aparecen en (E.4), aplicamos a la ecuación anterior un ket $|x\rangle$ y luego sumamos sobre esta variable

$$\sum_x \partial_t p_x(t) |x\rangle = \sum_x \gamma_1 p_{x-1}(t) |x\rangle + \sum_x \gamma_2 (x+1) p_{x+1}(t) |x\rangle - \sum_x [\gamma_1 + \gamma_2(x)] p_x(t) |x\rangle, \quad (\text{D.11})$$

En base a esta ecuación suponemos que nuestros operadores actúan de la siguiente manera (para el caso de procesos de vida-muerte)

$$a^- |x\rangle = x |x-1\rangle, \quad a^+ |x\rangle = |x+1\rangle, \quad (\text{D.12})$$

en el caso del operador de aniquilación se ve que la disminución es de manera lineal, mientras que la creación es constante, de estas relaciones se calcula que el corchete de Poisson de los operadores el cual es

$$[a^-, a^+] = 1. \quad (\text{D.13})$$

Además, es fácil mostrar la siguiente relación $a^- a^- |x\rangle = x |x\rangle$, esta será de utilidad a continuación. Con estos operadores ahora tendremos la siguiente ecuación

$$\frac{d}{dt} |G(t)\rangle = -L |G(t)\rangle. \quad (\text{D.14})$$

Donde el operador L esta dado como

$$-L = -\gamma_1 a^+ - \gamma_2 a^- + b^+ + b^- = a^+ a^- \gamma_2 - \gamma_1 a^+ - \gamma_2 a^- + \gamma_1 = \gamma_2 (a^+ - 1) \left(a^- - \frac{\gamma_1}{\gamma_2} \right). \quad (\text{D.15})$$

Ahora definimos unos operadores de aniquilación y creación

$$c^+ = (a^+ - 1), \quad c^- = \left(a^- - \frac{\gamma_1}{\gamma_2} \right), \quad (\text{D.16})$$

el objetivo de tener operadores de creación y aniquilación es poder resolver el problema de una manera similar al oscilador armónico, con estos operadores finalmente tenemos

$$L = -\gamma_2 c^+ c^-. \quad (\text{D.17})$$

Estos operadores tienen el siguiente conmutador

$$[c^-, c^+] = 1, \quad (\text{D.18})$$

y nos recuerdan mucho a los operadores escalera usados al solucionar el oscilador armónico, por lo que se tendría una base $|c\rangle$ tal que se cumple

$$L|c\rangle = \gamma_2 l_c |c\rangle, \quad (\text{D.19})$$

con esto evaluamos (D.9), por lo que ahora

$$U(t) = \exp(-Lt) = \sum_{c'} \sum_{c''} |c''\rangle \langle c''| \exp(-Lt) |c'\rangle \langle c'| = \sum_{c'} |c'\rangle \exp(-\gamma_2 l_{c'} t) \langle c'|, \quad (\text{D.20})$$

podemos expandir también al ket $|G(0)\rangle$ en la base del ket $|c\rangle$ de la siguiente manera $|G(0)\rangle = \sum_c A_c |c\rangle$, con esto la expresión (D.8) será

$$|G(t)\rangle = \sum_c \exp(-\gamma_2 l_c t) A_c |c\rangle. \quad (\text{D.21})$$

Ahora veremos cómo son estos eigenvalores, para esto vemos que los conmutadores serán

$$[L, c^+] = \gamma_2 c^+, \quad [L, c^-] = -\gamma_2 c^-, \quad (\text{D.22})$$

por lo que al desarrollar siguiente

$$Lc^+ |c\rangle = (c^+ L + \gamma_2 c^+) |c\rangle = \gamma_2 (l_c + 1) (c^+ |c\rangle), \quad (\text{D.23})$$

de este resultado podemos inferir (procedemos de manera similar con a^-)

$$c^+ |c\rangle = |c+1\rangle, \quad c^- |c\rangle = c |c-1\rangle, \quad (\text{D.24})$$

de estos resultados podemos obtener lo siguiente

$$|c\rangle = (c^+)^c |0\rangle, \quad (c^-) |0\rangle = 0, \quad (\text{D.25})$$

si recordamos lo que se hace en mecánica cuántica con los operadores escalera, existe una correspondencia entre los operadores escalera y la posición y la velocidad, de hecho, es la representación de los operadores en otra base,

$$a^+ \rightarrow s, \quad a^- \rightarrow \partial_s,$$

por lo que ahora suponemos que para nuestro sistema existe una base en donde la proyección es tal como en las relaciones anteriores, de hecho se puede mostrar que esta corresponde a la base $|s\rangle$, para verlo uno puede recordar lo que se hace cuando se usa la función generatriz para resolver la ecuación maestra, por lo que los otros operadores serán

$$c^- \rightarrow \left(\partial_s - \frac{\gamma_1}{\gamma_2} \right), \quad c^+ \rightarrow (s-1),$$

con ayuda de las relaciones anteriores y de (D.25) obtenemos

$$\langle s|c\rangle = \langle s| (c^+)^c |0\rangle = (s-1)^c \langle s|0\rangle, \quad (\text{D.26})$$

se ha aplicado el operador en el bra, para encontrar el $\langle s|0\rangle$, aplicaremos c^- como a continuación

$$\langle s|c^-|0\rangle = 0 = \left(\partial_s - \frac{\gamma_1}{\gamma_2}\right) \langle s|0\rangle, \quad (\text{D.27})$$

por un lado, lo aplicamos sobre el $|0\rangle$ y por el otro sobre $\langle s|$, esta es una ecuación diferencial tiene como solución $\langle s|0\rangle = N e^{\left(\frac{\gamma_1}{\gamma_2}\right)s}$, el factor N es una constante de normalización, con lo que finamente tenemos

$$\langle s|c\rangle = (s-1)^c N e^{\left(\frac{\gamma_1}{\gamma_2}\right)s}, \quad (\text{D.28})$$

ahora calculamos $\langle c|s\rangle$,

$$\delta_{c,c'} = \langle c|c'\rangle = \oint \frac{ds}{2\pi i} \langle c|s\rangle \langle s|c'\rangle = \oint \frac{ds}{2\pi i} \langle c|s\rangle (s-1)^c N e^{\left(\frac{\gamma_1}{\gamma_2}\right)s}, \quad (\text{D.29})$$

para que se cumpla el teorema integral de Cauchy se tiene que elegir

$$\langle c|s\rangle = \frac{1}{N} \frac{e^{-\left(\frac{\gamma_1}{\gamma_2}\right)s}}{(s-1)^{c+1}}, \quad (\text{D.30})$$

calculamos ahora lo siguiente $\langle c|x\rangle$ y $\langle x|c\rangle$ con ayuda de las relaciones (E.4), (D.28) y (D.30)

$$\langle c|x\rangle = \oint \frac{ds}{2\pi i} \langle c|s\rangle \langle s|x\rangle = \oint \frac{ds}{2\pi i} \frac{1}{N} \frac{e^{-\left(\frac{\gamma_1}{\gamma_2}\right)s}}{(s-1)^{c+1}} s^x = \oint \frac{ds'}{2\pi i} \frac{1}{N} \frac{e^{-\left(\frac{\gamma_1}{\gamma_2}\right)(s'+1)}}{(s')^{c+1}} (s'+1)^x, \quad (\text{D.31})$$

aplicando el teorema integral de Cauchy

$$\begin{aligned} \langle c|x\rangle &= \frac{1}{N} \frac{1}{c!} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)c} \partial_{s'}^c \left[e^{-\left(\frac{\gamma_1}{\gamma_2}\right)s'} (s'+1)^x \right]_{s'=0} \\ &= \frac{1}{N} \frac{1}{c!} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)c} \sum_{l=0}^c \frac{c!}{(c-l)!} \partial_{s'}^{c-l} \left[e^{-\left(\frac{\gamma_1}{\gamma_2}\right)s'} \right]_{s'=0} \partial_{s'}^l [(s'+1)^x]_{s'=0} \\ &= \frac{1}{N} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)c} \sum_{l=0}^c \frac{1}{(c-l)!} \left(-\frac{\gamma_1}{\gamma_2} \right)^{c-l} \frac{x!}{(x-l)!} \theta(x-l) \\ &= \frac{x!}{N} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)c} \left(-\left(\frac{\gamma_1}{\gamma_2}\right) \right)^c \sum_{l=0}^c \frac{1}{(c-l)!} \left(-\frac{\gamma_1}{\gamma_2} \right)^{-l} \frac{1}{(x-l)!} \theta(x-l). \end{aligned} \quad (\text{D.32})$$

