



Universidad Nacional de Rosario (UNR)
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Centro Internacional Franco Argentino de Ciencias de la
Información y Sistemas (CIFASIS) (CONICET-UNR)
Grupo de Ingeniería de Sistemas de Procesos (GISP)

Tesis Doctoral

Operación Óptima de Procesos Industriales. Selección de Estructuras de Control Según Criterios Económicos.

Bottari, Agustín

Director: Dr. Marchetti, Alejandro G.
Co-Director: Dr. Zumoffen, David A.R.
Miembros del Jurado: Dr. Bandoni, Alberto J.
Dr. Kofman, Ernesto J.
Dr. Martínez, Ernesto C.

Tesis presentada en la Facultad de Ciencias Exactas, Ingeniería y Agrimensura, en
cumplimiento parcial de los requisitos para optar al título de

Doctor en Ingeniería

16 de Diciembre de 2021

Certifico que el trabajo incluido en esta tesis es el resultado de tareas de investigación originales y que no ha sido presentado para optar a un título de posgrado en ninguna otra universidad o institución.

A handwritten signature in black ink, appearing to read 'Agustín Bottari', with a horizontal line underneath the name.

Agustín Bottari

GISP

Grupo de Ingeniería de Sistemas de Procesos.
CIFASIS (UNR-CONICET).
Blvd. 27 de Febrero 210 Bis.
2000 Rosario, Argentina.
TE: +54 341 4815569 Int. 300.
E-mail: bottari@cifasis-conicet.gov.ar
Año: 2021.

Agradecimientos

Gracias,

a mi familia: mi viejo, mis hermanos y hermanas por acompañarme, a mi vieja por ser modelo de perseverancia y dedicación. A mi tía y mi prima, no habría recorrido este camino sin ellas. A mi amiga, compañera y amor Sofía, que no encuentro palabras para abarcar todo lo que le debo. A Genaro el regalo más hermoso que podía darme la vida,

a los colegas del CIFASIS con quienes compartí el camino,

a Alejandro y David por su dedicación en la dirección del trabajo,

a quienes luchan por hacer del sistema científico nacional un lugar de trabajo más humano,

a Patricio por todos los momentos compartidos, porque de solo recordarte sonriente seguís alegrándome el día.

Resumen

La operación óptima de los procesos industriales multivariados sujetos a restricciones puede abordarse mediante la aplicación de sistemas de control. Existen diversos criterios y objetivos a perseguir al momento de diseñar el sistema de control. Es común que la atención de los ingenieros se centre en diseñar el nivel de control regulador en base a índices de estabilidad y controlabilidad, sin embargo, al hacer esto se deja de lado la economía de la operación como así también la posible existencia de un nivel de control supervisor. En esta tesis se aborda la operación óptima de procesos industriales en base a criterios económicos de estado estacionario. Se proponen distintas estrategias para la selección de las variables manipuladas, las variables controladas, el emparejamiento entre ellas y los valores de consigna adoptados para procesos multivariados con restricciones y sujetos a perturbaciones externas. Las diversas estrategias de selección de estructuras de control se formalizan mediante las herramientas de programación matemática.

Una primera propuesta para la selección de estructuras de control se logra luego de estudiar y comparar los métodos existentes basados en márgenes de seguridad y en control auto optimizante. A partir de la revisión de la bibliografía se logra encontrar los puntos en común y formular un problema del tipo mixto entero no lineal que nos permite encontrar la estructura de control óptima en base a criterios económicos de estado estacionario. Se demuestra que dicha formulación encuadra dentro de los métodos de márgenes de seguridad como así también dentro de los métodos de control auto optimizante.

Una segunda propuesta busca dar respuesta al diseño de sistemas de control multinivel. La idea que se persigue es incorporar un nivel de control supervisor capaz de producir cambios en la estructura de control en base a las mediciones de planta disponibles. Para ello se estudian distintas alternativas que van desde un optimizador en línea que modifica los valores de consigna hasta un sistema de control supervisor que divide la región operativa y realiza cambios en la estructura de control según se detecten variaciones en las condiciones operativas. La tesis muestra la importancia de resolver el diseño del sistema de control multinivel de manera integral, es decir, seleccionar el control regulador mediante un problema de optimización que contemple la acción del control supervisor.

En la tesis se arriba a una formulación generalizadora que nos permite buscar, analizar y seleccionar la estructura de control óptima en base a criterios económicos de estado estacionario para el proceso industrial en estudio. Dicha formulación provee una herramienta útil que nos permite diseñar y planificar potenciales mejoras del sistema de control.

Finalmente, la tesis se centra en el optimizador en tiempo real presente en el nivel supervisor del sistema de control y afronta la problemática de disponer un modelo deficiente de la planta. En los denominados métodos de optimización en tiempo real por adaptación de modificadores, este inconveniente se aborda iterativamente a través de la actualización y adaptación del modelo original seguido de una optimización. La adaptación del modelo es posible gracias a la información real de la planta obtenida mediante instrumentos de medición. Con los datos recolectados, se procede a comparar las diferencias existentes entre la planta y lo estimado por el modelo. De la diferencia planta versus

modelo surgen las modificaciones a realizar a la función objetivo y a las restricciones del modelo. Para un punto operativo, la diferencia entre la función costo de la planta y del modelo dan lugar al denominado modificador de orden cero. La diferencia entre el gradiente de la función costo de la planta y el gradiente de la función costo del modelo dan lugar a los modificadores de primer orden. De la misma manera se definen también los modificadores de las restricciones. La convergencia de los esquemas de adaptación por modificadores hacia el óptimo real de la planta puede verse limitado al utilizar solamente los denominados modificadores de primer orden (corrección de los gradientes), es por eso que en esta tesis se indaga en los métodos que calculan modificadores de hasta segundo orden (corrección de los hessianos) de la función costo y de las restricciones del problema de optimización. El uso de los modificadores de segundo orden busca sortear las limitaciones de convergencia que se pueden presentar en los esquemas de primer orden, sin embargo, tanto los esquemas de adaptación por modificadores de primer orden como de segundo orden comparten una dificultad adicional al problema de la convergencia: el cálculo de los gradientes y hessianos relativos a la planta real. En la tesis se propone un método de cálculo de modificadores a partir de la construcción de una aproximación cuadrática de la diferencia entre la planta y el modelo basada en puntos operativos pasados. La propuesta de cálculo es acompañada de un algoritmo que nos permite incorporar los modificadores en un esquema iterativo que converja a la cercanía del óptimo real de la planta.

Índice general

Resumen	VII
1. Introducción	1
1.1. Antecedentes y Motivación	1
1.2. Estado del Arte	3
1.2.1. Operación de una Planta en Estado Estacionario	4
1.2.2. Selección de Estructura de Control en Base a Criterios Económicos	5
1.2.2.1. Método de los Márgenes de Seguridad	6
1.2.2.2. Métodos de Control Auto-Optimizante	8
1.2.2.3. Control de Restricciones Activas	9
1.2.3. Optimización en Tiempo Real	10
1.2.3.1. Método Clásico de Dos Etapas	10
1.2.3.2. Método de Adaptación por Modificadores	10
1.3. Organización y Estructura de la Tesis	11
1.4. Principales Aportes y Publicaciones	11
2. Conceptos Preliminares y Formulaciones Matemáticas	13
2.1. Punto Operativo Óptimo en Estado Estacionario	13
2.1.1. Formulación del Problema de Optimización	13
2.1.2. Condiciones de Karush-Kuhn-Tucker	15
2.2. Problema de Selección de Estructura de Control	16
2.3. Problema de los Márgenes de Seguridad	17
2.3.1. Interpretación Conceptual del Problema de los Márgenes de Seguridad	17
2.3.2. Formulación Matemática del Método de los Márgenes de Seguridad	18
2.4. Problema de Control Auto-Optimizante	20
2.4.1. Pérdida Económica Promedio	20
2.4.2. Minimización del Costo Económico Promedio	20
2.4.3. Estructuras de Combinaciones Lineales	21
2.4.4. Método del Espacio Nulo	22
2.5. Problema de Optimización en Tiempo Real	23
2.5.1. Método Clásico de Dos Etapas	23
2.5.2. Método de Adaptación por Modificadores	24
3. Selección de Estructuras de Control Auto Optimizantes con Márgenes de Seguridad de Estado Estacionario	27
3.1. Introducción	27
3.1.1. Contribuciones de este Capítulo	27
3.2. Selección de la Mejor Estructura de Control Clásica	29
3.2.1. Formulación MINLP	29

3.2.2.	Restricciones Adicionales	30
3.2.3.	Estrategia de Solución	33
3.3.	Selección de Estructuras de Control con Combinaciones Lineales	34
3.3.1.	Restricciones Adicionales	35
3.4.	Aproximación como un Problema MIQP	35
3.4.1.	MIQP Global en Perturbaciones	35
3.4.2.	MIQP Local en Perturbaciones	37
3.4.3.	Calidad de la Aproximación y Aspectos a Tener en Cuenta en la Implementación	37
3.5.	Simulaciones y Resultados	38
3.5.1.	Sistema Lineal	38
3.5.1.1.	Selección de Estructura de Control para el Problema A	39
3.5.1.2.	Selección de Estructura de Control para el Problema B	40
3.5.1.3.	Estructura de Control de Combinaciones Lineales para el Problema B	41
3.5.1.4.	Conclusiones del Caso del Estudio 3.5.1	42
3.5.2.	Reactor Tanque Agitado Continuo	43
3.5.2.1.	Estructura de Control Clásica para el CSTR	44
3.5.2.2.	Estrategias de Solución Alternativas	46
3.5.2.3.	Estructuras de Control de Combinaciones Lineales para el CSTR	46
3.5.2.4.	Conclusiones del Caso de Estudio 3.5.2	47
3.5.3.	Evaporador	48
3.5.3.1.	Estructuras de Control Clásicas para el Evaporador	50
3.5.3.2.	Estrategias de Solución Alternativas para el Evaporador	50
3.5.3.3.	Estructura de Control de Combinaciones Lineales para el Evaporador	51
3.5.3.4.	Conclusiones del Caso de Estudio 3.5.3	51
3.6.	Conclusiones del Capítulo	52
4.	Selección de Estructuras de Control para Sistemas de Optimización en Tiempo Real	55
4.1.	Introducción	55
4.1.1.	Optimización en Tiempo Real en una Sistema de Control de Dos Capas	56
4.1.2.	Contribuciones de este Capítulo	57
4.2.	Selección de la Estructura de Control para un Sistema de Optimización en Tiempo Real	57
4.3.	Selección de Estructuras de Control Basada en Regiones de Restricciones Activas	60
4.3.1.	Regiones de Restricciones Activas	60
4.3.2.	Estructura de Control con Valores de Consigna Variables	61
4.3.3.	Estructura de Control Variable con Valores de Consigna Variables	62
4.4.	Formulación General para la Selección de Estructuras de Control en Sistemas con Optimización en Tiempo Real	63
4.4.1.	Estrategia de Solución del Problema General de Selección de Estructuras de Control con Optimización en Tiempo Real	64
4.5.	Aproximación como un Problema MIQP	67
4.6.	Consideraciones Adicionales	68
4.6.1.	Límites en la Matriz de Ganancias Relativas	68
4.6.2.	Múltiples Soluciones	68
4.6.3.	Control Indirecto de Restricciones Activas	68
4.6.4.	Identificación de las Regiones de Restricciones Activas	68
4.6.5.	Control de Restricciones Activas	69
4.6.6.	Dificultad Dimensional	70
4.6.7.	Estructuras de Control de Combinaciones Lineales de Variables	70
4.7.	Resultados de Simulación	70

4.7.1. Sistema Lineal	70
4.7.1.1. Solución Utilizando el Método de los Márgenes de Seguridad con Costo Promedio	72
4.7.1.2. Solución utilizando el Método de los Márgenes de Seguridad con Costo Promedio y Cambios en los Valores de Consigna	73
4.7.1.3. Comparación de los Métodos de los Márgenes de Seguridad	74
4.7.2. Evaporador	74
4.7.2.1. Regiones de Restricciones Activas para el Evaporador	76
4.7.2.2. Comparación de las Estrategias de Selección de Estructuras de Control	77
4.7.3. Sistema de Reacción-Separación con Reciclo	78
4.7.3.1. Descripción del Problema	78
4.7.3.2. Regiones de Restricciones Activas para el Sistema de Reacción-Separación con Reciclo	81
4.7.3.3. Estrategias de Solución para el Sistema de Reacción-Separación con Reciclo	81
4.8. Discusiones y Conclusiones del Capítulo	84
5. Optimización en Línea para Modelos Imperfectos	85
5.1. Introducción	85
5.1.1. Contribuciones de este Capítulo	86
5.2. Problema de Optimización en Línea por Adaptación de Modificadores para Modelos Imperfectos	87
5.3. Adaptación por Modificadores de Segundo Orden Mediante Interpolación Cuadrática	88
5.3.1. Cálculo de los Modificadores	88
5.3.2. Selección de Puntos	90
5.3.3. Algoritmo para Adaptación por Modificadores de Segundo Orden Mediante Interpolación Cuadrática	91
5.3.4. Aspectos de la implementación	92
5.4. Resultados de Simulación	93
5.4.1. Problema Sin Restricciones para el Reactor de Williams-Otto	94
5.4.2. Problema Con Restricciones para el Reactor de Williams-Otto	96
5.4.3. Problema Alternativo con Restricciones para el Reactor Williams-Otto	96
5.5. Discusiones y Conclusiones del Capítulo	100
6. Conclusiones Generales y Trabajos Futuros	103
6.1. Conclusiones Generales	103
6.2. Trabajos futuros	106
A. Modelos de Casos de Estudio	107
A.1. Modelo de Reactor Tanque Continuo Agitado (CSTR)	107
A.2. Modelo de Evaporador	109
A.3. Modelo Sistema de Reacción-Separación con Reciclo	111
A.4. Modelo de Reactor Tanque Continuo Agitado (CSTR) usado en Adaptación Por Modificadores	115
Bibliografía	117

Capítulo 1

Introducción

1.1 Antecedentes y Motivación

Los procesos industriales tienen como objetivo la producción de bienes y servicios que están sujetos a requisitos implícitos o explícitos. En un mundo globalizado y cada vez más competitivo, lograr reducir los costos operativos al tiempo que se mantienen las condiciones de calidad y seguridad preestablecidas es de sumo interés. Históricamente, las sugerencias del equipo de diseño, la experiencia en planta, y la prueba y error han sido los responsables de indicar el punto operativo de la planta. Con el advenimiento de modelos matemáticos más precisos, las técnicas de optimización matemática y el avance de los controladores, hace ya varias décadas que la heurística fue haciendo lugar a los desarrollos científicos y tecnológicos en el área. Sin embargo, la tarea no es sencilla ya que los procesos presentes en la industria, y sus respectivos modelos, son en general no lineales, poseen múltiples variables tanto de entrada como de salida y están sujetos a perturbaciones rápidas, lentas, conocidas y desconocidas, como así también a restricciones operativas, de seguridad y de calidad esperada. Para estos procesos multivariantes sujetos a perturbaciones, la implementación de una estructura de control (CS del inglés control structure) y la selección de los valores de consigna (SP del inglés set points), que dictaminarán el comportamiento de la planta controlada, deben elegirse en forma adecuada a fin de garantizar una operación segura y capaz de producir según los objetivos y la calidad requerida.

La selección de la estructura de control y los valores de consigna adoptados impactará en la operación del proceso y en consecuencia en la economía de este. Frecuentemente, el punto de operación óptimo del proceso se encuentra en el límite de ciertas restricciones del problema de optimización, por ejemplo, es común tender a un valor mínimo en el uso de la materia prima, mano de obra o en el consumo energético o a un valor máximo en la cantidad de productos de salida que respeten un mínimo de calidad establecido. Sin embargo, en un proceso que se encuentra operando al límite de las restricciones, el ingreso de perturbaciones externas pueden llevarlo a operar en una zona que ya no respeta las condiciones límites establecidas.

Una estrategia para evitar violar las restricciones cuando ingresa una perturbación radica en que los valores de consigna, para la estructura de control seleccionada, se determinen de modo tal que la planta opere en una zona factible aún ante la presencia de perturbaciones. Hecho esto, la operación será robusta ante el ingreso de perturbaciones y, aunque afectada por ellas, la planta seguirá operando en un punto que no violará las restricciones. Esto se conoce como el método de los márgenes de seguridad y estos valores de consigna robustos pueden calcularse para una estructura de control definida. En general, cuanto mayores sean los márgenes de seguridad a adoptar más nos alejamos del punto óptimo de operación. Seleccionar la estructura de control según el método de los márgenes de seguridad implica buscar la estructura de control que al adoptar dichos márgenes presente el mejor rendimiento

económico.

En lugar de preestablecer valores de consigna robustos, existe la alternativa de actualizar en línea los valores de consigna en tanto y en cuanto puedan estimarse los valores de las perturbaciones entrantes en el sistema. La actualización en línea de los valores de consigna es llevada a cabo por el nivel de Optimización en Tiempo Real (RTO) cuyo objetivo es obtener el punto operativo óptimo actual del proceso. La optimización en tiempo real se lleva a cabo en un nivel de control superior o supervisor. Dicho nivel se encarga de ejecutar la optimización de la planta y pasar sus resultados como valores de consigna al nivel de control inferior del tipo servo/regulador. Idealmente, el nivel de RTO cumple dos funciones importantes. La primera de ellas consiste en perseguir en tiempo real el óptimo de la planta que puede cambiar ante el efecto de las perturbaciones. La segunda función es la de actuar para corregir las consecuencias de tener como referencia un modelo de planta imperfecto.

Una tercera alternativa de interés son las técnicas desarrolladas en la bibliografía de control auto optimizante (SOC del inglés self-optimizing control). Las mismas buscan definir una estructura de control que en principio no requiera la actualización en línea de los valores de consigna pero que aun así se mantenga operando con una mínima pérdida de optimalidad económica ante el efecto de las perturbaciones. Para ello, los métodos SOC utilizan los grados de libertad del proceso en estudio de dos maneras distintas. En primer lugar se busca controlar las variables que se encuentren en los valores límites de las restricciones (restricciones activas) en el punto de operación óptimo definido. En segundo lugar, con los grados de libertad sobrantes, se diseñan lazos de control que fuerzan a la planta a perseguir el punto óptimo cuando este cambie al verse afectado por perturbaciones.

Cada una de las estrategias mencionadas encontraron sus seguidores en la academia, se han diversificado y han avanzado en resolver las dificultades que cada uno de los métodos presenta. El desafío de seleccionar y definir una estructura de control fija, el punto operativo elegido para el cálculo de los márgenes de seguridad, la existencia de modelos imperfectos, las perturbaciones que no siempre podrán medirse, la necesidad de grados de libertad sobrantes y los diseños que solo son confiables en la cercanía de un punto operativo son algunos de los retos con los que nos podremos encontrar al repasar la vasta bibliografía.

Considerando lo expuesto, y dentro de un objetivo general de desarrollar metodologías de selección de estructuras de control teniendo en cuenta criterios económicos, surgen algunas líneas de investigación promisorias que se abordaron en la presente tesis doctoral. En particular, dentro de los objetivos específicos de la tesis se encuentran los siguientes:

- Se propone examinar y comparar los métodos basados en márgenes de seguridad con los métodos de control auto optimizantes en el contexto de la selección óptima de estructuras de control en base a criterios económicos de estado estacionario.
- Se propone investigar cómo incorporar un mayor conocimiento acerca de la naturaleza de las perturbaciones presentes en el sistema para la estimación de los márgenes de seguridad. La división de las perturbaciones en perturbaciones de alta frecuencia o no medibles y perturbaciones de baja frecuencia o medibles, puede ser determinante para la selección de la estructura de control óptima en base a criterios económicos.
- Por otra parte, las estrategias de optimización en tiempo real disponibles determinan el punto de operación de estado estacionario óptimo, sin considerar los márgenes de seguridad necesarios para evitar violar las restricciones. En este contexto, un segundo objetivo de la tesis es el de aplicar en el contexto de optimización en tiempo real la técnica de selección de estructuras de control diseñada anteriormente.
- Por último, dado que es común en la práctica industrial encontrar diferencias entre el modelo disponible y la planta real, se buscará desarrollar un método de optimización en tiempo real por

adaptación de modificadores que nos permita obtener valores de consigna óptimos disponiendo de modelos de planta imperfectos.

1.2 Estado del Arte

En la industria química y petroquímica los procesos involucrados son en general no lineales, multivariantes y están sujetos a perturbaciones externas. Encontrar el punto operativo que satisface todos los requerimientos de cantidad y calidad productiva al mismo tiempo que se minimizan los costos, el consumo energético o la contaminación ambiental, sin violar ninguna restricción operativa o de seguridad, requiere sin dudas de un profundo estudio y de la adopción de una estrategia a tal fin. El diseño de estructuras de control permite dar respuestas a dichos requerimientos.

Comúnmente, los sistemas de control y optimización en las industrias de procesos continuos se dividen en distintos niveles que varían de acuerdo a los objetivos específicos que persiguen, las variables involucradas y los tiempos en que deben dar una respuesta al sistema. En la Figura 1.1 se ilustra un esquema característico. Para resolver asuntos a largo plazo (meses, días) se encuentra el nivel superior donde se realiza la planificación y programación de la producción. El resultado de este nivel son típicamente la tasa de producción, materia prima y mano de obra que se les pasará a los siguientes niveles de control como restricciones a satisfacer. En un nivel medio encontraremos un control supervisor, como por ejemplo un optimizador en tiempo real (RTO), que se encarga de encontrar el punto operativo óptimo del proceso en su estado actual (desde un punto de vista económico) y pasarlo al nivel de control regulador inferior. El optimizador en tiempo real se ejecuta con una frecuencia de horas a días. Finalmente, mediante la manipulación de los actuadores y la recolección de los datos procedentes de los instrumentos de medición, se encuentra el sistema de control regulador que busca alcanzar y mantener el proceso en los valores de referencia establecidos.

Para diseñar el sistema de control es necesario definir en primer lugar el o los objetivos que se persiguen. En base a dichos objetivos se seleccionan las variables manipuladas y las variables controladas, así como también la interconexión entre dichas variables. También es importante definir los niveles de control que estarán presentes en el sistema en base a los posibles múltiples objetivos y tecnología disponible. Por ejemplo, es muy común que se seleccione el control regulador siguiendo índices de estabilidad y controlabilidad según el criterio de error integral, el cálculo de la matriz de ganancias relativas o el número de condición [1]. Sin embargo, se deja de lado criterios económicos o la existencia de niveles superiores de control. Existen también métodos de diseño de la estructura de control apoyados en la optimización de estado estacionario o dinámico en base a criterios económicos, la satisfacción de restricciones y la existencia de perturbaciones ([2]; [3]). En esta tesis se hará hincapié en el desarrollo de estructuras de control que persiguen como objetivo la operación óptima de estado estacionario en base a criterios económicos.

El método de los márgenes de seguridad es uno de los principales métodos para la selección de la estructuras de control en base a criterios económicos. Desde un punto de vista económico el punto operativo óptimo comúnmente se encontrará en la intersección de algunas de las restricciones del proceso. Al estar sometido a perturbaciones externas, es de práctica habitual recurrir a un margen de seguridad operativo que nos preserve de violar las restricciones ante el efecto de estas perturbaciones. Dado que el margen de seguridad trae consigo una pérdida de optimalidad económica y puesto que este margen dependerá de la estructura de control instalada en la planta, se hace necesario el desarrollo de métodos que busquen seleccionar la estructura de control óptima que minimice las pérdidas económicas [4]. Dentro de los métodos de selección de estructuras de control en base a criterios económicos se han desarrollado también estrategias de control denominadas auto optimizantes. Estos métodos en principio abordaron la particular situación donde el conjunto de las restricciones activas no cambia

en presencia de las perturbaciones. En consecuencia, es óptimo desde el punto de vista económico controlar las restricciones activas en sus valores de referencia límites óptimos. Hecho esto, y si se disponen de grados de libertad adicionales luego de establecer los lazos de control mencionados, los métodos de control auto optimizantes [5] buscan seleccionar las variables de control restantes a fin de minimizar la pérdida económica ante el efecto de perturbaciones. Finalmente, otra categoría de control que persigue la operación óptima de las plantas industriales son los métodos de optimización en tiempo real. La optimización en tiempo real, que está situada en el nivel de control superior, utiliza las mediciones del proceso y actúa de manera iterativa actualizando el modelo imperfecto de la planta y optimizando el punto de referencia operativo. En [6] se propone utilizar una estrategia de adaptación por modificadores a fin de corregir el modelo y las restricciones para conducir iterativamente el punto de operación hacia el óptimo real de la planta.

El hecho de plantear niveles de control con distintos objetivos, distintas variables a disposición y distintos tiempos para dar respuestas, ha hecho que comúnmente el sistema se estudie y se diseñe de manera secuencial. Existen pocos trabajos donde, a pesar de su importancia, se diseñe de manera integral todo el sistema de control. En [7] se analizaron tres diseños para integrar la capa superior y la inferior mediante el uso de control predictivo. En [8] se propuso diseñar un sistema de optimización en tiempo real que sea capaz de no solo pasar los valores de consigna, sino también de modificar la estructura del control del nivel de control regulador inferior. En [9] se propone que la estrategia de control auto optimizante local se integre con una capa de control supervisor que guíe los cambios en la estructura según se detecten cambios en las restricciones activas.

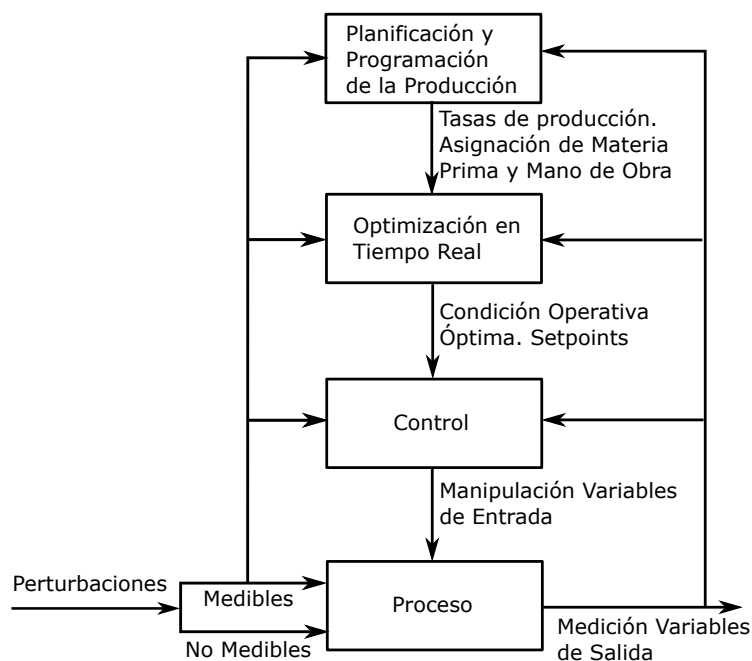


Figura 1.1: Jerarquía de niveles de decisión.

1.2.1 Operación de una Planta en Estado Estacionario

El punto de operación de estado estacionario de una planta industrial de operación continua puede estar dictado por la heurística, por el conocimiento ingenieril, la experiencia en el campo, la prueba y error, como así también puede ser determinado en base al planteo de un problema de programación matemática que busque optimizar la operación según un objetivo preestablecido. Generalmente esta última idea aplicada a procesos químicos y petroquímicos nos lleva a un programa de optimización

matemático no lineal (NLP, siglas en inglés de nonlinear program) con restricciones. Dicho problema de optimización no lineal con restricciones puede sintetizarse de la siguiente manera [10]:

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad (1.1a)$$

$$\text{s.t. } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \quad (1.1b)$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, \dots, l \quad (1.1c)$$

$$\mathbf{x} \in X, \quad (1.1d)$$

donde \mathbf{x} es el vector de n_x variables de decisión a definir. Las funciones $f, g_1, \dots, g_m, h_1, \dots, h_l$ están definidas en \mathbb{R}^{n_x} . El problema radica entonces en encontrar los valores de x^* que minimicen el funcional objetivo f mientras se satisfacen las restricciones de igualdad y desigualdad. Se dice que una restricción de desigualdad g_i se activa cuando la misma alcanza e iguala su cota.

Con frecuencia la solución óptima x^* se encuentra en la intersección de un conjunto de restricciones activas (ver Figura 1.2).

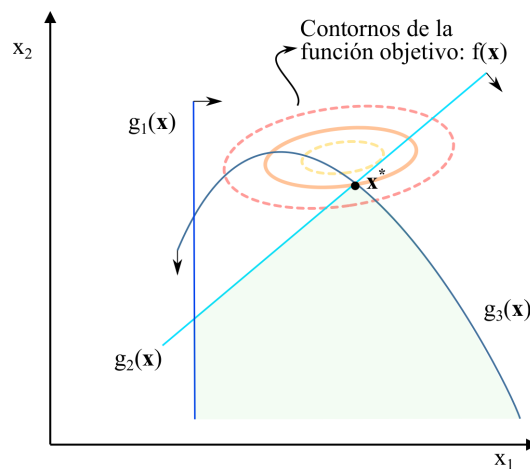


Figura 1.2: Problema de optimización con tres restricciones $g_1, g_2,$ y g_3 en el dominio bidimensional. El valor óptimo \mathbf{x}^* suele encontrarse en la intersección de un conjunto de restricciones activas.

La aplicación de la programación matemática para determinar el punto óptimo de operación de una planta industrial, en base a criterios económicos y para seleccionar la estructura de control y los valores de referencia, requiere disponer de un modelo de simulación del proceso basado en primeros principios. Dichos modelos incluyen balances de materia y de energía, ecuaciones de equilibrio líquido-vapor, cinética de reacciones químicas, etc. El modelo del proceso en estado estacionario generalmente consiste en un sistema de ecuaciones algebraicas no lineales. La formulación del problema de optimización requiere además establecer los límites operativos de las variables del proceso, restricciones que aseguren una determinada calidad del producto, restricciones de seguridad del proceso y, en nuestro enfoque, que la función objetivo se formule en relación al costo operativo.

1.2.2 Selección de Estructura de Control en Base a Criterios Económicos

En un proceso industrial nos encontraremos con sistemas que presentan múltiples variables de entrada, múltiples variables de salida y perturbaciones externas (Figura 1.3). La selección de la estructura de control de una planta consiste en definir las variables controladas (CVs), las variables manipuladas (MVs), los valores de consigna de las variables controladas (SPs, del inglés setpoints), y los lazos de control entre las variables controladas y las variables manipuladas. En los métodos de selección de estructuras de control en base a criterios económicos se busca minimizar el costo operativo

de la planta en presencia de perturbaciones sin violar las restricciones del proceso. Las dos principales metodologías existentes en la literatura para la selección de estructuras de control en base a criterios económicos son los métodos basados en el cálculo de márgenes de seguridad (backoff approaches) y los métodos de control auto optimizantes.

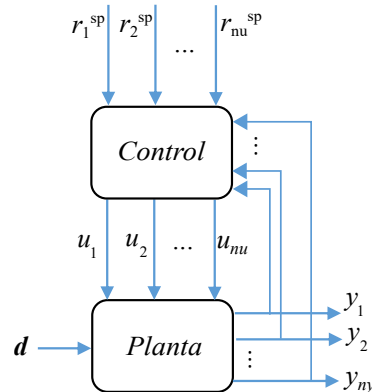


Figura 1.3: Procesos multivariables con sistema de control. Las variables u representan las variables manipulables, y las variables de salida, d las perturbaciones, y r^{sp} los valores de consigna.

1.2.2.1 Método de los Márgenes de Seguridad

La operación óptima de una planta industrial comúnmente se da en la intersección de las restricciones activas (restricciones que igualan la cota superior o inferior en el punto de operación óptima). Sin embargo, a consecuencia del efecto de las perturbaciones, en la operación del proceso será necesario alejarse de los límites de las restricciones para evitar posibles violaciones de estas.

La magnitud del retroceso que es necesario adoptar para ingresar a una zona de operación factible depende de la variabilidad de las variables restringidas para el sistema controlado a lazo cerrado. El método de los márgenes de seguridad para la selección de la estructura de control de procesos se basa en la idea de seleccionar la estructura de control que minimiza el costo económico asociado con los retrocesos en las restricciones.

Muchos de los métodos basados en márgenes de seguridad que se encuentran en la literatura se han formulado para la integración del diseño y control de los procesos. Al fijar las variables que definen el diseño y la topología de la planta a lazo abierto, estas formulaciones contemplarán el problema de selección de la estructura de control. En la Tabla 1.1 se proporciona una lista de los enfoques representativos de márgenes de seguridad que se encuentran en la literatura.

Como puede observarse los métodos de márgenes de seguridad han sido utilizados de diversas maneras. A continuación se proporciona una clasificación de los trabajos hallados en la literatura:

- *Economía de Estado Estacionario Vs. Economía Dinámica:* Un marco de estudio general consiste en formular el problema de selección de la estructura de control como un problema de optimización mixto entero dinámico (MIDO) en el que la economía dinámica de la planta controlada se optimiza para un período de tiempo determinado [1, 2, 26] (ver también Yuan et al. [27] y sus referencias). En estos enfoques, la dinámica de las perturbaciones se especifica a priori. Por otro lado, dado que las plantas que operan en forma continua están diseñadas típicamente para operar en condiciones de estado estacionario, muchos de los métodos de márgenes de seguridad evalúan la economía de la operación en estado estacionario [8, 11-25].
- *Costo Nominal Vs. Costo Promedio:* La evaluación del desempeño económico se ha incorporado utilizando diferentes criterios. Muchos autores proponen métodos de los márgenes de seguridad

Tabla 1.1: Literatura Representativa de los Métodos de Márgenes de Seguridad.