El término del sumatorio puede ser evaluado usando algún sistema computacional (en este caso se usa Mathematica) se encuentra que es proporcional a la función hipergeométrica generaliza

$$\langle c|x\rangle = \frac{(-1)^c x!}{N} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)c} \left(\frac{\gamma_1}{\gamma_2}\right)^c A(c, x), \quad (\text{D.33})$$

la z se define como $z = \min\{x, c\}$, la función $A(c, x)$ se define como a continuación

$$A(c, x) = \frac{\left(-\frac{\gamma_1}{\gamma_2}\right)^{-z} {}_2F_2\left(1, -z; c-z+1, x-z+1; \frac{\gamma_1}{\gamma_2}\right)}{z!(c-z)!(x-z)!} = \frac{\left(-\frac{\gamma_1}{\gamma_2}\right)^{-z} {}_1F_1\left(-z; |x-c|+1; \frac{\gamma_1}{\gamma_2}\right)}{z!(|x-c|)!}. \quad (\text{D.34})$$

En esta expresión se ve como depende de la función higrométrica confluyente y hay una variable que se tiene que introducir dependiendo de los valores de x y c . Ahora calculamos el siguiente termino y procedemos de una manera similar que el termino anterior,

$$\langle x|c\rangle = \oint \frac{ds}{2\pi i} \langle x|s\rangle \langle s|c\rangle = \oint \frac{ds}{2\pi i} \frac{(s-1)^c N e^{\left(\frac{\gamma_1}{\gamma_2}\right)s}}{s^{x+1}}, \quad (\text{D.35})$$

aplicando el teorema integral de Cauchy

$$\begin{aligned}
 \langle x|c \rangle &= \frac{N}{x!} \partial_s^x \left[(s-1)^c e^{\left(\frac{\gamma_1}{\gamma_2}\right)s} \right]_{s=0} \\
 &= \frac{N}{x!} \sum_{l=0}^x \frac{x!}{(x-l)!} \partial_s^{x-l} \left[e^{\left(\frac{\gamma_1}{\gamma_2}\right)s} \right]_{s=0} \partial_s^l [(s-1)^c]_{s=0} \\
 &= N \sum_{l=0}^x \frac{1}{(x-l)!} \left(\frac{\gamma_1}{\gamma_2} \right)^{x-l} \frac{c!}{(c-l)!} (-1)^{c-l} \theta(c-l),
 \end{aligned} \tag{D.36}$$

reacomodamos el termino dentro del sumatorio y volvemos a usar Mathematica con lo con lo que finalmente tenemos

$$\langle x|c \rangle = (-1)^c c! N \left(\frac{\gamma_1}{\gamma_2} \right)^x A(c, x). \tag{D.37}$$

Con esto encontramos que esta expresión aún depende de las funciones hipergeométricas. Con los cálculos anteriores ya se puede encontrar el valor de la variable A_c , por lo general podemos suponer que $p_x(0) = \delta_{x,y}$, es decir existe algún determinado número y de elementos en el instante $t = 0$, para calcular esta variable nos fijamos en el ket $|G(t)\rangle$ en $t = 0$ y luego aplicamos $\langle c|$,

$$A_c = \langle c|G(0)\rangle = \sum_x p_x(0) \langle c|x\rangle = \langle c|y\rangle = \frac{(-1)^c y!}{N} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)} \left(\frac{\gamma_1}{\gamma_2} \right)^c A(c, y), \tag{D.38}$$

regresando a la expresión (D.21) podemos encontrar a $p_x(t)$ al contraer con $\langle x|$, entonces

$$p_x(t) = \langle x|G(t)\rangle = \sum_c \exp(-\gamma_2 ct) A_c \langle x|c\rangle. \tag{D.39}$$

Sustituyendo (D.38) y (D.37) en esta última expresión

$$p_x(t) = \sum_{c=0}^{\infty} c! * y! * e^{-\gamma_2 ct} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)} \left(\frac{\gamma_1}{\gamma_2} \right)^{x+c} A(c, y) A(c, x). \tag{D.40}$$

Con este resultado también podemos ver que el factor N , es innecesario y podemos haberlo eliminado. El estado estacionario es aquel cuando el parámetro t crece (o tiende a infinito), por lo que en (D.40) casi todos los términos del sumatorio tienden a cero, el único termino que no tiende a cero es el que corresponde al eigenvalor $c = 0$, por lo que podemos considerarlo como el estado estacionario de nuestro sistema, entonces la p_x estacionaria de nuestro sistema será

$$p_x(S) = y! e^{-\left(\frac{\gamma_1}{\gamma_2}\right)} \left(\frac{\gamma_1}{\gamma_2} \right)^x A(0, y) A(0, x) = \frac{1}{x!} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)} \left(\frac{\gamma_1}{\gamma_2} \right)^x. \tag{D.41}$$

Este resultado es justo una distribución de Poisson, con la particularidad que las gammas pueden tomar muchos valores, además esta solo depende de la razón entre las gammas.

En la figura D.1 se puede observar una gráfica de la distribución del proceso de vida y muerte, aquí se observa como la distribución de probabilidad se acerca al estacionario relativamente rápido. Con ayuda de los resultados también es fácil obtener la ecuación generatriz de nuestro sistema, el cual nos ayudaría a obtener de manera más sencilla el promedio la variable x y su varianza. Para encontrar la función generatriz

$$G(s, t) = \langle s|G(t)\rangle = \sum_c \exp(-\gamma_2 ct) A_c \langle s|c\rangle. \tag{D.42}$$

Sustituyendo (D.38) y (D.26) en esta última expresión

$$G(s, t) = \sum_{c=0}^{\infty} (-1)^c y! * e^{-\gamma_2 ct} e^{-\left(\frac{\gamma_1}{\gamma_2}\right)} \left(\frac{\gamma_1}{\gamma_2} \right)^c A(c, y) (s-1)^c e^{\left(\frac{\gamma_1}{\gamma_2}\right)s}. \tag{D.43}$$

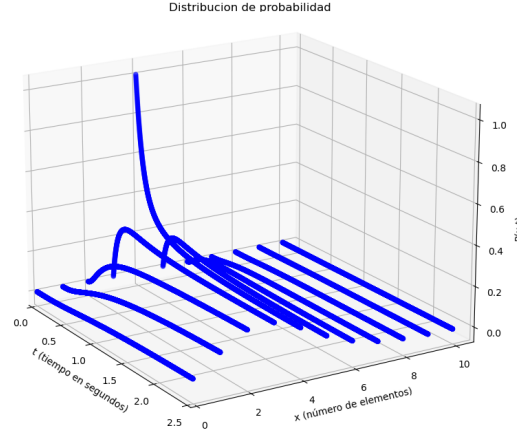


Figura D.1: Distribución del proceso de vida y muerte de una variable a lo largo del tiempo.