	Economía Estacionaria	Economía Dinámica	Costo Nominal	Costo Promedio	Backoff Estacionario	Backoff Dinámico	Problema Backoff	Problema Selección CS
Bahri et al. [11]	✓		✓		✓		✓	
Bahri et al. [12]	✓		✓			✓	✓	
Figueroa et al. [13]	✓		✓			✓	✓	
Narraway et al. [14]	✓		✓			✓	✓	
Loeblein y Perkins [15]	✓		✓			✓	✓	
Soliman et al. [16]	✓		✓			✓	✓	
Narraway y Perkins [17]	✓		✓			✓		✓
Heath et al. [18]	✓		✓			✓		✓
Psaltis et al. [19]	✓		✓			✓		✓
Zumoffen et al. [20]	✓		✓			✓		✓
Kookos y Perkins [21]	✓		✓			✓		✓
Kookos y Perkins [22]	✓		✓		✓			✓
Kookos y Perkins [23]	✓		✓			✓		✓
Kookos [8]	✓			✓	✓			✓
Sharifzadeh y Thornhill [24]	✓			✓	✓			✓
Sharifzadeh y Thornhill [25]	✓			✓		✓		✓
Mohideen et al. [2]		✓		✓		✓		✓
Bansal et al. [26]		✓		✓		✓		✓
Sakizlis et al. [1]		✓		✓		✓		✓
Yuan et al. [27]		✓		✓		✓		✓

que proceden minimizando el costo nominal [11–23]. En estos enfoques el costo económico se evalúa solo para los valores nominales de las perturbaciones, mientras que las variaciones de las perturbaciones con respecto a sus valores nominales se utilizan para calcular los márgenes de seguridad necesarios para no violar las restricciones de operación. Otro criterio es minimizar el costo esperado, que puede aproximarse como un costo promedio ponderado sobre un conjunto específico (o muestreado) de escenarios de perturbación [1, 2, 8, 24–26].

- *Backoff de Estado-Estacionario Vs. Backoff Dinámico*: Los métodos basados en márgenes de seguridad en estado estacionario solo se preocupan por evitar violaciones de las restricciones en estado estacionario. Los márgenes de seguridad se determinan calculando los valores de consigna (SP) óptimos de modo que la factibilidad de estado estacionario esté garantizada para todos los escenarios de perturbación [8, 22, 24]. Estos enfoques utilizan un modelo de estado estacionario del proceso y, si bien es necesario definir una estructura de control, no se requiere modelar el controlador ni ajustar sus parámetros. El único supuesto es que los controladores tienen acción integral, es decir, las variables controladas coinciden con sus valores de consigna en estado estacionario. Por otro lado, los enfoques de backoff dinámico prestan atención a las violaciones de las restricciones durante la operación transitoria (aún si la pérdida económica se evaluase para el estado estacionario) [1, 2, 12–21, 23, 25, 26]. Estos enfoques requieren modelar el comportamiento dinámico de la planta a lazo cerrado en presencia de perturbaciones. Para cada estructura de control, es necesario sintetizar analíticamente el conjunto de ecuaciones diferenciales y algebraicas de los controladores. Los parámetros del controlador pueden optimizarse o ajustarse automáticamente mediante reglas analíticas. Para evitar las complejidades planteadas por el modelado de los controladores, a menudo se ha asumido el control perfecto [17–19, 21, 25]. Aunque esta suposición podría ser útil para seleccionar la estructura de control, conduce a márgenes de seguridad dinámicos iguales a cero para las variables controladas, lo que puede conducir a violaciones de las restricciones en la práctica. Por lo tanto, cuando se asume un control perfecto, los márgenes de seguridad para la estructura de control seleccionada deben reevaluarse mediante simulaciones dinámicas después de modelar los controladores.
- *Problema de Backoff Vs. Problema de Selección de Estructura de Control*: El problema de los márgenes de seguridad se suele formular como un problema de programación no lineal (NLP), que

se resuelve para una estructura de control determinada (fija) [12–16]. El problema puede consistir en encontrar valores de consigna óptimos que garanticen la factibilidad para un conjunto dado de escenarios de perturbación, o en calcular directamente los márgenes de seguridad con respecto a la operación óptima nominal y luego evaluar la pérdida económica asociada. Basado en el enfoque de los márgenes de seguridad, el problema de selección de la estructura de control incorpora variables binarias que se utilizan para determinar la estructura de control óptima [1, 2, 8, 17–26]. El problema puede formularse como un problema MIDO o un problema de programación no lineal mixto entero (MINLP). Se han propuesto varias simplificaciones y aproximaciones para reducir la carga computacional del problema, por ejemplo, linealizar las ecuaciones del modelo dinámico o de estado estacionario y aproximar el problema MINLP por un problema mixto entero lineal (MILP) [17–19], e incluir la suposición de control perfecto [17–19, 21, 25].

1.2.2.2 Métodos de Control Auto-Optimizante

El control auto optimizante (SOC, del inglés self-optimizing control) tiene como objetivo seleccionar variables controladas y valores de consigna constantes de modo que, en presencia de perturbaciones, la planta a lazo cerrado opere en estado estacionario con una pérdida económica aceptable [5, 28, 29].

Tabla 1.2: Literatura Representativa de los Métodos de Control Auto Optimizante.

	Economía	Pérdida Económica	SOC	SOC	Restricciones Activas		Problema CS
	Estacionaria	Promedio	Local	Global	Invariantes	Variantes	Combinación Lineal
Skogestad [5]	✓	✓	✓		✓		✓
Halvorsen et al. [28]	✓		✓		✓		✓
Alstad y Skogestad [30]	✓	✓	✓		✓		✓
Kariwala et al. [31]	✓	✓	✓		✓		✓
Alstad et al. [32]	✓	✓	✓		✓		✓
Jäschke y Skogestad [33]	✓	✓	✓		✓		✓
Hu et al. [34]	✓	✓	✓			✓	✓
Manum y Skogestad [9]	✓	✓	✓			✓	✓
Yelchuru y Skogestad [35]	✓	✓	✓		✓		✓
Marchetti y Zumoffen [36]	✓	✓	✓		✓		✓
Ye et al. [37]	✓	✓		✓	✓		✓
Ye et al. [38]	✓	✓		✓		✓	✓

En la tabla 1.2 se sintetizan las características de los principales trabajos y autores de la literatura de control auto optimizante. La mayoría de las estrategias SOC propuestas en la literatura son métodos locales [29]. Los métodos SOC locales surgieron del supuesto de que el conjunto de restricciones activas no cambia con los valores de perturbación considerados. Con esta suposición, es óptimo (desde una perspectiva de estado estacionario) controlar las restricciones activas de las variables de entrada y salida en sus valores límite óptimos. Esta acción requiere asignar una serie de variables de entrada para controlar estas restricciones activas. Al hacer esto se puede ver el problema de optimización como un problema sin restricciones en el espacio reducido de las variables de entrada que queda después de satisfacer todas las restricciones activas:

$$\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d}) \quad (1.2)$$

Los métodos de control auto optimizantes locales tienen como objetivo seleccionar variables controladas adicionales de modo que estas entradas restantes (grados de libertad disponibles) alcancen valores cercanos al óptimo en estado estacionario cuando ocurren perturbaciones. Una particularidad de la mayoría de los enfoques de SOC es que estas variables controladas adicionales se seleccionan como combinaciones lineales de las variables de entrada y salida.

Basado en las condiciones necesarias de optimalidad (NCO) de primer orden para un problema de optimización sin restricciones y un modelo linealizado de las variables de salida, Halvorsen et al. [28] arribó a una expresión local para la pérdida económica en función de los valores de perturbación y el error de medición. Esta expresión de pérdida económica local se utilizó para obtener combinaciones lineales óptimas de variables medidas que minimizan la pérdida (local) en el peor escenario [28] y la pérdida promediada (local) [31], en presencia de perturbaciones. Otro método SOC local es el *método del espacio nulo* propuesto por Alstad y Skogestad [30], que consiste en seleccionar como variables controladas combinaciones lineales de variables medibles que se encuentran en el espacio nulo de la matriz de sensibilidad de las salidas óptimas con respecto a las perturbaciones. Más tarde, Alstad et al. [32] mostró que el método del espacio nulo lleva a cero la expresión de pérdida local de Halvorsen et al. [28]. Se ha establecido la equivalencia entre el método del espacio nulo y el control del gradiente [33], así como su equivalencia con el control de extremos (neighboring extremal control) [39]. En Marchetti y Zumoffen [36] se presenta una formulación del método del espacio nulo en todo el espacio de las entradas para problemas restringidos con restricciones activas. En los métodos SOC locales, [30–32, 36], los valores de consigna para las combinaciones lineales de variables medidas se seleccionan como los valores de estas combinaciones lineales evaluadas en el óptimo nominal.

Dado que en muchas aplicaciones prácticas el conjunto de restricciones activas cambia según los valores de la perturbación, se han propuesto variantes de los métodos SOC locales para manejar cambios en el conjunto de restricciones activas. Por ejemplo, Manum y Skogestad [9] emplearon un enfoque de programación paramétrica para determinar las regiones del conjunto de restricciones activas, y proponen seleccionar las variables controladas aplicando el método de espacio nulo dentro de cada región. El seguimiento de los valores que adoptan las variables a controlar de cada región se utiliza para decidir cuándo cambiar de región. En lugar de detectar el conjunto de restricciones activas y cambiar la estructura de control, Hu et al. [34] propusieron un enfoque local para encontrar una estructura de control única que garantice la viabilidad de todos los valores de perturbación dentro de límites preestablecidos. La estructura de control de combinación lineal óptima se obtiene minimizando la pérdida local promedio sujeto a restricciones linealizadas.

Recientemente, Ye et al. [37] propusieron un método de control auto optimizante con características globales que está restringido a problemas para los cuales el conjunto de restricciones activas es invariante. En lugar de depender de la pérdida local calculada únicamente en el óptimo nominal, se muestrean aleatoriamente varios escenarios de perturbación y se resuelve el problema de optimización para dichos escenarios. Para cada escenario de perturbación, se calcula la expresión de pérdida local correspondiente y se determina la combinación lineal óptima de variables medibles junto con sus correspondientes valores de consignas óptimos minimizando la pérdida económica promedio. Una variante de este enfoque, que también usa información local para cada escenario de perturbación, se extendió a problemas con restricciones que presenten cambios en los conjuntos de restricciones activas [38].

1.2.2.3 Control de Restricciones Activas

Maarleveld y Rijnsdorp [40] reconocieron hace 50 años la importancia de controlar las restricciones que están activas en sus valores límite para mantener un proceso en su punto de operación óptimo, esto es conocido como control de restricciones activas. El control de restricciones puede eliminar la necesidad de la optimización en tiempo real para casos donde las variaciones de las perturbaciones no cambian el conjunto de restricciones activas [41]. Sin embargo, el conjunto de restricciones activas del problema de optimización (2.5) puede cambiar para los diferentes valores de las perturbaciones \mathbf{d} . Cuando esto ocurre, es deseable que la estructura de control cambie al nuevo conjunto activo, lo que podría lograrse usando controladores selectivos (override control) [42]. En el contexto de la optimización en tiempo

real, también se ha argumentado que la capa RTO debe pasar a la capa de control regulatorio no solo los valores de consigna, sino también la propia estructura de control, basándose en la detección de cambios en el conjunto de restricciones activas [7, 8, 43]. Jacobsen y Skogestad [44] han estudiado la detección de regiones de restricciones activas para el diseño de estructuras de control en el contexto de la operación óptima del proceso.

1.2.3 Optimización en Tiempo Real

En el nivel de optimización en tiempo real típicamente nos encontraremos con un algoritmo diseñado para perseguir un punto operativo óptimo cambiante. La variación del punto de operación óptimo puede estar ligada a cambios provenientes del nivel de planificación y programación, al efecto de las perturbaciones externas, o a las diferencias (paramétricas o estructurales) entre el modelo utilizado y el proceso real de la planta. El sistema de optimización en tiempo real plantea un esquema a lazo cerrado de la planta real donde mediante el uso de los datos disponibles de las variables del proceso se corrige o adapta el modelo disponible para luego realizar una optimización del punto operativo en base a criterios económicos. Dos de los algoritmos comunmente encontrados en la vasta bibliografía de optimización en tiempo real son [45]: 1) el método clásico de dos etapas y 2) los esquemas de adaptación por modificadores .

1.2.3.1 Método Clásico de Dos Etapas

Con anterioridad se mencionó que el modelo que es utilizado para la optimización del punto operativo puede presentar diferencias con el proceso real de la planta. Un caso típico son las diferencias entre los valores de ciertos parámetros del modelo versus los valores reales de dichos parámetros. En los llamados métodos de dos etapas [43, 46], el algoritmo a implementar tiene como objetivo corregir el punto operativo obtenido con el modelo impreciso mediante corrección de los parámetros del mismo. Para ello el método realiza dos pasos: 1) Medición de variables de salida seleccionadas de la planta real. Estimación de los valores de los parámetros ajustables del modelo mediante la resolución de un problema de estimación de parámetros. 2) Optimización empleando el modelo con los parámetros actualizados y aplicación en la planta del punto operativo óptimo obtenido.

Si el modelo que se tiene de la planta presenta una diferencia paramétrica con la planta real, y los parámetros inciertos se pueden estimar, entonces el método de los dos pasos es adecuado y nos conducirá al óptimo real de la planta. Sin embargo, en presencia de una diferencia estructural entre el modelo y la planta el método de los dos pasos no nos garantiza la convergencia al óptimo real de la planta.

La posibilidad de no convergencia al óptimo del proceso real hizo necesario establecer el criterio para que un modelo sea adecuado para la utilización del esquema de optimización en tiempo real de dos etapas [47]. Se define como adecuado (para un esquema de dos etapas) a un modelo de proceso para el cual el óptimo de la planta es un punto fijo del algoritmo de optimización en tiempo real de dos etapas.

1.2.3.2 Método de Adaptación por Modificadores

Si el esquema de optimización en tiempo real nos conduce a un punto tal que las condiciones de Karush-Kuhn-Tucker (KKT) del modelo adaptado coinciden con las condiciones KKT de la planta real, entonces hemos encontrado un punto operativo que minimiza (o maximiza) la función objetivo (al menos localmente).

Los métodos de adaptación por modificadores (MA) plantean realizar correcciones al funcional costo y a las restricciones de modo que las condiciones KKT del modelo y la planta coincidan [6].

Las correcciones utilizadas en los esquemas de adaptación por modificadores no son entonces explícitamente los parámetros del modelo como en el método de dos etapas, sino que representan correcciones de primer (o segundo) orden del funcional costo y de las restricciones del problema de optimización.

Al igual que en el método de dos etapas es necesario definir la adecuación del modelo utilizado. Para métodos de adaptación por modificadores el modelo del proceso es adecuado si pueden hallarse valores de modificadores tales que en un punto fijo del esquema de RTO se coincida con el óptimo de la planta. En [6] el autor compara las condiciones de adecuación del modelo para la utilización del esquema de dos etapas versus las condiciones de adecuación del modelo para el método de adaptación por modificadores. En dicho trabajo, se muestra que las condiciones de adecuación del modelo para el método de adaptación por modificadores son menos exigentes para con el modelo de planta que en el caso del método de dos etapas.

La utilización progresiva de los distintos modificadores como así también el modo de calcularlos han nutrido las diversas investigaciones en el área de adaptación por modificadores [45, 48]. Por ejemplo, en [49] la adaptación se realiza en base a la diferencia entre el valor alcanzado por las restricciones del modelo versus los valores reales de la planta. En [50] se introduce el modificador del gradiente de la función costo. A su vez, los modificadores de gradientes, o de primer orden, para las restricciones se incorporaron en las formulaciones de Gao y Engell [51] y de Marchetti et al. [52]. Recientemente se planteó la idea de incorporar los denominados modificadores del Hessiano, o modificadores de segundo orden, tanto para las restricciones como para la función costo [53].

1.3 Organización y Estructura de la Tesis

La presente tesis está organizada de la siguiente forma: En el Capítulo 1 se presentaron las principales motivaciones que dieron origen a la tesis, como así también se presentaron los antecedentes más relevantes recopilados de la literatura que dan sustento a las ideas planteadas en la tesis. En el Capítulo 2 se introducen formulaciones preliminares y se ilustran conceptos necesarios para el correcto desarrollo y seguimiento del trabajo. En el Capítulo 3 se propone la formulación del problema de programación matemática para selección de estructuras de control mediante un método auto optimizante con márgenes de seguridad. El Capítulo 4 aborda la selección de la estructura de control teniendo en cuenta la presencia de un nivel de control superior. El Capítulo 5 es dedicado al estudio de la optimización en línea para modelos imperfectos de la planta. Finalmente, en el Capítulo 6 se encuentran las conclusiones y trabajos futuros que se desprenden de la realización de esta tesis.

1.4 Principales Aportes y Publicaciones

A nivel general, el principal aporte de esta tesis son las diversas formulaciones presentadas para el estudio y la selección de estructuras de control en base a criterios económicos para procesos industriales de operación continua, en estado estacionario. En lo particular, podemos detallar:

- Se vinculan los métodos de selección de estructuras de control basados en el cálculo de márgenes de seguridad con los métodos de control auto optimizante. En este contexto, se propuso una estrategia de selección de estructuras de control mediante la formulación de problemas matemáticos mixtos enteros no lineales.