Recordar que se cumple lo siguiente

$$\langle x(t) \rangle = \frac{\partial G(s, t)}{\partial s} \Big|_{s=1}, \quad \langle x^2(t) \rangle - \langle x(t) \rangle = \frac{\partial^2 G(s, t)}{\partial s^2} \Big|_{s=1}. \quad (\text{D.44})$$

Calculamos la primera expresión por lo que tenemos

$$\langle x(t) \rangle = \left(\frac{\gamma_1}{\gamma_2} \right) (1 - y!A(1, y)e^{-\gamma_2 t}) = \left(\frac{\gamma_1}{\gamma_2} \right) \left(1 + \left(\left(\frac{\gamma_2}{\gamma_1} \right) y - 1 \right) e^{-\gamma_2 t} \right), \quad (\text{D.45})$$

esta expresión corresponde al promedio de la variable x , esta expresión de hecho coincide exactamente con la solución determinista, ahora para calcular la varianza primero tenemos que calcular

$$\langle x^2(t) \rangle - \langle x(t) \rangle = \left(\frac{\gamma_1}{\gamma_2} \right)^2 (1 - 2y!A(1, y)e^{-\gamma_2 t} + y!A(2, y)e^{-2\gamma_2 t}), \quad (\text{D.46})$$

con ayuda de este resultado finalmente podemos calcular la varianza sera $(\sigma^2(t) = \langle x^2(t) \rangle - \langle x(t) \rangle^2)$

$$\sigma^2(t) = \left(\frac{\gamma_1}{\gamma_2} \right) (1 - y!A(1, y)e^{-\gamma_2 t}) + \left(\frac{\gamma_1}{\gamma_2} \right)^2 (A(2, y) - y!A^2(1, y))y!e^{-2\gamma_2 t}. \quad (\text{D.47})$$

Podemos ver como el promedio de variable x tiende a una constante a tiempos muy grandes, de igual manera que la varianza, de hecho, como hemos mostrado la distribución estacionaria es una de Poisson por lo que a tiempos grandes se obtiene un promedio y varianza correspondiente a la distribución de Poisson. Con este último resultado ya podemos calcular las fluctuaciones en el sistema, el cual sería

$$\eta^2 = \left(\frac{\gamma_1}{\gamma_2} \right) (1 - y!A(1, y)e^{-\gamma_2 t}) + \left(\frac{\gamma_1}{\gamma_2} \right)^2 (A(2, y) - y!A^2(1, y))y!e^{-2\gamma_2 t}, \quad (\text{D.48})$$

cuando tomamos el límite de t tendiendo al infinito se recupera las fluctuaciones en el estado estacionario. En esta expresión se puede observar que al principio las fluctuaciones están influenciadas por la condición inicial.

El formalismo de operadores resulta ser bastante conveniente para resolver de forma analítica la ecuación maestra, sin embargo, no siempre es posible hacer uso de los operadores de creación y aniquilación como los hemos usado aquí, aun así, existen varios problemas en los cuales se podría aplicar este formalismo.

Apéndice E

Modelo estocástico de un MTT

A continuación, haremos un análisis estocástico de un solo MTT, similar a lo que se encuentra en la figura E.1. Para el análisis nos enfocaremos en encontrar de manera analítica la distribución de probabilidad del sistema, más aún, encontraremos a la función generatriz de donde se pueden inferir varias cantidades estadísticas.

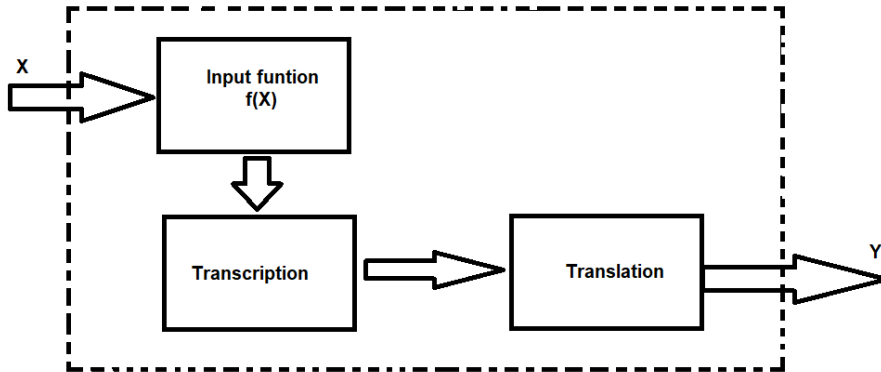
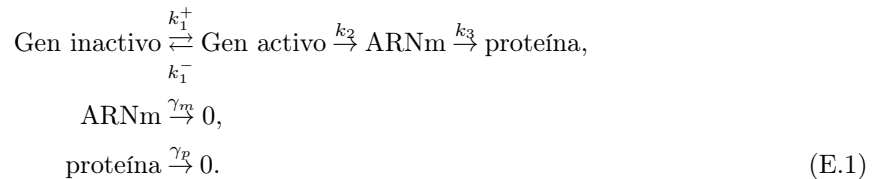


Figura E.1: Módulo de transcripción/traducción [8].

Nos centraremos en analizar el modelo estocástico que está en [30], en este ejemplo se analiza un módulo que se encuentra completamente aislado, donde los procesos que ocurren dentro son de la siguiente forma



La primera relación nos habla sobre como primero se activa y desactiva el gen a través de un proceso de tipo telégrafo aleatorio, después de que los genes están activos estos se copian para producir el ARm (proceso de transcripción), después de esto el ARm sintetiza la proteína (proceso de traducción), aunque este modelo no es tan realista principalmente en la parte de la activación de los genes, sin embargo, será suficiente para conocer un modelo estocástico del MTT. Todas esto se pone con más detalle en las siguientes reacciones, con g denotando a los genes activos, m al ARm, p a la proteína,

Activación genética: $g \xrightarrow{k_1^+} g + 1$

Desactivación genética: $g \xrightarrow{k_1^-} g - 1$

Trascrición: $m \xrightarrow{k_2} m + 1$

Degradación de ARm: $m \xrightarrow{\gamma_m} m - 1$

Traducción: $p \xrightarrow{k_3} p + 1$

Degradación de la proteína: $p \xrightarrow{\gamma_p} p - 1$

De las reacciones anteriores se desprende la siguiente ecuación maestra, asumiendo que el proceso es Markoviano y además los siguientes tiempos de actividad genética $\tau_1 = \frac{1}{k_1^+} + k_1^-$, mientras que $\tau_2 = \frac{1}{\gamma_m}$ y $\tau_3 = \frac{1}{\gamma_p}$ son los tiempos de degradación del ARm y la proteína respectivamente, hemos supuesto que hay un número máximo de copias de un gen que pueden ser activados g^{max} , en muchos modelos g^{max} oscila entre 1-8,

$$\begin{aligned} \frac{dP(g, m, p)}{dt} = & k_1^+(g^{max} - g + 1)P(g - 1, m, p) - k_1^+(g^{max} - g)P(g, m, p) \\ & + k_1^-(g + 1)P(g + 1, m, p) - k_1^-gP(g, m, p) \\ & + \frac{m + 1}{\tau_2}P(g, m + 1, p) - \frac{m}{\tau_2}P(g, m, p) + k_2gP(g, m + 1, p) - k_2gP(g, m, p) \\ & + k_3mP(g, m, p - 1) - k_3mP(g, m, p) + \frac{p + 1}{\tau_3}P(g, m, p + 1) - \frac{p}{\tau_3}P(g, m, p). \end{aligned} \quad (E.2)$$

Ahora resolveremos estas ecuaciones desde la perspectiva de operadores, similar a las ideas presentadas en [11], además resolvemos en el apéndice un ejemplo con este método y de donde tomaremos algunas ideas. Consideramos las siguientes convenciones

$$|G(t)\rangle = \sum_{x_i} p_{x_i}(t) |x_i\rangle = \sum_{g=0}^{g^{max}} \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} p_{(g,m,p)}(t) |g, m, p\rangle, \quad (E.3)$$

los sub-índices engloban tanto al gen(g), el ARNm(m) y a la proteína(p), en este orden en específico. Se ha adoptado una notación que es muy común en mecánica cuántica [21], se puede recuperar la notación con la función generatriz de la siguiente manera (estos resultados se obtienen de acuerdo con la notación de Dirac)

$$\begin{aligned} \langle s|G(t)\rangle &= G(s, t), & \langle x|G(t)\rangle &= p_x(t), \\ \langle s|x\rangle &= s^x, & \langle x|s\rangle &= \frac{1}{s^{x+1}}, \\ \langle x|x'\rangle &= \delta_{x,x'}, \end{aligned} \quad (E.4)$$

a la ecuación (E.2) se le aplica un ket y se le suma sobre todos los posibles $x_i = g, m, p$, obteniendo

$$\begin{aligned} \sum_{x_i} \frac{dP(x_i)}{dt} |x_i\rangle = & - \sum_{x_i} \left(k_1^+(g - 1 - g^{max})P(g - 1, m, p) |x_i\rangle + k_1^+(g - g^{max})P(g, m, p) |x_i\rangle \right. \\ & + k_1^-(g + 1)P(g + 1, m, p) |x_i\rangle - k_1^-gP(g, m, p) |x_i\rangle \\ & + \frac{m + 1}{\tau_2}P(g, m + 1, p) |x_i\rangle - \frac{m}{\tau_2}P(g, m, p) |x_i\rangle + k_2gP(g, m + 1, p) |x_i\rangle - k_2gP(g, m, p) |x_i\rangle \\ & \left. + k_3mP(g, m, p - 1) |x_i\rangle - k_3mP(g, m, p) |x_i\rangle + \frac{p + 1}{\tau_3}P(g, m, p + 1) |x_i\rangle - \frac{p}{\tau_3}P(g, m, p) |x_i\rangle \right). \end{aligned} \quad (E.5)$$

Ahora definimos el siguiente conjunto de operadores

$$\begin{aligned}
 a_1^- |g\rangle &= g |g-1\rangle, & a_1^+ |g\rangle &= |g+1\rangle, \\
 a_2^- |m\rangle &= m |m-1\rangle, & a_2^+ |m\rangle &= |m+1\rangle, \\
 a_3^- |p\rangle &= p |p-1\rangle, & a_3^+ |p\rangle &= |p+1\rangle, \\
 b_1 |g\rangle &= -g |g+1\rangle, & &
 \end{aligned} \tag{E.6}$$

se puede ver como los operadores actúan cada uno sobre diferentes variables, para ser precisos es una notación tensorial la que se usa (para más detalles ver [26]), por lo que ahora será suficiente con analizar el corchete de Poisson de la siguiente manera,

$$[a_1^-, a_1^+] = [a_2^-, a_2^+] = [a_3^-, a_3^+] = 1. \tag{E.7}$$

Con estos resultados ahora reescribimos a (E.5), simplificamos esta expresión introduciendo $g = a_1^+ a_1^-$,

$$\begin{aligned}
 \frac{d}{dt} |G(t)\rangle &= k_1^+ g^{max} a_1^+ |G(t)\rangle + k_1^+ b_1 |G(t)\rangle + k_1^+ (a_1^+ a_1^- - g^{max}) |G(t)\rangle + k_1^- a_1^- |G(t)\rangle - k_1^- a_1^+ a_1^- |G(t)\rangle \\
 &+ \frac{1}{\tau_2} a_2^- |G(t)\rangle - \frac{1}{\tau_2} a_2^+ a_2^- |G(t)\rangle + k_2 a_1^+ a_1^- a_2^+ |G(t)\rangle - a_1^+ a_1^- |G(t)\rangle \\
 &+ k_3 a_3^+ a_2^+ a_2^- |G(t)\rangle - k_3 a_2^+ a_2^- |G(t)\rangle + \frac{1}{\tau_3} a_3^- |G(t)\rangle - \frac{1}{\tau_3} a_3^+ a_3^- |G(t)\rangle.
 \end{aligned} \tag{E.8}$$

reacomodamos términos ($\gamma = \frac{k_1^-}{k_1^+}$)

$$\begin{aligned}
 \frac{d}{dt} |G(t)\rangle &= k_1^+ \left(-(\gamma - 1) \left(a_1^+ + \frac{\gamma}{1 - \gamma} \right) \left(a_1^- + \frac{g^{max}}{1 - \gamma} \right) - \frac{g^{max}}{1 - \gamma} + b_1 \right) |G(t)\rangle \\
 &+ \frac{1}{\tau_2} (1 - a_2^+) a_2^- |G(t)\rangle + k_2 a_1^+ a_1^- (a_2^+ - 1) |G(t)\rangle + k_3 (a_3^+ - 1) a_2^+ a_2^- |G(t)\rangle + \frac{1}{\tau_3} (1 - a_3^+) a_3^- |G(t)\rangle.
 \end{aligned} \tag{E.9}$$

introducimos el siguiente conjunto de operadores inspirados en lo que se ha hecho en el apéndice para el proceso de vida y muerte, con el objetivo de formar operadores escalera similares a los que se encuentran en el problema del oscilador armónico cuántico

$$\begin{aligned}
 b_1^- &= a_1^- + \frac{g^{max}}{1 - \gamma}, & b_1^+ &= a_1^+ + \frac{\gamma}{1 - \gamma}, \\
 b_2^- &= a_2^-, & b_2^+ &= a_2^+ - 1, \\
 b_3^- &= a_3^-, & b_3^+ &= a_3^+ - 1,
 \end{aligned} \tag{E.10}$$

ahora tendremos

$$\begin{aligned}
 \frac{d}{dt} |G(t)\rangle &= k_1^+ \left(-(\gamma - 1) b_1^+ b_1^- - \frac{g^{max}}{1 - \gamma} + b_1 - \frac{1}{\tau_2} b_2^+ b_2^- \right) |G(t)\rangle \\
 &+ \left(k_2 \left(b_1^+ - \frac{\gamma}{1 - \gamma} \right) \left(b_1^- - \frac{g^{max}}{1 - \gamma} \right) b_2^+ + k_3 b_3^+ (b_2^+ b_2^- + b_2^-) - \frac{1}{\tau_3} b_3^+ b_3^- \right) |G(t)\rangle,
 \end{aligned} \tag{E.11}$$

suponemos que existe una base $(|j, k, l\rangle)$ de tal manera que los operador actúan aquí de manera similar a los operadores escalera, por la definición de (E.10) estos operadores aun cumplen corchetes de Poisson similar a los operadores (E.6) por lo que de inmediato se encuentran las siguientes relaciones,

$$\begin{aligned}
 b_1^- |j\rangle &= j |j-1\rangle, & b_1^+ |j\rangle &= |j+1\rangle, \\
 b_2^- |k\rangle &= k |k-1\rangle, & b_2^+ |k\rangle &= |k+1\rangle, \\
 b_3^- |l\rangle &= l |l-1\rangle, & b_3^+ |l\rangle &= |l+1\rangle.
 \end{aligned} \tag{E.12}$$