- Se formula un problema de programación no lineal que encuadra dentro de las estructuras auto optimizantes globales utilizando combinaciones lineales de variables medibles como variables controladas.
- Se formulan simplificaciones de los problemas de programación no lineal mixta entera, planteados para la selección de la estructura de control, que dan como resultado problemas de programación cuadrática mixta entera convexos.
- Se formulan estrategias de selección de estructuras de control que integran el nivel de optimización en línea con el nivel de control regulador.
- Se formula un problema de programación matemática mixta entero no lineal que integra y generaliza las distintas alternativas desarrolladas anteriormente.
- Se propone un método para el cálculo de modificadores de segundo orden aplicable a las estrategias de optimización en tiempo real por adaptación de modificadores.

El trabajo de investigación que permitió el desarrollo de la presente tesis de doctorado fue realizado en el Grupo de Ingeniería de Sistemas de Procesos (GISP), perteneciente al CIFASIS. Los resultados parciales que forman parte de los capítulos de esta tesis configuraron publicaciones en congresos nacionales y artículos en revistas internacionales indexadas.

Revistas internacionales con referato

- [I] Economic Control Structure Selection for Two-Layered Real-Time Optimization Systems, Agustín Bottari, David A. R. Zumoffen, y Alejandro G. Marchetti, *Industrial & Engineering Chemistry Research* 2020 59 (49), 21413-21428 DOI: 10.1021/acs.iecr.0c02591.
- [II] Self-Optimizing Steady-State Back-Off Approach for Control Structure Selection, Agustín Bottari, Pablo A. Marchetti, y Alejandro G. Marchetti, *Industrial & Engineering Chemistry Research* 2019 58 (30), 13699-13717, DOI: 10.1021/acs.iecr.8b06296.

Congresos nacionales e internacionales con referato

- [III] Second-Order Modifier Adaptation via Quadratic Interpolation, Agustín Bottari, Alejandro G. Marchetti, IX Congreso Argentino de Ingeniería Química, Lugar: Bahía Blanca; Año: 2017.
- [IV] Optimal Control Structure Selection for Real-Time Optimization Systems, Agustín Bottari, Alejandro G. Marchetti, 26º Congreso Argentino de Control Automático, Lugar: Ciudad Autónoma de Buenos Aires; Año: 2018.
- [V] Control secundario para una microrred CC basado en la teoría de asignación de control, Pablo G. Rullo, Patricio A. Luppi, Mauro Patrone, Agustín Bottari, Diego Feroldi, y David A. R. Zumoffen, X Congreso Argentino de Ingeniería Química, Lugar: Santa Fe; Año: 2019.
- [VI] Self-Optimizing Control Structure Selection Based on Active Constraint Regions, Agustín Bottari, David A. R. Zumoffen, y Alejandro G. Marchetti, X Congreso Argentino de Ingeniería Química, Lugar: Santa Fe; Año: 2019.
- [VII] Second Order Modifier Adaptation via Quadratic Approximation for Constrained Problems, Agustín Bottari, Alejandro G. Marchetti, XIX Reunión de Trabajo en Procesamiento de la Información y Control, Lugar: San Juan; Año: 2021.

Capítulo 2

Conceptos Preliminares y Formulaciones Matemáticas

2.1 Punto Operativo Óptimo en Estado Estacionario

2.1.1 Formulación del Problema de Optimización

El comportamiento de una planta industrial se puede representar mediante un sistema de ecuaciones diferenciales y algebraicas:

$$\mathbf{f}_D(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) = \mathbf{0}, \quad (2.1a)$$

$$\mathbf{f}_A(\mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) = \mathbf{0}, \quad (2.1b)$$

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}), \quad (2.1c)$$

donde $\mathbf{x} \in \mathbb{R}^{n_x}$ es el vector de variables de estado, $\mathbf{w} \in \mathbb{R}^{n_w}$ es el vector de variables algebraicas, $\mathbf{u} \in \mathbb{R}^{n_u}$ es el vector de variables de entrada (o de decisión), $\mathbf{d} \in \mathbb{R}^{n_d}$ es el vector de perturbaciones y parámetros inciertos, \mathbf{f}_D es el vector de n_x ecuaciones diferenciales, \mathbf{f}_A es el vector de n_w ecuaciones algebraicas, e $\mathbf{y} \in \mathbb{R}^{n_y}$ es el vector de variables de salida medibles.

Las plantas de operación continua típicamente se diseñan para operar en condiciones de estado estacionario ($\dot{\mathbf{x}} = \mathbf{0}$). El punto operativo óptimo de estado estacionario se puede obtener resolviendo un problema de programación no lineal (NLP):

$$\min_{\mathbf{x}, \mathbf{w}, \mathbf{u}} \Phi(\mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) \quad (2.2a)$$

$$\text{s.t. } \mathbf{f}_D(\mathbf{0}, \mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) = \mathbf{0}, \quad (2.2b)$$

$$\mathbf{f}_A(\mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) = \mathbf{0}, \quad (2.2c)$$

$$\mathbf{y}^L \leq \mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{w}, \mathbf{u}, \mathbf{d}) \leq \mathbf{y}^U, \quad (2.2d)$$

$$\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U, \quad (2.2e)$$

donde Φ es el funcional costo económico a minimizar, \mathbf{y}^L e \mathbf{y}^U son las cotas inferiores y superiores de las variables de salida, mientras que \mathbf{u}^L y \mathbf{u}^U son las cotas inferiores y superiores de las variables de entrada.

Para unos valores definidos de \mathbf{u} y \mathbf{d} la solución al sistema (2.2b, 2.2c) puede escribirse como:

$$(\mathbf{x}, \mathbf{w}) = \boldsymbol{\xi}(\mathbf{u}, \mathbf{d}) \quad (2.3)$$

donde ξ es un operador que expresa el mapeo en estado estacionario entre (\mathbf{u}, \mathbf{d}) y (\mathbf{x}, \mathbf{w}) . Definimos también los mapeos entrada-salida $\mathbf{y}(\mathbf{u}, \mathbf{d})$ y $J(\mathbf{u}, \mathbf{d})$ correspondientes al comportamiento en estado estacionario del proceso,

$$J(\mathbf{u}, \mathbf{d}) = \Phi(\xi(\mathbf{u}, \mathbf{d}), \mathbf{u}, \mathbf{d}), \quad (2.4a)$$

$$\mathbf{y}(\mathbf{u}, \mathbf{d}) = \mathbf{F}(\xi(\mathbf{u}, \mathbf{d}), \mathbf{u}, \mathbf{d}) \quad (2.4b)$$

La evaluación de estos mapeos requiere resolver (2.2b, 2.2c) para luego evaluar \mathbf{y} de (2.2d) y Φ de (2.2a). Al usar estos mapeos, el Problema (2.2) puede escribirse de manera compacta como:

$$\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d}) \quad (2.5a)$$

$$\text{s.t. } \mathbf{y}^L \leq \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^U, \quad (2.5b)$$

$$\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \quad (2.5c)$$

El Problema (2.5) está definido para un valor dado del vector de perturbaciones \mathbf{d} . Indicaremos mediante $\mathbf{u}^*(\mathbf{d})$ a las variables de entrada óptimas del Problema (2.5) en función del valor de las perturbaciones. En la Figura 2.1 se ilustran estos conceptos para un ejemplo particular. En la mencionada imagen, se grafican los límites de las restricciones para las salidas y_1 (en verde) e y_2 (en azul) para valores nominales de las perturbaciones, $\mathbf{d} = \mathbf{d}_n$. Debajo de dichas curvas, y sombreado en verde, se representa la zona de operación factible. Se incluye también los contornos de la función objetivo. El punto operativo óptimo $\mathbf{u}^*(\mathbf{d}_n)$ se encuentra en la intersección de las restricciones y_1 e y_2 .

Vale hacer aquí una aclaración, y es que el conjunto de restricciones que se activarán, en el óptimo de dicho problema, puede variar para distintos valores de las perturbaciones \mathbf{d} . Este tema se tratará en detalle en la Sección 4.3.1 dado que es importante para el desarrollo de las estrategias propuestas en el Capítulo 4 de la tesis.

Es común que se considere existente un valor nominal del vector de perturbaciones, indicado como \mathbf{d}_n . Dichos valores de las perturbaciones pueden representar, según el caso de estudio, los valores más frecuentes o los valores esperados de cada perturbación. En la tesis, se asume que las perturbaciones pertenecen a un conjunto \mathcal{D} que contiene todas las perturbaciones posibles que pueden presentarse. El conjunto de perturbaciones \mathcal{D} típicamente está definido considerando restricciones del tipo caja, $\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{n_d} : \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U\}$, donde \mathbf{d}^L y \mathbf{d}^U son las cotas inferiores y superiores de las perturbaciones, respectivamente.

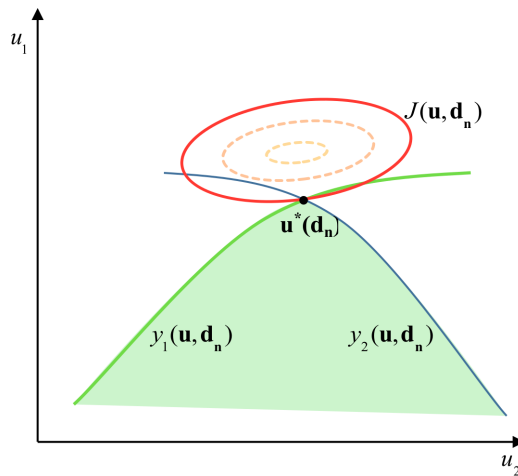


Figura 2.1: Punto operativo óptimo

2.1.2 Condiciones de Karush-Kuhn-Tucker

Consideremos el problema de optimización (2.5) para un valor predefinido del vector de perturbaciones y agrupando las restricciones sobre $\mathbf{y}(\mathbf{u})$ en un único vector, es decir:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (2.6)$$

donde el vector $\mathbf{G}(\mathbf{u})$ incluye tanto $\mathbf{y}^L - \mathbf{y}(\mathbf{u})$, como también $\mathbf{y}(\mathbf{u}) - \mathbf{y}^U$. Las condiciones de Karush-Kuhn-Tucker para el problema de optimización (2.6) pueden entonces expresarse mediante los siguientes teoremas:

Teorema 1 (Condiciones Necesarias de Karush-Kuhn-Tucker). Sea \mathbf{u}^* un óptimo (local) del Problema (2.6) y suponga que cumple la calificación de independencia lineal de las restricciones (LICQ, del inglés linear independence constraint qualification) en \mathbf{u}^* , es decir, que los gradientes de las restricciones activas en el punto óptimo son linealmente independientes. Entonces existen valores para los vectores de multiplicadores de Lagrange $\mu \in \mathbb{R}^{n_g}, \zeta^U, \zeta^L \in \mathbb{R}^{n_u}$, tales que las condiciones de primer orden de Karush-Kuhn-Tucker se satisfacen en \mathbf{u}^* :

$$\mathbf{G}(\mathbf{u}) \leq \mathbf{0}, \quad \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \quad (2.7a)$$

$$\mu^T \mathbf{G}(\mathbf{u}) = 0, \quad \zeta^{UT} (\mathbf{u} - \mathbf{u}^U) = 0, \quad \zeta^{LT} (\mathbf{u}^L - \mathbf{u}) = 0 \quad (2.7b)$$

$$\mu \geq 0, \quad \zeta^U \geq 0, \quad \zeta^L \geq 0 \quad (2.7c)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{u}} = \frac{\partial J}{\partial \mathbf{u}} + \mu^T \frac{\partial \mathbf{G}}{\partial \mathbf{u}} + \zeta^U - \zeta^L = 0, \quad (2.7d)$$

donde la funcion Lagrangiana se define como:

$$\mathcal{L}(\mathbf{u}, \mu, \zeta^U, \zeta^L) = J(\mathbf{u}) + \mu^T \mathbf{G}(\mathbf{u}) + \zeta^{UT} (\mathbf{u} - \mathbf{u}^U) + \zeta^{LT} (\mathbf{u}^L - \mathbf{u}) \quad (2.8)$$

Las condiciones (2.7a) se conocen como las condiciones de factibilidad primal, la condiciones (2.7b) como las condiciones de holgura complementaria y las condiciones (2.7c), (2.7d) como las condiciones de factibilidad duales. Expresadas de manera conjunta las condiciones (2.7a-2.7d) se denominan las condiciones de KKT. Si un punto \mathbf{u}^* satisface las condiciones KKT se dice que dicho punto es un punto KKT. Si se cumple la calificación de las restricciones, los puntos mínimos del problema de optimización son necesariamente puntos KKT, sin embargo, al ser condiciones necesarias puede resultar que un punto satisfaga las condiciones de KKT pero que no sea un mínimo sino un máximo o un punto de ensilladura. Las condiciones suficientes de segundo orden nos permitirán identificar si el punto KKT se trata efectivamente de un mínimo local estricto.

Definamos primero el vector de multiplicadores de Lagrange como:

$$\boldsymbol{\nu}^T = [\mu^T, \zeta^{UT}, \zeta^{LT}] \quad (2.9)$$

A su vez definimos el vector que recolecta todas las restricciones de desigualdad como:

$$\mathbf{v}(\mathbf{u}) := \begin{bmatrix} \mathbf{G}(\mathbf{u}) \\ \mathbf{u}^L - \mathbf{u} \\ \mathbf{u} - \mathbf{u}^U \end{bmatrix} \leq \mathbf{0}, \quad (2.10)$$

Utilizando esta notación, las condiciones suficientes de segundo orden se presentan a continuación.

Teorema 2 (Condiciones Suficientes de Segundo Orden de KKT). Sea \mathbf{u}^* un punto KKT del Problema (2.6) con los multiplicadores de Lagrange $\nu^{*T} = [\mu^{*T}, \zeta^{U^{*T}}, \zeta^{L^{*T}}]$. Sea $\mathcal{A}_v := \{i : v_i(\mathbf{u}^*) = 0, i = 1, \dots, n_v\}$, con $n_v = n_g + 2n_u$. Especificamos las restricciones fuertemente activas como $\mathcal{A}_v^+ := \{i \in \mathcal{A}_v : \nu_i^* > 0\}$ y las restricciones débilmente activas como $\mathcal{A}_v^0 := \{i \in \mathcal{A}_v : \nu_i^* = 0\}$. Definimos la función Lagrangiana restringida como:

$$L(\mathbf{u}) := \mathcal{L}(\mathbf{u}, \mu^*, \zeta^{U^*}, \zeta^{L^*}) = J(\mathbf{u}) + \sum_{i \in \mathcal{A}_v} \nu_i^* v_i(\mathbf{u}) \quad (2.11)$$

y su Hessiano en \mathbf{u}^* como:

$$\frac{\partial^2 L}{\partial \mathbf{u}^2}(\mathbf{u}^*) = \frac{\partial^2 J}{\partial \mathbf{u}^2}(\mathbf{u}^*) + \sum_{i \in \mathcal{A}_v} \nu_i^* \frac{\partial^2 v_i}{\partial \mathbf{u}^2}(\mathbf{u}^*). \quad (2.12)$$

Definimos el cono:

$$C = \left\{ \mathbf{d} \neq 0 : \frac{\partial v_i}{\partial \mathbf{u}}(\mathbf{u}^*) \mathbf{d} = 0 \text{ for } i \in \mathcal{A}_v^+; \quad \frac{\partial v_i}{\partial \mathbf{u}}(\mathbf{u}^*) \mathbf{d} \leq 0 \text{ for } i \in \mathcal{A}_v^0 \right\} \quad (2.13)$$

Entonces, si $\mathbf{d}^T \frac{\partial^2 L}{\partial \mathbf{u}^2}(\mathbf{u}^*) \mathbf{d} > 0$ para todo $\mathbf{d} \in C$, \mathbf{u}^* es un mínimo local estricto para el Problema (2.6).