Con esto en mente podemos expandir a $|G(t)\rangle = \sum_{jkl} G_{j,k,l}^t |j, k, l\rangle$, podemos ir un poco más lejos y proponemos el siguiente anzatz,

$$|G(t)\rangle = \sum_{jkl} e^{-\left(k_1^+ \left((\gamma-1)j + \frac{g^{max}}{1-\gamma}\right) + \frac{k}{\tau_2} + \frac{l}{\tau_3}\right)t} G_{j,k,l}(t) |j, k, l\rangle, \quad (E.13)$$

la estructura está inspirada en la solución del problema de vida y muerte que aparece en el apéndice, sustituimos nuestro anzatz en la ecuación (E.11), con lo cual obtenemos

$$\sum_{jkl} \dot{G}_{j',k',l'}(t) |j', k', l'\rangle = \sum_{j',k',l'} \left(k_1^+ b_1 + k_2 \left(b_1^+ - \frac{\gamma}{1-\gamma} \right) \left(b_1^- - \frac{g^{max}}{1-\gamma} \right) b_2^+ + k_3 b_3^+ (b_2^+ b_2^- + b_2^-) \right) G_{j',k',l'}(t) |j', k', l'\rangle. \quad (E.14)$$

aplicamos a la relación anterior un bra $\langle l, k, j|$ por la izquierda, con lo que finalmente obtenemos

$$\dot{G}_{j,k,l}(t) = k_1^+ \sum_{j'} \Gamma_{j,j'}^1 G_{j',k,l}(t) + k_2 \sum_{j',k'} \Gamma_{j,j'}^2 \delta_{k-1,k'} G_{j',k',l}(t) + k_3 \sum_{k',l'} \Gamma_{k,k'}^3 \delta_{l-1,l'} G_{j,k',l'}(t). \quad (E.15)$$

Donde

$$\begin{aligned} \Gamma_{j,j'}^1 &= \langle j|b_1|j'\rangle = \sum_{g=0}^{g^{max}} \langle j|b_1|g\rangle \langle g|j'\rangle = - \sum_{g=0}^{g^{max}} g \langle j|g+1\rangle \langle g|j'\rangle, \\ \Gamma_{j,j'}^2 &= \langle j| \left(b_1^+ - \frac{\gamma}{1-\gamma} \right) \left(b_1^- - \frac{g^{max}}{1-\gamma} \right) |j'\rangle = \sum_{g=0}^{g^{max}} \langle j| \left(b_1^+ - \frac{\gamma}{1-\gamma} \right) |g\rangle \langle g| \left(b_1^- - \frac{g^{max}}{1-\gamma} \right) |j'\rangle, \\ \Gamma_{k,k'}^3 &= \langle k|b_2^+ b_2^- + b_2^- |k'\rangle = k' (\delta_{k,k'} + \delta_{k,k'-1}). \end{aligned} \quad (E.16)$$

La estructura de (E.15) es muy similar a una ecuación diferencial matricial por lo que su solución sera.

$$\vec{G}(t) = G_{j,k,l}(t) = e^{\left(k_1^+ \Gamma^1 + k_2 \Gamma^2 I_k + k_3 \Gamma^3 I_l\right)t} \vec{A}. \quad (E.17)$$

Donde $I_k = \delta_{k-1,k'}$ y $I_l = \delta_{l-1,l'}$ son matrices, por lo que ahora solo queda calcular los términos $\vec{A} = A_{j,k,l}$, omitiremos gran parte del desarrollo debido a que ya hemos la hemos calculado en el apéndice, para ello solo daremos algunos detalles de cómo proceder. Con (E.17) nuestra solución sera

$$|G(t)\rangle = \sum_{jkl} e^{-\left(k_1^+ \left((\gamma-1)j + \frac{g^{max}}{1-\gamma}\right) + \frac{k}{\tau_2} + \frac{l}{\tau_3}\right)t} G_{j,k,l}(t) |j, k, l\rangle, \quad (E.18)$$

para calcular el término $A_{j,k,l}$ primero nos fijamos en lo siguiente $\langle l, k, j|G(0)\rangle = A_{j,k,l}$, mientras que por otro lado $|G(0)\rangle = \sum_{g,m,p} p_{g,m,p}(0) |g, m, p\rangle$, en general muchas veces para las condiciones iniciales se asume que nuestro sistema tiene algún número determinado de elementos, tal como que todas las variables empiezan en cero, aunque esto aplicaría para este problema, nosotros tomaremos como condición inicial algo más general $p_{g,m,p}(0) = \delta_{g,g_0} \delta_{m,m_0} \delta_{p,p_0}$, juntando todo tenemos

$$A_{j,k,l} = \langle l, k, j|G(0)\rangle = \sum_{g,m,p} p_{g,m,p}(0) \langle l, k, j|g, m, p\rangle = \langle l, k, j|g_0, m_0, p_0\rangle, \quad (E.19)$$

por la forma de esta ecuación uno puede calcular cada una de manera independiente, es decir cambiaremos un poco la notación (para más detalles ver [21, 26])

$$\langle l, k, j|g_0, m_0, p_0\rangle = \langle j|g_0\rangle \langle k|m_0\rangle \langle l|p_0\rangle. \quad (E.20)$$

En la parte final se realizan los cálculos detallados para cada una de estas cantidades, se obtiene lo siguiente ($\gamma_3 = \frac{1}{1-\gamma}$)

$$\begin{aligned}\langle j|g\rangle &= e^{-\gamma\gamma_3(g^{max}\gamma_3)} (\gamma\gamma_3)^j (-\gamma_3g^{max})^g A_1(j, g, z), & \langle g|j\rangle &= j!(-\gamma\gamma_3)^g (\gamma_3g^{max})^j A_1(j, g, z), \\ \langle k|m\rangle &= \frac{m!}{k!(m-k)!} \theta(m-k), & \langle m|k\rangle &= \frac{1}{m!} \frac{k!}{(k-m)!} (-1)^{k-m} \theta(k-m), \\ \langle l|p\rangle &= \frac{p!}{l!(p-l)!} \theta(p-l), & \langle p|l\rangle &= \frac{1}{p!} \frac{l!}{(l-p)!} (-1)^{k-m} \theta(k-m),\end{aligned}\quad (E.21)$$

la z se define como $z = \min\{g, j\}$, $A_1(j, g, z) = \frac{(-\gamma\gamma_3^2g^{max})^{-z} {}_1F_1(-z; |g-j|+1; \gamma\gamma_3^2g^{max})}{z!(|g-j)!}$, con estos resultados finalmente tendremos

$$A_{j,k,l} = e^{-\gamma\gamma_3(g^{max}\gamma_3)} (\gamma\gamma_3)^j (-\gamma_3g^{max})^{g_0} A_1(j, g_0, z_1) \left(\frac{m_0!}{k!(m_0-k)!} \theta(m_0-k) \right) \frac{p_0!}{l!(p_0-l)!} \theta(p_0-l),$$

de las relaciones (E.21) se calcula de inmediato las siguientes cantidades

$$\begin{aligned}\Gamma_{j,j'}^1 &= j'! e^{-\gamma\gamma_3(g^{max}\gamma_3)} \sum_{g=0}^{g^{max}} g (\gamma\gamma_3)^{j+g} (\gamma_3g^{max})^{g+1+j'} A_1(j, g+1, z) A_1(j', g, z), \\ \Gamma_{j,j'}^2 &= g^{max} \gamma_3^2 \delta_{j,j'} + \sum_{g=0}^{g^{max}} ((g+1) \langle j|g+1\rangle \langle g+1|j'\rangle - \gamma\gamma_3 \langle j|g\rangle \langle g+1|j'\rangle) + \gamma\gamma_3 \langle j|g+1\rangle \langle g|j'\rangle,\end{aligned}$$

ya con todos estos resultados y con ayuda de (E.18) calculamos la distribución de probabilidades, al multiplicar por un bra $\langle p, m, g|$ del lado izquierdo, obteniendo

$$p_{g,m,p}(t) = \sum_{j,k,l} e^{-\left(k_1^+((\gamma-1)j+\gamma_3g^{max})+\frac{k}{\tau_2}+\frac{l}{\tau_3}\right)t} G_{j,k,l}(t) \langle p, m, g|j, k, l\rangle. \quad (E.22)$$