2.2 Problema de Selección de Estructura de Control

Incorporar los lazos de control en el problema de optimización implica expresarlos como ecuaciones. Para ello podemos construir un vector $\mathbf{r}(\mathbf{u}, \mathbf{y}) \in \mathbb{R}^{n_u}$ que esté formado por las variables de salida controladas y las variables de entrada que permanecerán fijas. A su vez, las variables de entrada que no se fijan en el vector $\mathbf{r}(\mathbf{u}, \mathbf{y})$ son las variables manipuladas seleccionadas. Del mismo modo introducimos un vector $\mathbf{r}^{\text{sp}} \in \mathbb{R}^{n_u}$, que incluye los valores de consigna para las variables de salidas controladas y los valores a los que se mantendrán fijas las variables de entrada seleccionadas. De este modo el comportamiento de la planta controlada en estado estacionario puede tenerse en cuenta incluyendo las siguientes n_u ecuaciones al problema de optimización:

$$\mathbf{r}(\mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d})) = \mathbf{r}^{\text{sp}}. \quad (2.14)$$

Ejemplo 1. Consideremos una planta con $n_y = 4$ variables de salida y $n_u = 3$ variables de entrada. Una posible estructura de control se ilustra en la Figura 2.2 donde la salida y_1 se controla manipulado u_3 , y_2 se controla manipulado u_2 y la restante variable manipulada u_1 se fija a un valor de consigna seleccionado. Entonces la ecuación 2.14 a incorporar como restricciones en el problema de optimización a lazo cerrado se define con:

$$\mathbf{r} = \begin{bmatrix} u_1 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_1^{\text{sp}} \\ y_1^{\text{sp}} \\ y_2^{\text{sp}} \end{bmatrix} = \mathbf{r}^{\text{sp}} \quad (2.15)$$

Los valores de consigna $\mathbf{r}^{\text{sp}} \in \mathbb{R}^{n_u}$, pasan a ser las variables a optimizar en un problema a lazo cerrado. Es decir, el problema de operación óptima a lazo cerrado se puede expresar como:

$$\underset{\mathbf{r}^{\text{sp}}}{\text{mín}} \quad J(\mathbf{u}, \mathbf{d}) \quad (2.16a)$$

$$\text{s.t.} \quad \mathbf{r}(\mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d})) = \mathbf{r}^{\text{sp}} \quad (2.16b)$$

$$\mathbf{y}^L \leq \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^U, \quad (2.16c)$$

$$\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \quad (2.16d)$$

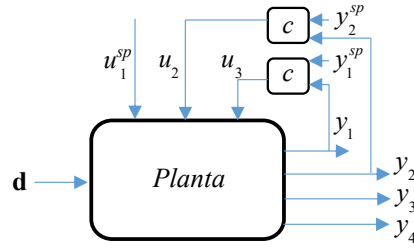


Figura 2.2: Ejemplo de estructura de control en procesos multivariables.

Este Problema (2.16) está definido para un valor dado del vector de perturbaciones \mathbf{d} y para una estructura de control $\mathbf{r}(\mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d}))$ establecida.

2.3 Problema de los Márgenes de Seguridad

2.3.1 Interpretación Conceptual del Problema de los Márgenes de Seguridad

Se comentó en la introducción que, a consecuencia del efecto de las perturbaciones en el proceso, comúnmente es necesario alejarse de los límites de las restricciones para evitar posibles violaciones de las mismas. En esta sección, mediante ejemplos conceptuales presentaremos una interpretación gráfica de dicha idea.

Ejemplo 2. Consideremos primero para el problema de operación óptima el comportamiento de una planta a lazo abierto como el mostrado en la Figura 2.3. En la imagen de la izquierda de la Figura 2.3 se ilustran los límites de las restricciones para las variables de salida y_1 e y_2 , en verde y en azul respectivamente, y para dos valores diferentes del vector de perturbaciones, \mathbf{d}_k y \mathbf{d}_n . La zona de operación factible se encuentra por debajo de las curvas de contorno correspondientes a los valores límite de las restricciones. Los contornos de la función objetivo se ilustra para el valor de la perturbación nominal \mathbf{d}_n .

Si fijamos la perturbación en su valor nominal \mathbf{d}_n y resolvemos el problema de optimización (2.5), entonces tendremos como sugerencia (desde el punto de vista de un funcional económico) poner la planta a operar en el punto $\mathbf{u}^*(\mathbf{d}_n) = [u_{1,n}, u_{2,n}]^T$. No obstante, al producirse un cambio en los valores de la perturbaciones, es decir, al pasar de \mathbf{d}_n a \mathbf{d}_k , los límites de las restricciones y_1 e y_2 se mueven y dicho punto operativo $\mathbf{u}^*(\mathbf{d}_n)$ queda fuera de la región factible. Para evitar esto, podemos seleccionar un punto operativo a lazo abierto considerando un margen de seguridad. Este nuevo punto operativo $\mathbf{u}^{sp} = [u_{1,k}, u_{2,k}]^T$ será seguro para las perturbaciones conocidas (supongamos simplemente \mathbf{d}_k y \mathbf{d}_n) pero para el valor de perturbación \mathbf{d}_n se tendrá un rendimiento económico subóptimo. Una alternativa, que está a la vista en este ejemplo para operar en forma óptima a pesar de la ocurrencia de perturbaciones y evitar violar las restricciones, es colocar lazos de control sobre y_1 e y_2 y para los valores analizados aquí de \mathbf{d}_k y \mathbf{d}_n , se podría fijar como valores de consigna los valores límites superiores de las restricciones y operar la planta sin márgenes de seguridad y sin pérdida de optimalidad tal como se muestra en la Figura 2.4.

Podría pensarse que la necesidad de incluir márgenes de seguridad en estado estacionario es un inconveniente exclusivo de operar la planta a lazo abierto. Para notar que esto no es cierto, bastará con complejizar el ejemplo anterior con una nueva restricción.

Ejemplo 3. Incorporamos una tercera restricción y_3 como se muestra en la Figura 2.5 y ampliamos el rango de la perturbación considerando también un valor de perturbación \mathbf{d}_l . Ahora al ingresar el escenario de perturbación \mathbf{d}_l , los límites de las restricciones se mueven según se muestran en la Figura

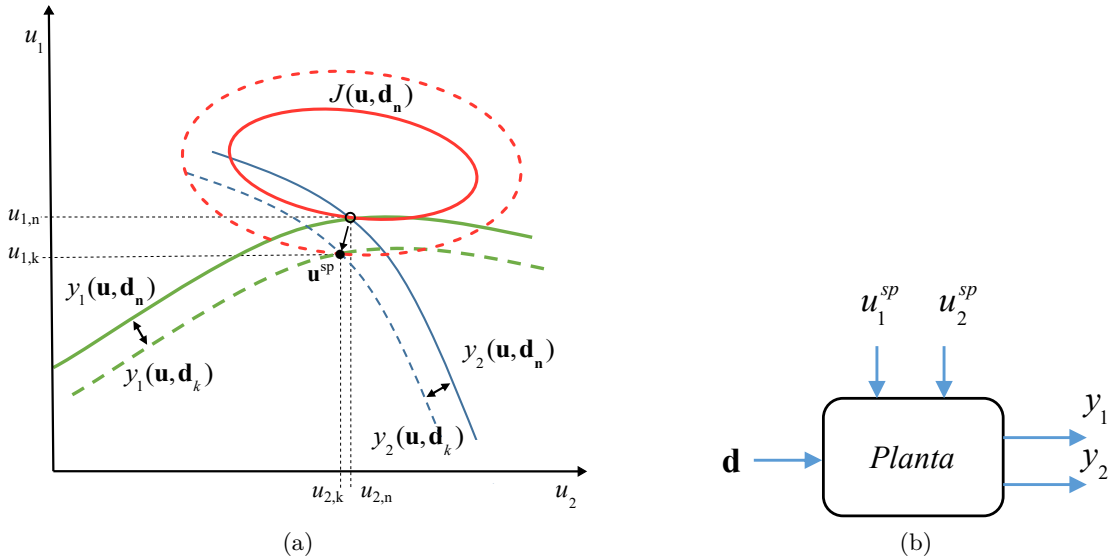


Figura 2.3: Ejemplo margen de seguridad a lazo abierto. En la imagen de la derecha se ilustra el sistema a considerar. En la imagen de la izquierda, la flecha negra hacia \mathbf{u}^{sp} indica el movimiento que debe hacer el punto operativo para evitar violar las restricciones y para mantener la factibilidad ante la presencia de la perturbación \mathbf{d}_k .

2.5 y, como puede observarse, la estructura de control que persigue los límites superiores de y_1 e y_2 deberá adoptar un margen de seguridad para evitar violar la restricción sobre y_3 . Nuevamente la operación de la planta, esta vez a lazo cerrado, tendrá una pérdida económica asociada al margen de seguridad que se seleccione.

La magnitud del retroceso hacia la zona factible que hay que adoptar depende de la variabilidad de las variables restringidas para el sistema controlado a lazo cerrado. El enfoque de los márgenes de seguridad para la selección de la estructura de control de procesos se basa en la idea de seleccionar la estructura de control que minimiza el costo económico asociado con los retrocesos en las restricciones.

2.3.2 Formulación Matemática del Método de los Márgenes de Seguridad

El objetivo del problema de los márgenes de seguridad en estado estacionario es determinar la estructura de control y los valores de consigna que minimizan el costo económico para valores nominales $(\mathbf{u}_n, \mathbf{d}_n)$ de perturbación, al tiempo que garantizan la factibilidad en estado estacionario para todas las perturbaciones posibles $\mathbf{d} \in \mathcal{D}$ [4]. Sea \mathbf{z} el vector de variables binarias que determina la elección de la estructura de control. Para un \mathbf{z} dado, es posible construir el vector $\mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y}) \in \mathbb{R}^{n_u}$, que incluye las variables de salida controladas seleccionadas y las variables de entrada fijas. Las variables de entrada que no están incluidas en $\mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y})$ son las variables manipuladas seleccionadas. Asimismo, introducimos el vector $\mathbf{r}^{sp} \in \mathbb{R}^{n_u}$, que incluye los valores de consigna para las variables de salida controladas y los valores de consigna para las variables de entrada fijas. De esta manera, el efecto de los lazos de control sobre el comportamiento de estado estacionario de la planta controlada, según la estructura definida por \mathbf{z} , puede modelarse introduciendo las siguientes n_u ecuaciones:

$$\mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d})) = \mathbf{r}^{sp}. \quad (2.17)$$

Aquí se asume que el controlador está diseñado con acción integral y que las variables manipuladas \mathbf{u} de la planta controlada en estado estacionario se pueden calcular resolviendo (2.17) para \mathbf{d} y \mathbf{r}^{sp} dados. Notar que a diferencia de la ecuación (2.14), esta nueva expresión (2.17) posee el vector de

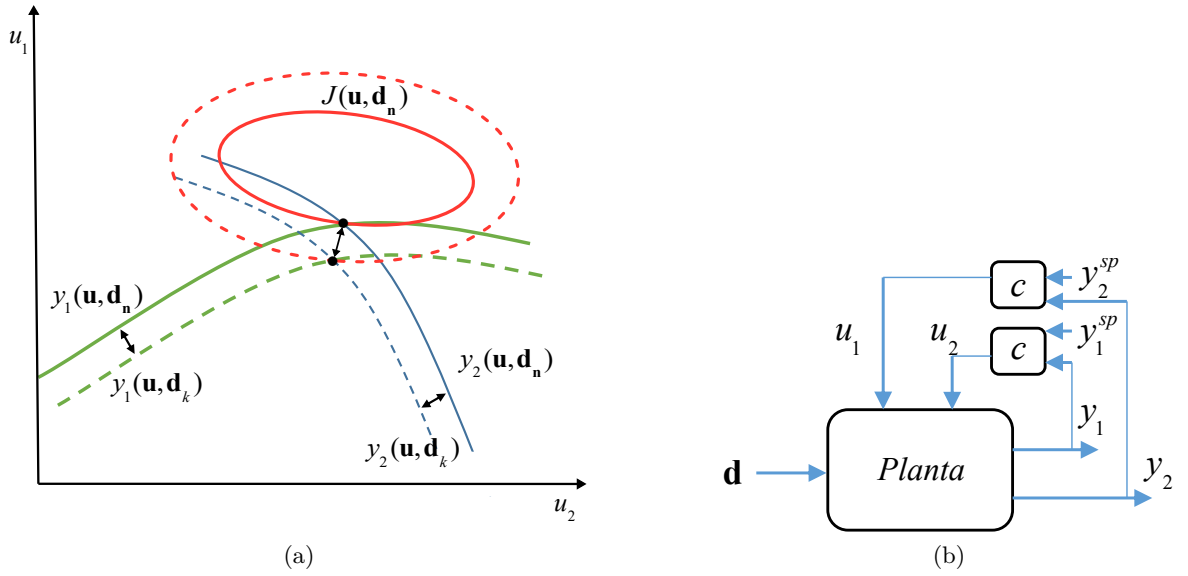


Figura 2.4: Ejemplo de lazo cerrado sin la necesidad de márgenes de seguridad. Al controlar y_1 e y_2 , para el problema mostrado, se evita el uso de márgenes de seguridad. El punto operativo varía según se mueven los límites de las restricciones de modo que siempre es un punto factible en estado estacionario. El rendimiento económico también mejora en comparación con la estructura de lazo abierto.

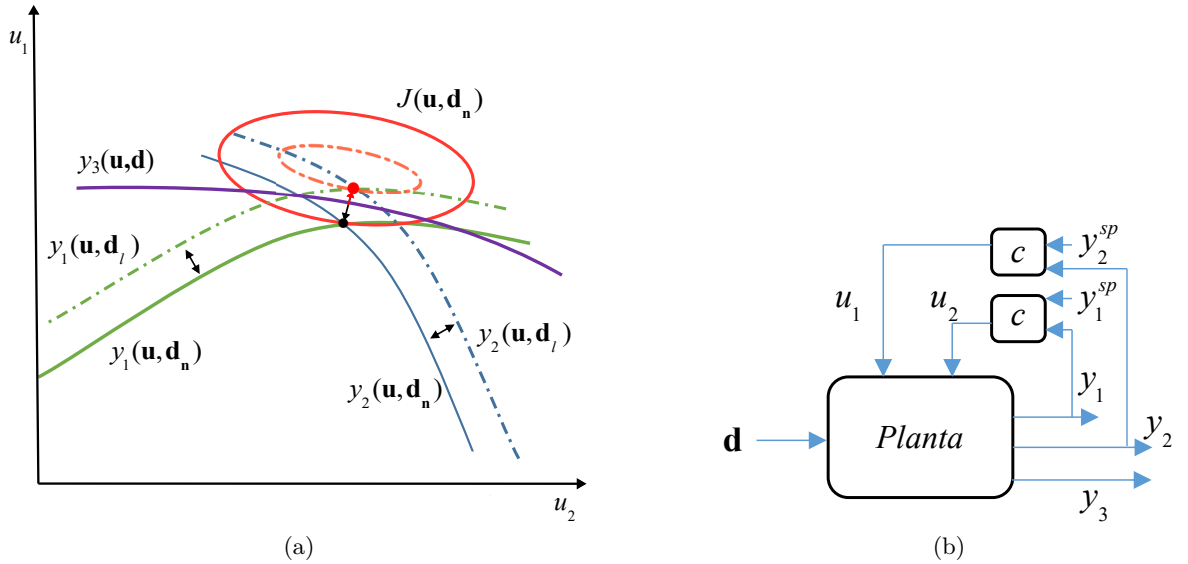


Figura 2.5: Estructura a lazo cerrado. El punto operativo infactible se muestra en la imagen de la izquierda y está representado con un punto rojo. La estructura necesitará un margen de seguridad para evitar violar la restricción y_3 al ingresar la perturbación \mathbf{d}_l .

variables binarias \mathbf{z} . El problema de los márgenes de seguridad en estado estacionario a lazo cerrado entonces se puede formular de la siguiente manera:

$$\begin{aligned}
 & \min_{\mathbf{z}, \mathbf{r}^{sp}} J(\mathbf{u}_n, \mathbf{d}_n) \\
 & \text{s.t.} \quad \mathbf{r}(\mathbf{z}, \mathbf{u}_n, \mathbf{y}(\mathbf{u}_n, \mathbf{d}_n)) = \mathbf{r}^{sp}, \quad \mathbf{d}_n \in \mathcal{D} \\
 & \quad \mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d})) = \mathbf{r}^{sp}, \quad \forall \mathbf{d} \in \mathcal{D} \\
 & \quad \mathbf{y}^L \leq \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^U, \quad \forall \mathbf{d} \in \mathcal{D} \\
 & \quad \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U.
 \end{aligned} \tag{2.18}$$

Esta formulación corresponde a un problema de optimización semi-infinito no lineal mixto entero. Los valores factibles de las variables de decisión \mathbf{z} y \mathbf{r}^{SP} requieren que para cada perturbación $\mathbf{d} \in \mathcal{D}$ exista una entrada \mathbf{u} de manera que se satisfagan las restricciones. Notar que el Problema (2.18) es una versión de estado estacionario del problema de los márgenes de seguridad dinámicos a lazo cerrado presentado en Bahri et al. [12].

2.4 Problema de Control Auto-Optimizante

2.4.1 Pérdida Económica Promedio

Los métodos de control auto optimizantes (SOC) propuestos en la literatura se basan en la idea de seleccionar las variables controladas de forma de minimizar la pérdida económica de la planta controlada con respecto al funcionamiento óptimo de la planta en presencia de perturbaciones [5, 29]. Esta pérdida viene dada por:

$$L(\mathbf{z}, \mathbf{d}) = J(\mathbf{u}(\mathbf{z}, \mathbf{d}), \mathbf{d}) - J(\mathbf{u}^*(\mathbf{d}), \mathbf{d}), \quad (2.19)$$

donde $\mathbf{u}(\mathbf{z}, \mathbf{d})$ son los valores de las variables de entrada alcanzados en el estado estacionario de la planta controlada usando la estructura de control \mathbf{z} , y $\mathbf{u}^*(\mathbf{d})$ son los valores óptimos de las variables de entrada obtenidos mediante la resolución del Problema (2.5) para los valores de las perturbaciones dados por \mathbf{d} .