Podemos calcular la función generatriz de nuestro sistema, para ello usamos (E.18) y al multiplicamos por un bra $\langle s_1, s_2, s_3|$, con lo que obtenemos

$$G(s_1, s_2, s_3, t) = \sum_{j,k,l} e^{-\left(k_1^+((\gamma-1)j+\gamma_3g^{max})+\frac{k}{\tau_2}+\frac{l}{\tau_3}\right)t} G_{j,k,l}(t) (s_1-1)^j e^{-\gamma\gamma_3 s_1} (s_2-1)^k (s_3-1)^l. \quad (E.23)$$

Con esta función resulta relativamente sencillo calcular los promedios y varianzas de cada una de las variables, lo que posteriormente facilitaría cuantificar las fluctuaciones del sistema. Primero calculáremos el promedio de cada una de estas variables:

$$\begin{aligned}\langle g\rangle &= \left(e^{-\left(k_1^+(\gamma-1)\right)t} G_{1,0,0}(t) - \gamma\gamma_3 G_{0,0,0}(t) \right) e^{-\gamma\gamma_3} e^{-\left(k_1^+\gamma_3g^{max}\right)t}, \\ \langle m\rangle &= e^{-\left(k_1^+\gamma_3g^{max}+\frac{1}{\tau_2}\right)t} G_{0,1,0}(t) e^{-\gamma\gamma_3}, \\ \langle p\rangle &= e^{-\left(k_1^+\gamma_3g^{max}+\frac{1}{\tau_3}\right)t} G_{0,0,1}(t) e^{-\gamma\gamma_3},\end{aligned}\quad (E.24)$$

mientras que para calcular la varianza primero necesitaremos lo siguiente,

$$\begin{aligned}\langle g^2\rangle &= \left(2e^{-2\left(k_1^+(\gamma-1)\right)t} G_{2,0,0}(t) + 2\gamma\gamma_3 e^{-\left(k_1^+(\gamma-1)\right)t} G_{1,0,0}(t) + \gamma\gamma_3(\gamma\gamma_3-1) G_{0,0,0}(t) \right) e^{-\gamma\gamma_3} e^{-\left(k_1^+\gamma_3g^{max}\right)t}, \\ \langle m^2\rangle &= \left(2e^{-\left(\frac{1}{\tau_2}\right)t} G_{0,2,0}(t) + G_{0,1,0}(t) \right) e^{-\gamma\gamma_3} e^{-\left(k_1^+\gamma_3g^{max}+\frac{1}{\tau_2}\right)t}, \\ \langle p^2\rangle &= \left(2e^{-\left(\frac{1}{\tau_3}\right)t} G_{0,0,2}(t) + G_{0,0,1}(t) \right) e^{-\gamma\gamma_3} e^{-\left(k_1^+\gamma_3g^{max}+\frac{1}{\tau_3}\right)t},\end{aligned}\quad (E.25)$$

por lo que ahora la varianza de cada una de las variables será $\sigma^2 = \langle x^2\rangle - \langle x\rangle^2$, pero como nos interesa las fluctuaciones, de inmediato las calculamos como $\eta^2 = \frac{\langle x^2\rangle}{\langle x\rangle^2} - 1$

$$\begin{aligned}
\eta_g^2 &= \frac{\left(2e^{-2(k_1^+(\gamma-1))t}G_{2,0,0}(t) + 2\gamma\gamma_3e^{-(k_1^+(\gamma-1))t}G_{1,0,0}(t) + \gamma\gamma_3(\gamma\gamma_3 - 1)G_{0,0,0}(t)\right)}{\left(e^{-(k_1^+(\gamma-1))t}G_{1,0,0}(t) - \gamma\gamma_3G_{0,0,0}(t)\right)^2 e^{-\gamma\gamma_3}e^{-(k_1^+\gamma_3g^{max})t}} - 1, \\
\eta_m^2 &= \frac{\left(2e^{-\left(\frac{1}{\tau_2}\right)t}G_{0,2,0}(t) + G_{0,1,0}(t)\right)}{e^{-(k_1^+\gamma_3g^{max} + \frac{1}{\tau_2})t}G_{0,1,0}^2(t)e^{-\gamma\gamma_3}} - 1, \\
\eta_p^2 &= \frac{\left(2e^{-\left(\frac{1}{\tau_2}\right)t}G_{0,0,2}(t) + G_{0,0,1}(t)\right)}{e^{-(k_1^+\gamma_3g^{max} + \frac{1}{\tau_2})t}G_{0,0,1}^2(t)e^{-\gamma\gamma_3}} - 1,
\end{aligned} \tag{E.26}$$

estas relaciones muestran las fluctuaciones sobre cada una de las variables que están presentes en el sistema, si uno quisiera conocer las fluctuaciones de todo el sistema sumaría las tres cantidades anteriores y luego sacaría la raíz cuadrada.

E.1. Cálculos para el módulo de transcripción

A continuación, detallaremos los cálculos para $\langle j|g\rangle$, $\langle k|m\rangle$ y $\langle l|p\rangle$, de igual manera para $\langle g|j\rangle$, $\langle m|k\rangle$ y $\langle p|0\rangle$, dos de ellos son parecidos por los cálculos se reducen, primero nos enfocaremos en $\langle j|g\rangle$, para ello partimos de las siguientes relaciones

$$b_1^+ |j\rangle = |j+1\rangle, \quad b_1^- |j\rangle = j |j-1\rangle, \tag{E.27}$$

de estos resultados podemos obtener lo siguiente usando estos resultados se tiene

$$|j\rangle = (b_1^+)^j |0\rangle, \quad (b_1^-) |0\rangle = 0, \tag{E.28}$$

si recordamos lo que se hace en mecánica cuántica con los operadores escalera, existe la siguiente correspondencia

$$a^+ \rightarrow s, \quad a^- \rightarrow \partial_s,$$

por lo que nuestros operadores serán ($\gamma_3 = \frac{1}{1-\gamma}$)

$$b_1^- \rightarrow (\partial_s + \gamma\gamma_3), \quad b_1^+ \rightarrow (s + g^{max}\gamma_3), \tag{E.29}$$

por lo que ahora suponemos que existe una base en donde la proyección es tal como en las relaciones anteriores, se puede mostrar que esta corresponde a la base $|s\rangle$, entonces

$$\langle s|j\rangle = \langle s|(b_1^+)^j |0\rangle = (s + g^{max}\gamma_3)^j \langle s|0\rangle, \tag{E.30}$$

para encontrar el $\langle s|0\rangle$, aplicaremos b_1^- como a continuación

$$\langle s|b_1^- |0\rangle = 0 = (\partial_s + \gamma\gamma_3) \langle s|0\rangle, \tag{E.31}$$

Esta es una ecuación diferencial tiene como solución $\langle s|0\rangle = Ne^{-\gamma\gamma_3 s}$, el factor N es una constante de normalización, con lo que finalmente tenemos

$$\langle s|j\rangle = (s + g^{max}\gamma_3)^j Ne^{-\gamma\gamma_3 s}, \tag{E.32}$$

ahora calculamos $\langle c|s\rangle$,

$$\delta_{j,j'} = \langle j|j'\rangle = \oint \frac{ds}{2\pi i} \langle j|s\rangle \langle s|j'\rangle = \oint \frac{ds}{2\pi i} \langle j|s\rangle (s + g^{max}\gamma_3)^j Ne^{-\gamma\gamma_3 s}, \tag{E.33}$$

para que se cumpla el teorema integral de Cauchy se tiene que elegir

$$\langle j|s \rangle = \frac{1}{N} \frac{e^{\gamma\gamma_3 s}}{(s + g^{max}\gamma_3)^{j+1}}, \quad (\text{E.34})$$

calculamos ahora lo siguiente $\langle j|g \rangle$ y $\langle g|j \rangle$ con ayuda de las relaciones (E.4), (E.32) y (E.34)

$$\langle j|g \rangle = \oint \frac{ds}{2\pi i} \langle j|s \rangle \langle s|g \rangle = \oint \frac{ds}{2\pi i} \frac{1}{N} \frac{e^{\gamma\gamma_3 s}}{(s + g^{max}\gamma_3)^{j+1}} s^g = \oint \frac{ds'}{2\pi i} \frac{1}{N} \frac{e^{\gamma\gamma_3(s' - g^{max}\gamma_3)}}{(s')^{j+1}} (s' - g^{max}\gamma_3)^g, \quad (\text{E.35})$$

aplicando el teorema integral de Cauchy

$$\begin{aligned} \langle j|g \rangle &= \frac{1}{N} \frac{1}{j!} \partial_{s'}^j \left[e^{\gamma\gamma_3(s' - g^{max}\gamma_3)} (s' - g^{max}\gamma_3)^g \right]_{s'=0} \\ &= \frac{1}{N} \frac{1}{j!} e^{-\gamma\gamma_3(g^{max}\gamma_3)} \sum_{l=0}^j \frac{j!}{(j-l)!l!} \partial_{s'}^{j-l} \left[e^{\gamma\gamma_3 s'} \right]_{s'=0} \partial_{s'}^l \left[(s' - g^{max}\gamma_3)^g \right]_{s'=0} \\ &= \frac{1}{N} \frac{1}{j!} e^{-\gamma\gamma_3(g^{max}\gamma_3)} \sum_{l=0}^j \frac{j!}{(j-l)!l!} (\gamma\gamma_3)^{j-l} \frac{g!}{(g-l)!} \theta(g-l) (-g^{max}\gamma_3)^{g-l}. \end{aligned} \quad (\text{E.36})$$

El término del sumatorio puede ser evaluado usando algún sistema computacional (en este caso use Mathematica) obteniendo

$$\langle j|g \rangle = e^{-\gamma\gamma_3(g^{max}\gamma_3)} (\gamma\gamma_3)^j (-\gamma_3 g^{max})^g A_1(j, g, z), \quad (\text{E.37})$$

la z se define como $z = \min\{g, j\}$, además la función $A_1(j, g, z)$ está definida como

$$A_1(j, g, z) = \frac{(-\gamma\gamma_3^2 g^{max})^{-z} {}_2F_2(1, -z; g-z+1, j-z+1; \gamma\gamma_3^2 g^{max})}{z!(g-z)!(j-z)!} = \frac{(-\gamma\gamma_3^2 g^{max})^{-z} {}_1F_1(-z; |g-j|+1; \gamma\gamma_3^2 g^{max})}{z!(|g-j|)!}. \quad (\text{E.38})$$

Ahora calculamos el siguiente termino

$$\langle g|j \rangle = \oint \frac{ds}{2\pi i} \langle g|s \rangle \langle s|j \rangle = \oint \frac{ds}{2\pi i} \frac{(s + g^{max}\gamma_3)^j N e^{-\gamma\gamma_3 s}}{s^{g+1}}, \quad (\text{E.39})$$

aplicando el teorema integral de Cauchy

$$\begin{aligned} \langle g|j \rangle &= \frac{N}{g!} \partial_s^g \left[(s + g^{max}\gamma_3)^j e^{-\gamma\gamma_3 s} \right]_{s=0} \\ &= \frac{N}{g!} \sum_{l=0}^g \frac{g!}{(g-l)!l!} \partial_s^{g-l} \left[e^{-\gamma\gamma_3 s} \right]_{s=0} \partial_s^l \left[(s + g^{max}\gamma_3)^j \right]_{s=0} \\ &= N \sum_{l=0}^g \frac{1}{(g-l)!l!} (-\gamma\gamma_3)^{g-l} \frac{j!}{(j-l)!} (g^{max}\gamma_3)^{j-l} \theta(g-l), \end{aligned} \quad (\text{E.40})$$

reacomodamos el término dentro del sumatorio y volvemos a usar Mathematica con lo con lo que finalmente tenemos

$$\langle g|j \rangle = j! (-\gamma\gamma_3)^g (\gamma_3 g^{max})^j A_1(j, g, z_1). \quad (\text{E.41})$$

Procedemos de manera similar para calcular los brakets relacionados con b_2^- (válido también para b_3^-)

$$b_2^+ |c\rangle = |k+1\rangle, \quad b_2^- |c\rangle = k |k-1\rangle, \quad (\text{E.42})$$

de estos resultados podemos obtener lo siguiente usando estos resultados se tiene

$$|k\rangle = (b_2^+)^k |0\rangle, \quad (b_2^-) |0\rangle = 0, \quad (\text{E.43})$$

si recordamos lo que se hace en mecánica cuántica con los operadores escalera, existe la siguiente correspondencia

$$a^+ \rightarrow s, \quad a^- \rightarrow \partial_s,$$

por lo que los otros operadores serán

$$b_2^- \rightarrow \partial_s, \quad b_2^+ \rightarrow (s-1), \quad (\text{E.44})$$

por lo que ahora suponemos que existe una base en donde la proyección es tal como en las relaciones anteriores, se puede mostrar que esta corresponde a la base $|s\rangle$, entonces

$$\langle s|k\rangle = \langle s|(b_2^+)^k|0\rangle = (s-1)^k \langle s|0\rangle, \quad (\text{E.45})$$

para encontrar el $\langle s|0\rangle$, aplicaremos b_2^- como a continuación

$$\langle s|b_2^-|0\rangle = 0 = \partial_s \langle s|0\rangle, \quad (\text{E.46})$$