Los llamados *métodos de pérdida promedio* [31, 34, 37, 38] tienen como objetivo seleccionar las variables controladas que minimizan el valor esperado (o el valor promedio) de la pérdida económica para un conjunto de perturbaciones determinado \mathcal{D} . La pérdida esperada viene dada por:

$$L_{\text{exp}}(\mathbf{z}) = \mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}(\mathbf{z}, \mathbf{d}), \mathbf{d}) - J(\mathbf{u}^*(\mathbf{d}), \mathbf{d}) \right\} \quad (2.20)$$

2.4.2 Minimización del Costo Económico Promedio

Consideremos nuevamente la pérdida económica esperada (2.20) usada en los métodos de control auto optimizantes y observemos que:

$$\mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}(\mathbf{z}, \mathbf{d}), \mathbf{d}) - J(\mathbf{u}^*(\mathbf{d}), \mathbf{d}) \right\} = \mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}(\mathbf{z}, \mathbf{d}), \mathbf{d}) \right\} - \mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}^*(\mathbf{d}), \mathbf{d}) \right\}. \quad (2.21)$$

Observación 1 (Pérdida Económica Esperada Vs. Costo Esperado). Dado que el segundo término en el lado derecho de la ecuación (óptimos en función de la perturbación) es una constante que no depende de la estructura de control \mathbf{z} , se deduce que, para el propósito de la selección de la estructura de control, minimizar la pérdida económica esperada es equivalente a minimizar el costo económico esperado.

Los métodos de control auto optimizante se han desarrollado con el objetivo de minimizar la pérdida económica con respecto a la operación óptima del proceso [28, 30, 31] y no incluyen márgenes de seguridad de estado estacionario o dinámicos. Por otra parte, los métodos de márgenes de seguridad buscan minimizar el costo económico. En esta tesis, y haciendo referencia a la observación anterior, mostraremos los vínculos existentes entre ambos métodos como así también la potencial ventaja de utilizar una formulación general para la selección de estructuras de control.

2.4.3 Estructuras de Combinaciones Lineales

La mayoría de los métodos de control auto optimizante seleccionan como variables controladas combinaciones lineales de las variables de entrada y salida. En este caso, el vector \mathbf{z} que determina la elección de la estructura de control incluye los coeficientes de las combinaciones lineales, y las variables controladas $\mathbf{r} \in \mathbb{R}^{n_u}$ están dadas por:

$$\mathbf{r} = \mathcal{Q}^T \begin{bmatrix} \mathbf{y}(\mathbf{u}, \mathbf{d}) \\ \mathbf{u} \end{bmatrix}, \quad (2.22)$$

donde $\mathcal{Q} \in \mathbb{R}^{(n_u+n_y) \times n_u}$ es la matriz de combinaciones lineales. En esta tesis, distinguiremos entre *estructuras de control clásicas* y *estructuras de control con combinaciones lineales*.

Definición 1 (Estructura de Control Clásica). Una estructura de control clásica es una estructura de control cuadrada que no incluye combinaciones lineales de variables a controlar. En las estructuras de control clásicas, el vector \mathbf{z} incluye solo variables binarias, y las columnas de la matriz \mathcal{Q} son vectores unitarios con exactamente un elemento igual a uno y el resto de los elementos iguales a cero, mientras que las filas en \mathcal{Q} son vectores unitarios o vectores nulos.

Definición 2 (Estructura de Control de Combinaciones Lineales). Una estructura de control de combinaciones lineales es una estructura de control (cuadrada) que selecciona como variable controlada al menos una combinación lineal de las variables de salida y de entrada, o fija al menos una combinación lineal de variables de entrada. En estas estructuras de control, al menos una columna de la matriz \mathcal{Q} tiene más de un elemento diferente de cero.

Ejemplo 4. Consideremos un sistema con dos entradas ($n_u = 2$) y tres salidas medibles ($n_y = 3$), con las siguientes estructuras de control:

a) Una estructura de control clásica definida por $\mathbf{z}^0 = [0 \ 1 \ 0]$ y $\mathbf{z}^1 = [0 \ 1]$. Para esta estructura de control la variable de salida controlada es y_2 mediante la manipulación de la variable u_1 , mientras que la variable de entrada u_2 permanecerá fija en un valor de referencia adoptado:

$$\mathbf{r} = \mathcal{Q}^T \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} y_2 \\ u_2 \end{bmatrix} \quad (2.23)$$

b) Una estructura de combinación lineal:

$$\mathbf{r} = \mathcal{Q}^T \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0.3} & \mathbf{1.4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} (0.3)y_2 + (1.4)y_3 \\ u_2 \end{bmatrix} \quad (2.24)$$

Notar que en este caso el vector \mathbf{z} podría construirse concatenando los valores de la matriz de combinaciones lineales. La matriz \mathcal{Q}^T es la manera general y mas práctica para expresar ambas posibilidades.

2.4.4 Método del Espacio Nulo

El método del espacio nulo [30] es un buen método para explicar por qué se seleccionan combinaciones lineales de variables de salida y de entrada como variables controladas en los métodos de control auto optimizante. Para este propósito, será suficiente la consideración de un problema de optimización sin restricciones.

La condición necesaria de optimalidad de primer orden del problema sin restricciones, $\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{d})$, es que el gradiente de la función costo sea igual a cero en el punto de solución, es decir, $\frac{\partial J}{\partial \mathbf{u}}(\mathbf{u}, \mathbf{d}) = \mathbf{0}$. La variación de primer orden de esta condición evaluada en el óptimo nominal está dada por:

$$\frac{\partial^2 J}{\partial \mathbf{u}^2} \Big|_n \delta \mathbf{u} + \frac{\partial^2 J}{\partial \mathbf{u} \partial \mathbf{d}} \Big|_n \delta \mathbf{d} = \mathbf{0} \quad (2.25)$$

Suponiendo que el óptimo nominal $\mathbf{u}^*(\mathbf{d}_n)$ es un mínimo local estricto, la matriz hessiana es positiva definida, y por lo tanto no singular, y podemos escribir:

$$\delta \mathbf{u}^*(\mathbf{d}) = \mathbf{K} \delta \mathbf{d}, \quad \text{con } \mathbf{K} = - \left(\frac{\partial^2 J}{\partial \mathbf{u}^2} \Big|_n \right)^{-1} \frac{\partial^2 J}{\partial \mathbf{u} \partial \mathbf{d}} \Big|_n \quad (2.26)$$

A su vez, la variación de primer orden de las variables de salida viene dada por:

$$\delta \mathbf{y} = \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \Big|_n \delta \mathbf{u} + \frac{\partial \mathbf{y}}{\partial \mathbf{d}} \Big|_n \delta \mathbf{d} \quad (2.27)$$

Usando (2.26) y (2.27) podemos escribir:

$$\begin{bmatrix} \delta \mathbf{y} \\ \delta \mathbf{u} \end{bmatrix} = \mathbf{S} \delta \mathbf{d}, \quad \text{con } \mathbf{S} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \Big|_n \mathbf{K} + \frac{\partial \mathbf{y}}{\partial \mathbf{d}} \Big|_n \\ \mathbf{K} \end{bmatrix}. \quad (2.28)$$

Supongamos que $\mathbf{S} \in \mathbb{R}^{(n_y+n_u) \times n_d}$ es de rango completo (columnas linealmente independientes), y que las columnas en $\mathcal{N} \in \mathbb{R}^{(n_y+n_u) \times n_u}$ son un conjunto de n_u vectores ortogonales que se encuentran en el espacio nulo izquierdo de \mathbf{S} . Por lo tanto, $\mathcal{N}^T \mathbf{S} = \mathbf{0}$ y de (2.28) tenemos,

$$\delta \mathbf{r} = \mathcal{N}^T \begin{bmatrix} \delta \mathbf{y} \\ \delta \mathbf{u} \end{bmatrix} = \mathcal{N}^T \mathbf{S} \delta \mathbf{d} = \mathbf{0}. \quad (2.29)$$

El método del espacio nulo consiste en seleccionar la matriz de combinación $\mathcal{Q} = \mathcal{N}$, y controlar $\delta \mathbf{r} = \mathbf{0}$ mediante control por retroalimentación [30]. Observe que la dimensión del espacio nulo de \mathbf{S} debe ser mayor o igual que n_u , lo cual requiere $n_y + n_u - n_d \geq n_u$. Esta última desigualdad se reduce a $n_y \geq n_d$. Es decir, el número de variables de salida medibles debe ser mayor o igual al número de perturbaciones [36]. En presencia de perturbaciones, el método del espacio nulo conduce el punto de operación del proceso hacia un punto de operación cercano al punto óptimo de operación. Al ser un método SOC local, la calidad de dicha aproximación puede empeorar para valores de las perturbaciones distantes de los valores nominales. De hecho, usando la variación de primer orden (2.27) se puede demostrar que hacer cumplir $\delta \mathbf{r} = \mathbf{0}$ da como resultado $\delta \mathbf{u} = \mathbf{K} \delta \mathbf{d}$ en estado estacionario [32].

2.5 Problema de Optimización en Tiempo Real

En los problemas de optimización en tiempo real de procesos industriales rara vez se dispone de un modelo de planta perfecto. Será necesario entonces distinguir entre el problema de optimización (2.6) referido a la planta real y el referido al modelo disponible. Notaremos con el subíndice $(.)_p$ a la representación perfecta de la planta. De este modo el problema de optimización de la planta real está dado por:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J_p(\mathbf{u}) \\ \text{s.t.} \quad & \mathbf{G}_p(\mathbf{u}) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U, \end{aligned} \quad (2.30)$$

donde se asume que $J_p(\mathbf{u})$ y $\mathbf{G}_p(\mathbf{u})$ son funciones conocidas de las variables de entrada \mathbf{u} y de las variables de salida $\mathbf{y}_p(\mathbf{u})$, y por lo tanto, se puede conocer sus valores a través del conocimiento de las variables de entrada \mathbf{u} aplicadas y las variables de salida medidas $\mathbf{y}_p(\mathbf{u})$. Sin embargo, en la práctica, el mapeo de entrada-salida de estado estacionario de la planta $\mathbf{y}_p(\mathbf{u})$ es en general desconocido, y solo un modelo aproximado está disponible. El mapeo de entrada-salida de estado estacionario del modelo se puede denotar por $\mathbf{y}(\mathbf{u}, \boldsymbol{\theta})$, donde $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ es el conjunto de parámetros ajustables del modelo.

El problema de optimización basado en modelo se puede plantear como:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (2.31)$$

Indicaremos a la solución óptima del Problema (2.30) como \mathbf{u}_p^* . Debido al error de modelado, la solución óptima del Problema (2.31), denotada \mathbf{u}^* , generalmente no coincide con el óptimo de planta \mathbf{u}_p^* de (2.30).

2.5.1 Método Clásico de Dos Etapas

El método clásico de dos etapas procede iterativamente estimando los parámetros del modelo disponible y optimizando el modelo actualizado de la planta [43, 46]. En un primer paso de la i -ésima iteración, se estiman los parámetros $\boldsymbol{\theta}_i^*$ que minimizan la diferencia entre las salidas medibles de la planta versus las salidas del modelo en estado estacionario, es decir:

$$\begin{aligned} \min_{\boldsymbol{\theta}} \quad & \phi^{id} \\ \text{donde} \quad & \phi^{id} := [\mathbf{y}_p(\mathbf{u}_i) - \mathbf{y}(\mathbf{u}_i, \boldsymbol{\theta})]^T \mathbf{R} [\mathbf{y}_p(\mathbf{u}_i) - \mathbf{y}(\mathbf{u}_i, \boldsymbol{\theta})], \end{aligned} \quad (2.32)$$

donde \mathbf{R} es una matriz cuadrada (en general la inversa de la matriz de covarianza de los errores de medición).

El segundo paso consiste en resolver el problema de optimización para obtener las próximas entradas óptimas \mathbf{u}_{i+1}^* del modelo actualizado:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}_i^*) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}_i^*) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U, \end{aligned} \quad (2.33)$$

La solución óptima \mathbf{u}_{i+1}^* suele filtrarse mediante un filtro exponencial de primer orden antes de aplicarse a la planta:

$$\begin{aligned}\mathbf{u}_{i+1} &:= (\mathbf{I} - \mathbf{K})\mathbf{u}_i + \mathbf{K}\mathbf{u}_{i+1}^* \\ i &:= i + 1,\end{aligned}\tag{2.34}$$

donde \mathbf{K} es una matriz diagonal de ganancia con elementos diagonales dentro del intervalo $(0, 1]$.

Como se mencionó anteriormente se define como adecuado (para un esquema de dos etapas) a un modelo de proceso para el cual el óptimo de la planta es un punto fijo del algoritmo de optimización en tiempo real de estimación de parámetros o de dos etapas.

2.5.2 Método de Adaptación por Modificadores

Ante una diferencia planta-modelo estructural es posible adaptar el modelo disponible mediante el uso de modificadores a fin de que se converja a un punto donde las condiciones KKT de optimalidad del modelo y de la planta real sean coincidentes.

En los esquemas de adaptación por modificadores, para manejar la diferencia entre el modelo y la planta, el problema de optimización (2.31) se adapta iterativamente. Partiendo de la optimización del modelo para la iteración i , obtenemos el punto operativo \mathbf{u}_{i+1} y el vector $\mathbf{C}(\mathbf{u}_{i+1})$. El vector $\mathbf{C}(\mathbf{u}_{i+1})$ del modelo contiene los valores actuales del funcional costo, las restricciones, como también los gradientes evaluados en \mathbf{u}_{i+1} , es decir:

$$\mathbf{C}(\mathbf{u}_{i+1}, \boldsymbol{\theta}) = [G_1(\mathbf{u}_{i+1}, \boldsymbol{\theta}), \dots, G_{n_g}(\mathbf{u}_{i+1}, \boldsymbol{\theta}), \frac{\partial G_1}{\partial \mathbf{u}}(\mathbf{u}_{i+1}, \boldsymbol{\theta}), \dots, \frac{\partial G_{n_g}}{\partial \mathbf{u}}(\mathbf{u}_{i+1}, \boldsymbol{\theta}), J(\mathbf{u}_{i+1}, \boldsymbol{\theta}), \frac{\partial J}{\partial \mathbf{u}}(\mathbf{u}_{i+1}, \boldsymbol{\theta})]^T.\tag{2.35}$$

El punto operativo sugerido se aplica a la planta para obtener las salidas reales del sistema en estado estacionario, las mismas se recolectan en el vector $\mathbf{C}_p(\mathbf{u}_{i+1})$, esto es:

$$\mathbf{C}_p(\mathbf{u}_{i+1}) = [G_{p,1}(\mathbf{u}_{i+1}), \dots, G_{p,n_g}(\mathbf{u}_{i+1}), \frac{\partial G_{p,1}}{\partial \mathbf{u}}(\mathbf{u}_{i+1}), \dots, \frac{\partial G_{p,n_g}}{\partial \mathbf{u}}(\mathbf{u}_{i+1}), J_p(\mathbf{u}_{i+1}), \frac{\partial J_p}{\partial \mathbf{u}}(\mathbf{u}_{i+1})]^T,\tag{2.36}$$

donde el subíndice $(\cdot)_p$ hace referencia a los valores de la planta real. Notar aquí que los gradientes de la planta no son una salida directa del sistema y deben obtenerse experimentalmente.

Los modificadores entonces se obtienen al realizar la diferencia entre los valores recolectados de la planta versus los estimados por el modelo.

$$\boldsymbol{\varepsilon}_{i+1}^J = J_p(\mathbf{u}_{i+1}) - J(\mathbf{u}_{i+1}, \boldsymbol{\theta}),\tag{2.37}$$

$$(\boldsymbol{\lambda}_{i+1}^J)^T = \left. \frac{\partial J_p}{\partial \mathbf{u}} \right|_{\mathbf{u}_{i+1}} - \left. \frac{\partial J}{\partial \mathbf{u}} \right|_{\mathbf{u}_{i+1}},\tag{2.38}$$

$$\boldsymbol{\varepsilon}_{i+1}^G = G_p(\mathbf{u}_{i+1}) - G(\mathbf{u}_{i+1}, \boldsymbol{\theta}),\tag{2.39}$$

$$(\boldsymbol{\lambda}_{i+1}^G)^T = \left. \frac{\partial G_p}{\partial \mathbf{u}} \right|_{\mathbf{u}_{i+1}} - \left. \frac{\partial G}{\partial \mathbf{u}} \right|_{\mathbf{u}_{i+1}},\tag{2.40}$$

donde ε_{i+1}^J es la diferencia entre las funciones de costo de la planta y del modelo (modificador de orden cero) y $(\lambda_{i+1}^J)^T$ representa la diferencia entre el gradiente del costo de la planta y del modelo (modificador de primer orden). Los modificadores de las restricciones, ε_{i+1}^G y $(\lambda_{i+1}^G)^T$ se definen de manera similar.

Los modificadores comúnmente se presentan agrupados en arreglos vectoriales:

$$\Lambda_{i+1}^J = [\varepsilon_{i+1}^J \quad \text{vec}(\lambda_{i+1}^J)]^T \quad (2.41)$$

$$\Lambda_{i+1}^G = [\varepsilon_{i+1}^G \quad \text{vec}(\lambda_{i+1}^G)]^T \quad (2.42)$$

A su vez estos vectores pueden agruparse y representarse como:

$$\Lambda_{i+1} = [\Lambda_{i+1}^G \quad \Lambda_{i+1}^J]^T = [C_p(\mathbf{u}_{i+1}) - C(\mathbf{u}_{i+1}, \boldsymbol{\theta})] \quad (2.43)$$

Típicamente, estos modificadores pasan por un filtro exponencial de primer orden antes de usarse para adaptar el modelo en la próxima iteración [6]:

$$\Lambda_{i+1} := (\mathbf{I} - \mathbf{K})\Lambda_i + \mathbf{K}[C_p(\mathbf{u}_{i+1}) - C(\mathbf{u}_{i+1}, \boldsymbol{\theta})] \quad (2.44)$$

donde \mathbf{K} es una matriz diagonal de ganancia con elementos diagonales dentro del intervalo $(0, 1]$.

Los últimos pasos del esquema de adaptación por modificadores son la actualización del índice de la iteración ($i = i + 1$) y resolver el problema de optimización con el modelo adaptado. El problema de optimización con modificadores se escribe de la siguiente manera:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^J + (\lambda_i^J)^T (\mathbf{u} - \mathbf{u}_i) \\ \text{s.t.} \quad & G(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^G + (\lambda_i^G)^T (\mathbf{u} - \mathbf{u}_i) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (2.45)$$

En la Figura 2.6 se ilustra el esquema de adaptación por modificadores descripto.

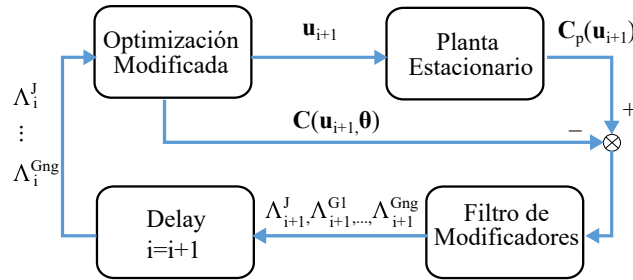


Figura 2.6: Esquema del método de adaptación por modificadores

Las estrategias para el cálculo y aplicación de los modificadores tanto en el funcional costo como en las restricciones varían según los autores y sus investigaciones en el área [54]. El método más común para la obtención de los gradientes es la utilización de diferencias finitas. En este método, para el punto operativo actual, se realizan perturbaciones en cada una de las variables de entrada a fin de estimar el gradiente. El método de diferencias finitas requiere por ejemplo para el esquema en adelante (FFD del inglés forward-finite-differencing) de n_u perturbaciones en cada iteración de RTO [52, 53]. Es posible también utilizar puntos pasados para estimar los gradientes, esto puede observarse en ciertos trabajos para los métodos de ISOPE dual o de adaptación de modificadores dual [55, 56]. Otra

manera de aprovechar los puntos pasados para la estimación de los gradientes es mediante el esquema Broyden [57]. Recientemente Gao et al. [58] propusieron calcular los modificadores para los esquemas de optimización en tiempo real mediante la utilización de aproximaciones cuadráticas basadas en puntos operativos pasados de la función costo y de las restricciones de la planta.

Además de un apropiado método para el cálculo de gradientes, comentamos en la Sección 1.2.3.2 la necesidad de disponer de un modelo adecuado para los esquemas de adaptación por modificadores. Según Marchetti et al. [54] la adecuación del modelo para la utilización de adaptación por modificadores de primer orden puede traducirse en la siguiente proposición:

Proposición 1. Sea \mathbf{u}_p^* un punto regular para las restricciones y el único óptimo de la planta. Sea $\nabla_r^2 \mathcal{L}(\mathbf{u}_p^*)$ el Hessiano reducido del Lagrangiano del problema de optimización en el punto \mathbf{u}_p^* . Entonces:

(i) Si $\nabla_r^2 \mathcal{L}(\mathbf{u}_p^*)$ es positivo definido, entonces el modelo del proceso es adecuado para el uso del método de adaptación por modificadores.

(ii) Si $\nabla_r^2 \mathcal{L}(\mathbf{u}_p^*)$ no es positivo semi-definido, entonces el modelo del proceso es inadecuado para el uso del método de adaptación por modificadores.

(iii) Si $\nabla_r^2 \mathcal{L}(\mathbf{u}_p^*)$ es positivo semi-definido, entonces las condiciones de segundo orden no son concluyentes con respecto a la adecuación del modelo para el uso del método de adaptación por modificadores.

La utilización de modificadores para el hessiano del modelo (denominados modificadores de segundo orden) nos permitiría, en principio, poder utilizar modelos que resulten inadecuados según la Proposición 1 para esquemas de adaptación por modificadores de primer orden.

De manera similar a la definición de los modificadores de orden cero y de primer orden, se puede definir modificadores de segundo orden representados por la diferencia entre las matrices hessianas de la planta y del modelo, es decir:

$$\mathbf{H}_{i+1}^J = \left. \frac{\partial^2 J_p}{\partial \mathbf{u}^2} \right|_{\mathbf{u}_{i+1}} - \left. \frac{\partial^2 J}{\partial \mathbf{u}^2} \right|_{\mathbf{u}_{i+1}} \quad (2.46)$$

$$\mathbf{H}_{i+1}^G = \left. \frac{\partial^2 G_p}{\partial \mathbf{u}^2} \right|_{\mathbf{u}_{i+1}} - \left. \frac{\partial^2 G}{\partial \mathbf{u}^2} \right|_{\mathbf{u}_{i+1}} \quad (2.47)$$

Entonces se puede completar el problema adaptado (2.45) como:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^J + (\boldsymbol{\lambda}_i^J)^T (\mathbf{u} - \mathbf{u}_i) + \frac{1}{2} (\mathbf{u} - \mathbf{u}_i)^T \mathbf{H}_i^J (\mathbf{u} - \mathbf{u}_i) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^G + (\boldsymbol{\lambda}_i^G)^T (\mathbf{u} - \mathbf{u}_i) + \frac{1}{2} (\mathbf{u} - \mathbf{u}_i)^T \mathbf{H}_i^G (\mathbf{u} - \mathbf{u}_i) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (2.48)$$

En [53], los autores proponen resolver el Problema (2.48) estimando el hessiano de la diferencia planta-modelo mediante el cálculo y la actualización simétrica de rango uno (SR1 de symmetric rank-one) y los gradientes mediante diferencias finitas.

En esta tesis se propone en el Capítulo 5 un esquema basado en modificadores de segundo orden obtenidos de una aproximación cuadrática de la diferencia planta-modelo que utiliza puntos operativos pasados de estado estacionario.

Capítulo 3

Selección de Estructuras de Control Auto Optimizantes con Márgenes de Seguridad de Estado Estacionario

3.1 Introducción

La selección de la estructura de control (CS) implica determinar las variables controladas (CVs), las variables manipuladas (MVs), la política de control (descentralizada o centralizada), la tecnología del controlador (clásica o avanzada), el emparejamiento entre las CVs y las MVs en el caso de un controlador descentralizado, y el ajuste de los controladores. Históricamente, el diseño de un sistema de control de planta completa se ha basado en la heurística, la comprensión matemática y el conocimiento del proceso [59–61].

Esta tesis considera el problema de la selección de la estructura de control basado en criterios económicos y en presencia de perturbaciones y restricciones de operación. Este problema ha sido estudiado por muchos investigadores en los últimos 35 años, utilizando diferentes estrategias para incorporar criterios económicos en el problema de diseño, o para combinar objetivos económicos con otros objetivos como la estabilidad y controlabilidad, y utilizando diferentes simplificaciones y aproximaciones para formular el problema de selección de estructura de control. Existe también una amplia literatura en la que el problema de selección de estructura de control se ha integrado con el problema de diseño y dimensionamiento del proceso [3, 27, 62]. A su vez, dado que el desempeño dinámico de un proceso depende en gran medida de su diseño y debido a los conflictos entre la economía y la controlabilidad de los procesos químicos, existen fuertes incentivos para integrar el diseño de la planta y la selección de la estructura de control [62].

Dos estrategias para la selección de la estructura de control basadas en la economía operativa del proceso que han sido estudiadas independientemente por diferentes investigadores son la *metodología de márgenes de seguridad* y el *control auto-optimizante*.

3.1.1 Contribuciones de este Capítulo

La estrategia de selección de estructuras de control en base a criterios económicos presentada en esta sección se basa en la idea de los márgenes de seguridad para la operación en estado estacionario. Por lo tanto, nos preocupamos por evitar violaciones de restricciones para la operación en estado estacionario y en presencia de perturbaciones. Se presenta un estudio comparativo analizando la diferencia entre

minimizar el costo nominal y el costo promedio. Se observa que el enfoque de márgenes de seguridad de estado estacionario y el control auto optimizante comparten los mismos objetivos, y se argumenta que el método de los márgenes de seguridad con minimización del costo promedio representa un método de control auto optimizante global. Los aportes de este capítulo son los siguientes:

- Mediante casos de estudio se presenta una comparación entre los métodos de márgenes de seguridad de estado estacionario con minimización del costo nominal y con minimización del costo promedio.
- Dado que para el propósito de la selección de la estructura de control, minimizar el costo promedio es equivalente a minimizar la pérdida de optimalidad promedio, se argumenta que el problema de márgenes de seguridad en estado estacionario con minimización del costo promedio es un método de control auto optimizante global. Esto se verifica mediante un estudio comparativo sobre el funcionamiento de un reactor tanque continuo agitado (CSTR), donde se compara el desempeño del enfoque propuesto con el de un método SOC local (el método del espacio nulo).
- Se consideraron dos formulaciones diferentes del problema de márgenes de seguridad en estado estacionario con minimización del costo promedio. La primera formulación consiste en un MINLP que selecciona la mejor estructura de control clásica, mientras que la segunda formulación es un NLP que selecciona combinaciones lineales óptimas de variables de entrada y salida como variables controladas. Ambas formulaciones son métodos SOC globales que encuentran valores de consigna óptimos para las variables controladas y para las entradas fijas que no se utilizan como variables manipuladas, al tiempo que garantizan la satisfacción de las restricciones para un conjunto de escenarios de perturbaciones. La formulación MINLP guarda similitudes con la formulación para la selección de la estructura de control basada en la economía de operación propuesta en Kookos [8].
- Se incluye en la formulación MINLP restricciones lineales adicionales que se utilizan para evaluar la matriz de ganancia relativa (RGA) de la estructura de control seleccionada. Se muestra que estas restricciones adicionales pueden ayudar a evitar la obtención de una estructura de control inadecuada.
- Se presentan dos aproximaciones alternativas de la formulación MINLP como problemas de programación cuadrática mixta entera (MIQP) convexa. La primera aproximación MIQP es local en las entradas y global en las perturbaciones, mientras que la segunda aproximación MIQP es local tanto en las entradas como en las perturbaciones.
- Mediante tres casos de estudio, se compara el desempeño obtenido por la mejor estructura de control clásica con el desempeño de una estructura de control de combinación lineal óptima. Mientras que en la literatura de SOC la selección de estructuras de control empleando combinaciones lineales ha estado enfocada en casos de estudio donde se optimizan las variables de entrada restantes luego de controlar las restricciones activas, en nuestro primer ejemplo mostramos que también son útiles en los casos en que el óptimo está completamente determinado por la intersección de restricciones activas, y el conjunto de restricciones activas cambia con los valores de las perturbaciones.
- En base al estudio de los resultados obtenidos, se proporcionan una serie de recomendaciones para la implementación de métodos de selección de estructuras de control basados en márgenes de seguridad de estado estacionario y control auto optimizante.

3.2 Selección de la Mejor Estructura de Control Clásica

El problema de selección de la estructura de control clásica se puede escribir (sin simplificaciones ni aproximaciones) de la siguiente manera:

$$\begin{aligned}
& \min_{\mathbf{z}, \mathbf{r}^{\text{SP}}} \mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}, \mathbf{d}) \right\} \\
& \text{s.t.} \quad \mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y}(\mathbf{u}, \mathbf{d})) = \mathbf{r}^{\text{SP}}(\mathbf{z}), \quad \forall \mathbf{d} \in \mathcal{D} \\
& \quad \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^{\text{U}}, \quad \forall \mathbf{d} \in \mathcal{D} \\
& \quad \mathbf{u}^{\text{L}} \leq \mathbf{u} \leq \mathbf{u}^{\text{U}},
\end{aligned} \tag{3.1}$$

Note que la diferencia entre este problema y el problema de los márgenes de seguridad en estado estacionario a lazo cerrado (2.18) es que se minimiza el costo esperado en lugar de minimizar el costo para los valores nominales de las perturbaciones. De este modo la estructura buscada y los valores de consigna se seleccionan en función de optimizar el rendimiento económico para todo el rango de perturbaciones. Como se verá en los casos de estudio, esta diferencia convierte el problema de márgenes de seguridad en un problema SOC global.

3.2.1 Formulación MINLP

Para la formulación del problema con estructuras de control clásicas podemos utilizar el vector de variables binarias \mathbf{z} , y el vector de valores de consigna \mathbf{r}^{SP} para las variables de entrada y salida. Dichos vectores se parametrizarán de la siguiente manera:

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}^{\text{O}} \\ \mathbf{z}^{\text{I}} \end{bmatrix}, \quad \mathbf{z}^{\text{O}} \in \mathbb{B}^{n_y}, \quad \mathbf{z}^{\text{I}} \in \mathbb{B}^{n_u}, \tag{3.2}$$

$$\mathbf{r}^{\text{SP}} = \begin{bmatrix} \mathbf{y}^{\text{SP}} \\ \mathbf{u}^{\text{SP}} \end{bmatrix}, \quad \mathbf{y}^{\text{SP}} \in \mathbb{R}^{n_y}, \quad \mathbf{u}^{\text{SP}} \in \mathbb{R}^{n_u}. \tag{3.3}$$

Los componentes binarios con valores iguales a uno en los vectores \mathbf{z}^{O} y \mathbf{z}^{I} corresponden a las variables de salida controladas seleccionadas y las variables de entrada fijas, respectivamente. Los componentes restantes en \mathbf{z}^{I} con valores iguales a cero corresponden a las variables manipuladas seleccionadas.

Los problemas (2.18) y (3.1) se aproximan considerando un número finito de valores del vector de perturbaciones (escenarios de perturbaciones). Sea $\underline{\mathcal{D}}$ una discretización representativa de dimensión N del conjunto de perturbaciones \mathcal{D} , es decir, $\underline{\mathcal{D}} = \{\mathbf{d}_1, \dots, \mathbf{d}_N\}$.

Por conveniencia, definimos ahora los siguientes conjuntos de índices:

$$\mathcal{U} = \{1, 2, \dots, n_u\}, \quad \mathcal{Y} = \{1, 2, \dots, n_y\}, \quad \mathcal{K} = \{1, 2, \dots, N\}.$$

El problema de selección de la estructura de control clásica se formula como un MINLP de la

siguiente manera:

$$\min_{\mathbf{z}, \mathbf{r}^{\text{SP}}, \mathbf{u}} \sum_{k=1}^N p_k J(\mathbf{u}_k, \mathbf{d}_k) \quad (3.4a)$$

$$\text{s.t.} \quad \sum_{i=1}^{n_y} z_i^{\text{O}} + \sum_{j=1}^{n_u} z_j^{\text{I}} = n_u \quad (3.4b)$$

$$y_i^{\text{L}} z_i^{\text{O}} \leq y_i^{\text{SP}} \leq y_i^{\text{U}} z_i^{\text{O}}, \quad \forall i \in \mathcal{Y}, \quad (3.4c)$$

$$u_j^{\text{L}} z_j^{\text{I}} \leq u_j^{\text{SP}} \leq u_j^{\text{U}} z_j^{\text{I}}, \quad \forall j \in \mathcal{U}, \quad (3.4d)$$

$$y_i^{\text{L}}(1 - z_i^{\text{O}}) \leq y_i(\mathbf{u}_k, \mathbf{d}_k) - y_i^{\text{SP}} \leq y_i^{\text{U}}(1 - z_i^{\text{O}}), \quad \forall i \in \mathcal{Y}, \quad k \in \mathcal{K}, \quad (3.4e)$$

$$u_j^{\text{L}}(1 - z_j^{\text{I}}) \leq u_{j,k} - u_j^{\text{SP}} \leq u_j^{\text{U}}(1 - z_j^{\text{I}}), \quad \forall j \in \mathcal{U}, \quad k \in \mathcal{K}, \quad (3.4f)$$

El valor esperado de la función objetivo se reemplaza por una suma ponderada sobre el conjunto especificado de escenarios de perturbación, donde p_k son las ponderaciones para cada escenario de perturbación. Si las realizaciones de las perturbaciones se distribuyen uniformemente en \mathcal{D} , y $\underline{\mathcal{D}}$ es una discretización uniforme de \mathcal{D} , entonces tendremos $p_k = \frac{1}{N}$ para $k = 1, \dots, N$. La ecuación (3.4b) establece que el número total de variables fijadas debe ser igual al número total de variables de entrada. El número de variables binarias iguales a uno en \mathbf{z}^{O} define la dimensión del sistema de control (es decir, el número de lazos de control). Téngase en cuenta que el problema de control siempre es cuadrado, es decir, el número de variables manipuladas es igual al número de variables de salida controladas. Por medio de (3.4c) y (3.4d), los valores de consigna de las variables de salida controladas y los valores de las variables de entrada fijas están restringidos a estar dentro de sus valores límites correspondientes, y si las variables no están controladas o fijas, sus valores en el vector \mathbf{r}^{SP} se establecen arbitrariamente en cero. Las restricciones (3.4e) y (3.4f) obligan a las variables de salida controladas y las variables de entrada fijas a coincidir con sus valores de consigna, mientras que las salidas no controladas y las entradas manipuladas tendrán que respetar sus respectivos valores límites, para todos los escenarios de perturbación $\mathbf{d}_k \in \underline{\mathcal{D}}$.