Esta es una ecuación diferencial tiene como solución $\langle s|0\rangle = N$, el factor N es una constante de normalización, con lo que finamente tenemos

$$\langle s|k\rangle = (s-1)^k N, \quad (\text{E.47})$$

ahora calculamos $\langle k|s\rangle$,

$$\delta_{k,k'} = \langle k|k'\rangle = \oint \frac{ds}{2\pi i} \langle k|s\rangle \langle s|k'\rangle = \oint \frac{ds}{2\pi i} \langle k|s\rangle (s-1)^k N, \quad (\text{E.48})$$

para que se cumpla el teorema integral de Cauchy se tiene que elegir

$$\langle c|s\rangle = \frac{1}{N} \frac{1}{(s-1)^{k+1}}, \quad (\text{E.49})$$

calculamos ahora lo siguiente $\langle k|m\rangle$ y $\langle m|k\rangle$ con ayuda de las relaciones (E.4), (E.47) y (E.49)

$$\langle k|m\rangle = \oint \frac{ds}{2\pi i} \langle k|s\rangle \langle s|m\rangle = \oint \frac{ds}{2\pi i} \frac{1}{N} \frac{1}{(s-1)^{k+1}} s^m = \oint \frac{ds'}{2\pi i} \frac{1}{N} \frac{1}{(s')^{k+1}} (s'+1)^m, \quad (\text{E.50})$$

aplicando el teorema integral de Cauchy

$$\langle k|m\rangle = \frac{1}{N} \frac{1}{k!} \partial_{s'}^k [(s'+1)^m]_{s'=0} = \frac{1}{N} \frac{m!}{k!(m-k)!} \theta(m-k) \quad (\text{E.51})$$

Ahora calculamos el siguiente termino

$$\langle m|k\rangle = \oint \frac{ds}{2\pi i} \langle m|s\rangle \langle s|k\rangle = \oint \frac{ds}{2\pi i} \frac{(s-1)^k N}{s^{m+1}}, \quad (\text{E.52})$$

aplicando el teorema integral de Cauchy

$$\langle m|k\rangle = \frac{N}{m!} \partial_s^m [(s-1)^k]_{s=0} = \frac{N}{m!} \frac{k!}{(k-m)!} (-1)^{k-m} \theta(k-m). \quad (\text{E.53})$$

Bibliografía

- [1] Sayama, H. (2015). Introduction to the modeling and analysis of complex systems. Open SUNY Textbooks.
- [2] Roces, M. E. Á. B., Martínez-García, J. C., Dávila-Velderrain, J., Domínguez-Hüttinger, E., & Martínez-Sánchez, M. E. (2018). Modeling Methods for Medical Systems Biology.
- [3] Del Muñoz-Miranda, L. A., Higuera-Ciapara, I., Gschaedler-Mathis, A. C., Rodríguez-Zapata, L. C., Pereira-Santana, A., & Figueroa-Yáñez, L. J. (2019). Breve Descripción de la Biología Sintética y la Importancia de su Relación con otras Disciplinas. *Revista mexicana de ingeniería biomédica*, 40(1).
- [4] Alon, U. (2019). An introduction to systems biology: design principles of biological circuits. CRC press.
- [5] Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335-338.
- [6] Johan, K., & Murray, R. (2009). Feedback systems. An introduction for scientist and engineers.
- [7] De Jong, H. (2002). Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology*, 9(1), 67-103.
- [8] Iglesias, P. A., & Ingalls, B. P. (Eds.). (2010). Control theory and systems biology. MIT press.
- [9] Del Vecchio, D., Ninfa, A. J., & Sontag, E. D. (2008). Modular cell biology: retroactivity and insulation. *Molecular systems biology*, 4(1), 161.
- [10] Érdi, P., & Lente, G. (2016). Stochastic chemical kinetics (Vol. 1). Springer-Verlag New York.
- [11] Walczak, A. M., Mugler, A., & Wiggins, C. H. (2012). Analytic methods for modeling stochastic regulatory networks. *Computational Modeling of Signaling Networks*, 273-322.k.
- [12] Oxford University Press (OUP). (2021). Módulo. Lexico.com. <https://www.lexico.com/es/definicion/modulo>
- [13] Evans, L. C. (1983). An introduction to mathematical optimal control theory version 0.2. Lecture notes available at <http://math.berkeley.edu/~evans/control.course.pdf>.
- [14] Brunton, S. L., Nathan Kutz, J., Manohar, K., Aravkin, A. Y., Morgansen, K., Klemisch, J., ... & McDonald, D. (2021). Data-driven aerospace engineering: reframing the industry with machine learning. *AIAA Journal*, 1-26.
- [15] Del Vecchio, D., & Murray, R. M. (2014). Biomolecular feedback systems. In *Biomolecular Feedback Systems*. Princeton University Press.
- [16] Del Vecchio, D. (2007, July). Design and analysis of an activator-repressor clock in *E. coli*. In 2007 american control conference (pp. 1589-1594). IEEE.

- [17] Celis, J. E., Madsen, P., & Ryazanov, A. G. (1990). Increased phosphorylation of elongation factor 2 during mitosis in transformed human amnion cells correlates with a decreased rate of protein synthesis. *Proceedings of the National Academy of Sciences*, 87(11), 4231-4235.
- [18] Gardiner, C. (2009). *Stochastic methods* (Vol. 4). Berlin: Springer.
- [19] Anderson, D. F., Craciun, G., & Kurtz, T. G. (2010). Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin of mathematical biology*, 72(8), 1947-1970.
- [20] Hong, H., Kim, J., Ali Al-Radhawi, M., Sontag, E. D., & Kim, J. K. (2021). Derivation of stationary distributions of biochemical reaction networks via structure transformation. *Communications biology*, 4(1), 1-10.
- [21] Sakurai, J. J., & Napolitano, J. (2014). *Modern Quantum Mechanics*. 2-nd edition. Person New International edition.
- [22] Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4), 403-434.
- [23] Lecca, P. (2013). Stochastic chemical kinetics. *Biophysical reviews*, 5(4), 323-345.
- [24] Scott, M. (2013). *Applied stochastic processes*. Lecture Notes.
- [25] Gomez-Uribe, C. A., & Verghese, G. C. (2007). Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *The Journal of chemical physics*, 126(2), 024109.
- [26] Jeevanjee, N. (2011). *An introduction to tensors and group theory for physicists*. Birkhäuser.
- [27] Raj, A., & Van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2), 216-226.
- [28] Thattai, M., & Van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15), 8614-8619.
- [29] Eldar, A., & Elowitz, M. B. (2010). Functional roles for noise in genetic circuits. *Nature*, 467(7312), 167-173.
- [30] Paulsson, J. (2005). Models of stochastic gene expression. *Physics of life reviews*, 2(2), 157-175.
- [31] Swain, P. S., Elowitz, M. B., & Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20), 12795-12800.
- [32] Thomas, P., Straube, A. V., & Grima, R. (2012). The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC systems biology*, 6(1), 1-23.
- [33] Walls, D. F., & Milburn, G. J. (2007). *Quantum optics*. Springer Science & Business Media.
- [34] Gómez-Uribe, C. A., Verghese, G. C., & Tzafiriri, A. R. (2008). Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics. *The Journal of chemical physics*, 129(24), 244112.
- [35] Santillán, M. (2014). *Chemical kinetics, stochastic processes, and irreversible thermodynamics* (Vol. 2014). Heidelberg: Springer.

- [36] Alleva, K., & Federico, L. (2013). Análisis estructuralista de las teorías de Hill: una elucidación de explicación en bioquímica. *Scientiae Studia*, 11, 333-353.
- [37] Weiss, J. N. (1997). The Hill equation revisited: uses and misuses. *The FASEB Journal*, 11(11), 835-841.
- [38] Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2), 167-256.
- [39] Loinger, A., & Biham, O. (2007). Stochastic simulations of the repressilator circuit. *Physical Review E*, 76(5), 051917.
- [40] Loinger, A., Lipshtat, A., Balaban, N. Q., & Biham, O. (2007). Stochastic simulations of genetic switch systems. *Physical Review E*, 75(2), 021904.