Observación 2. El problema de los márgenes de seguridad en estado estacionario con minimización de costos nominales (2.18) se puede formular como un MINLP de una manera similar, reemplazando el costo promedio en (3.4a) por el costo nominal $J(\mathbf{u}_n, \mathbf{d}_n)$. Se puede encontrar una formulación similar en Kookos y Perkins [22].

3.2.2 Restricciones Adicionales

Número fijo de lazos de control. El Problema (3.4) puede resolverse para un número fijo de lazos de control $n_q \in \mathbb{I}^+$, $1 \leq q \leq \min\{n_y, n_u\}$, mediante la adición de la siguiente restricción:

$$\sum_{i=1}^{n_y} z_i^{\text{O}} = n_q. \quad (3.5)$$

Emparejamiento entrada-salida basado en la matriz RGA.

Dentro del contexto de diseño de estructuras de control, Braccia et al. [63] propusieron limitar los valores de la matriz de ganancias relativas (RGA) a modo de asegurar que, para la estructura de control seleccionada, exista un emparejamiento de entrada-salida que de como resultado un grado aceptable de interacción cuando se implementa el control descentralizado.

Sea $\mathbf{G} \in \mathbb{R}^{n_y \times n_u}$ la matriz de ganancia de entrada-salida completa del sistema a lazo abierto, y sea $\mathbf{G}_s(\mathbf{z}) \in \mathbb{R}^{n_q \times n_q}$ la sub-matriz cuadrada determinada por las variables de salida controladas seleccionadas (y_i para las cuales $z_i^O = 1$) y las variables manipuladas (u_j para las cuales $z_j^I = 0$). Se propone evaluar las siguientes restricciones [63]:

$$\Lambda(\mathbf{z}) = \mathbf{G}_s(\mathbf{z}) \otimes [\mathbf{G}_s^{-1}(\mathbf{z})]^\top \quad (3.6a)$$

$$\Lambda_p(\mathbf{z}, \mathbf{Z}^p) = \Lambda(\mathbf{z}) \otimes \mathbf{Z}^p(\mathbf{z}) \quad (3.6b)$$

$$\delta^L \leq \sum_{j=1}^{n_q} \lambda_{p,i,j}(\mathbf{z}, \mathbf{Z}^p(\mathbf{z})) \leq \delta^U, \quad i = 1, \dots, n_q \quad (3.6c)$$

$$\sum_{i=1}^{n_q} z_{i,j}^p = \sum_{j=1}^{n_q} z_{i,j}^p = 1 \quad (3.6d)$$

donde $\Lambda(\mathbf{z})$ es la matriz de ganancia relativa correspondiente a la estructura de control \mathbf{z} ; $\mathbf{Z}^p \in \mathbb{B}^{n_q \times n_q}$ es una matriz de variables de decisión binarias que determina el emparejamiento entrada-salida; la matriz $\Lambda_p \in \mathbb{R}^{n_q \times n_q}$ contiene los elementos de Λ seleccionados por \mathbf{Z}^p y ceros en las otras posiciones; $\lambda_{p,i,j}$ son los elementos de Λ_p , y $z_{i,j}^p$ son los elementos de \mathbf{Z}^p . En (3.6c), los límites inferior y superior de las ganancias relativas, δ^L y δ^U (parámetros de diseño) permiten hacer cumplir un grado deseado de interacción en una estructura de control descentralizada.

Dado que el Problema (3.4) utiliza un modelo no lineal del sistema, proponemos aplicar las restricciones (3.6) linealizando el modelo en el óptimo nominal. Es decir, seleccionamos $\mathbf{G} = \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \Big|_n$ como la matriz de ganancia de estado estacionario calculada en el óptimo nominal $\mathbf{u}^*(\mathbf{d}_n)$. En general, esto no garantiza que los límites (3.6c) se satisfagan para valores de las perturbaciones diferentes a los nominales, o cuando se opere lejos del óptimo nominal debido a la presencia de los márgenes de seguridad. En la práctica, sin embargo, la evaluación de estos límites en el óptimo nominal ya permite descartar algunas estructuras de control inapropiadas.

Braccia et al. [63] propusieron una reformulación de (3.6) como un conjunto de restricciones lineales únicamente. A continuación, se adapta esta reformulación para que las restricciones resultantes se ajusten al Problema (3.4). La reformulación requiere introducir las siguientes variables y parámetros:

Variables Continuas: inversa de la matriz de ganancia a lazo cerrado $\tilde{\mathbf{G}} \in \mathbb{R}^{n_u \times n_y}$ con los elementos $\tilde{g}_{j,i}$; matriz auxiliar $\tilde{\mathbf{B}} \in \mathbb{R}^{n_y \times n_y}$ con elementos $\tilde{b}_{l,i}$; matriz auxiliar $\Lambda^r \in \mathbb{R}^{n_y \times n_u}$ con elementos $\lambda_{i,j}^r$.

Variables Binarias: matriz $\mathbf{Z}^{IO} \in \mathbb{B}^{n_y \times n_u}$ con elementos $z_{i,j}^{IO}$.

Parámetros: matriz de ganancia a lazo abierto $\mathbf{G} \in \mathbb{R}^{n_y \times n_u}$ con elementos $g_{i,j}$; valores grandes M ; cota inferior δ^L de los elementos de la matriz RGA y cota superior δ^U de los elementos de la matriz RGA .

Estas variables y parámetros adicionales se incluyen en el Problema (3.4) junto con las siguientes

restricciones lineales:

$$-M(1 - z_j^I) \leq \tilde{g}_{j,i} \leq M(1 - z_j^I) \quad \forall j \in \mathcal{U}, \quad i \in \mathcal{Y}, \quad (3.7a)$$

$$-M(z_i^O) \leq \tilde{g}_{j,i} \leq M(z_i^O) \quad \forall j \in \mathcal{U}, \quad i \in \mathcal{Y}, \quad (3.7b)$$

$$-M(2 - z_l^O - z_i^O) \leq \tilde{b}_{l,i} \leq M(2 - z_l^O - z_i^O) \quad \forall l \in \mathcal{Y}, \quad i \in \mathcal{Y}, \quad (3.7c)$$

$$\sum_j g_{l,j} \tilde{g}_{j,i} = \tilde{b}_{l,i} + I_{l,i} z_i^O \quad \forall l \in \mathcal{Y}, \quad i \in \mathcal{Y}, \quad (3.7d)$$

$$\sum_j z_{i,j}^{IO} = z_i^O \quad \forall i \in \mathcal{Y}, \quad (3.7e)$$

$$\sum_i z_{i,j}^{IO} = (1 - z_j^I) \quad \forall j \in \mathcal{U}, \quad (3.7f)$$

$$-M(1 - z_{i,j}^{IO}) \leq \lambda_{i,j}^r - g_{i,j} \tilde{g}_{j,i} \leq M(1 - z_{i,j}^{IO}) \quad \forall i \in \mathcal{Y}, \quad j \in \mathcal{U}, \quad (3.7g)$$

$$\delta^L z_{i,j}^{IO} \leq \lambda_{i,j}^r \leq \delta^U z_{i,j}^{IO} \quad \forall i \in \mathcal{Y}, \quad j \in \mathcal{U}. \quad (3.7h)$$

Como mencionamos, $\mathbf{G}_s(\mathbf{z})$ es la submatriz (cuadrada) de \mathbf{G} obtenida para una estructura de control particular \mathbf{z} . Sea $\tilde{\mathbf{G}}_s(\mathbf{z})$ la submatriz correspondiente de $\tilde{\mathbf{G}}$. Las ecuaciones (3.7a) y (3.7b) llevan a cero todos los elementos de la matriz $\tilde{\mathbf{G}}$ a excepción de aquellos correspondientes a la submatriz seleccionada $\tilde{\mathbf{G}}_s(\mathbf{z})$. Considere la matriz cuadrada $\mathbf{B} = \mathbf{G}\tilde{\mathbf{G}}$, y sea $\mathbf{B}_s(\mathbf{z})$ la submatriz de \mathbf{B} correspondiente a las variables de salida seleccionadas a controlar \mathbf{z}^O . Las ecuaciones (3.7c) y (3.7d) son usadas para establecer $\mathbf{G}_s(\mathbf{z})\tilde{\mathbf{G}}_s(\mathbf{z}) = \mathbf{B}_s(\mathbf{z}) = \mathbf{I}$. Esto nos permite obtener la matriz inversa $\tilde{\mathbf{G}}_s(\mathbf{z})$ para cualquier estructura de control \mathbf{z} mediante ecuaciones lineales. Note que, por medio de (3.7d) tendremos $\mathbf{G}\tilde{\mathbf{G}} = \mathbf{B} = \tilde{\mathbf{B}} + \text{diag}(\mathbf{z}^O)$. Las ecuaciones (3.7e) y (3.7f) son usadas para construir la matriz de emparejamiento entrada-salida \mathbf{Z}^{IO} . Note que la matriz $\mathbf{Z}^P(\mathbf{z})$ en (3.6) es la submatriz cuadrada de \mathbf{Z}^{IO} correspondiente a la estructura de control \mathbf{z} . En (3.7g), los elementos de la matriz RGA son calculados y alojados en los elementos $\lambda_{i,j}^r$ de la matriz Λ^r en base al emparejamiento seleccionado \mathbf{Z}^{IO} . Estos elementos están restringidos a satisfacer los límites definidos en (3.7h). Observar que Λ_p en (3.6) es la submatriz cuadrada de Λ^r correspondiente a la estructura de control \mathbf{z} y al emparejamiento \mathbf{Z}^{IO} . El siguiente ejemplo ilustra la implementación de las restricciones (3.7) para una estructura de control definida.

Ejemplo 5 (Límites en la matriz RGA). Consideremos un sistema con cuatro entradas ($n_u = 4$) y seis salidas medibles ($n_y = 6$), y consideremos la estructura de control definida por $\mathbf{z}^O = [0 \ 1 \ 0 \ 1 \ 0 \ 0]$ y $\mathbf{z}^I = [0 \ 1 \ 1 \ 0]$. Para esta estructura de control de dimension $n_q = 2$ las variables controladas son y_2 y y_4 y las variables manipuladas son u_1 y u_4 . Por medio de (3.7a) y (3.7b) la matriz $\tilde{\mathbf{G}}$ toma la siguiente forma:

$$\tilde{\mathbf{G}} = \begin{bmatrix} 0 & \tilde{g}_{1,2} & 0 & \tilde{g}_{1,4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tilde{g}_{4,2} & 0 & \tilde{g}_{4,4} & 0 & 0 \end{bmatrix}.$$

La multiplicación matricial $\mathbf{G}\tilde{\mathbf{G}} = \mathbf{B}$ nos lleva a:

$$\mathbf{B} = \begin{bmatrix} 0 & b_{1,2} & 0 & b_{1,4} & 0 & 0 \\ 0 & b_{2,2} & 0 & b_{2,4} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & b_{6,2} & 0 & b_{6,4} & 0 & 0 \end{bmatrix},$$

con $b_{i,j} = g_{i,1}\tilde{g}_{1,j} + g_{i,4}\tilde{g}_{4,j}$. Esta matriz captura la información de la multiplicación del subproceso seleccionado en los elementos $b_{2,2}$, $b_{2,4}$, $b_{4,2}$, y $b_{4,4}$. Para visualizar esto, considere las sub matrices \mathbf{G}_s , $\tilde{\mathbf{G}}_s$ y \mathbf{B}_s para las variables controladas y manipuladas seleccionadas:

$$\mathbf{G}_s = \begin{bmatrix} g_{2,1} & g_{2,4} \\ g_{4,1} & g_{4,4} \end{bmatrix}, \quad \tilde{\mathbf{G}}_s = \begin{bmatrix} \tilde{g}_{1,2} & \tilde{g}_{1,4} \\ \tilde{g}_{4,2} & \tilde{g}_{4,4} \end{bmatrix}, \quad \mathbf{G}_s\tilde{\mathbf{G}}_s = \mathbf{B}_s = \begin{bmatrix} b_{2,2} & b_{2,4} \\ b_{4,2} & b_{4,4} \end{bmatrix}.$$

La multiplicación matricial $\mathbf{G}_s\tilde{\mathbf{G}}_s$ se lleva a cabo en (3.7d) con $\mathbf{B} = \tilde{\mathbf{B}} + \text{diag}(\mathbf{z}^O)$. Las restricciones para $\tilde{\mathbf{B}}$ en (3.7c) implican que:

$$\mathbf{B} = \begin{bmatrix} 0 & b_{1,2} & 0 & b_{1,4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & b_{3,2} & 0 & b_{3,4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & b_{5,2} & 0 & b_{5,4} & 0 & 0 \\ 0 & b_{6,2} & 0 & b_{6,4} & 0 & 0 \end{bmatrix},$$

de donde puede observarse que $\mathbf{B}_s = \mathbf{I}$, y por lo tanto, $\tilde{\mathbf{G}}_s$ es la inversa de \mathbf{G}_s .

El emparejamiento basado en la matriz binaria \mathbf{Z}^{IO} tiene, para este caso, dos alternativas en base a la estructura de control definida por \mathbf{z}^O y \mathbf{z}^I :

$$\mathbf{Z}_a^{\text{IO}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \text{obien} \quad \mathbf{Z}_b^{\text{IO}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Estas alternativas posibles se generan mediante las ecuaciones (3.7e) y (3.7f), e implica que para \mathbf{Z}_a^{IO} , por ejemplo, los lazos propuestos son $y_2 - u_1$ y $y_4 - u_4$. Los elementos de la matriz RGA resultantes ($\lambda_{i,j}^r = g_{i,j}\tilde{g}_{j,i}$) son almacenados en la matriz auxiliar Λ^r por medio de la ecuación (3.7g), y sus valores están restringidos a satisfacer los límites predefinidos en (3.7h). Por ejemplo, para el emparejamiento \mathbf{Z}_a^{IO} tendremos:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ \delta^L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta^L \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leq \Lambda_a^r = \begin{bmatrix} 0 & 0 & 0 & 0 \\ g_{2,1}\tilde{g}_{1,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_{4,4}\tilde{g}_{4,4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \leq \begin{bmatrix} 0 & 0 & 0 & 0 \\ \delta^U & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta^U \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (3.8)$$

3.2.3 Estrategia de Solución

El Problema MINLP (3.4) se puede programar en GAMS y utilizar los solvers disponibles como BARON u otros. Sin embargo, la formulación puede ser difícil de resolver para modelos de procesos complejos. La principal limitación del Problema (3.4) proviene de las no linealidades y no convexidades en las ecuaciones del modelo utilizadas para calcular los mapeos $J(\mathbf{u}, \mathbf{d})$ y $y_i(\mathbf{u}, \mathbf{d})$, $i = 1, \dots, n_y$. Además, el problema se vuelve engorroso a medida que aumenta la discretización del número de escenarios.

Una vez que se determina la estructura de control óptima (\mathbf{z}^*), el siguiente paso es seleccionar la interacción y la tecnología del controlador. A su vez el sistema de control de toda la planta resultante debe validarse mediante pruebas de estabilidad, controlabilidad y robustez, posiblemente requiriendo simulaciones dinámicas. Si la estructura de control óptima resultante ha sido validada, se puede aplicar a la planta. Por el contrario, si se encuentra que la estructura de control óptima no es adecuada, se introduce un corte en las variables enteras correspondientes y el Problema (3.4) se resuelve nuevamente hasta que se obtiene una estructura de control apropiada.

En muchos casos, el uso de una discretización gruesa de las perturbaciones en el problema MINLP es suficiente para revelar la estructura de control óptima correcta. Sin embargo, debido a la aproximación, los valores de consigna obtenidos pueden no ser adecuados. A fin de evitar el riesgo de violar las restricciones para algunos escenarios de perturbación, se sugiere que los valores de consigna vuelvan a calcularse fijando la estructura de control \mathbf{z} en el Problema (3.4) según la estructura óptima obtenida, y luego resolviendo el Problema (3.4) como un NLP usando una discretización mas fina del conjunto de perturbaciones \mathcal{D} .

3.3 Selección de Estructuras de Control con Combinaciones Lineales

El problema de seleccionar combinaciones lineales de variables de entrada y salida como variables controladas se puede formular de la siguiente manera:

$$\begin{aligned}
& \underset{\mathcal{Q}, \mathbf{r}^{\text{sp}}}{\text{mín}} && \mathbb{E}_{\mathbf{d} \in \mathcal{D}} \left\{ J(\mathbf{u}, \mathbf{d}) \right\} \\
& \text{s.t.} && \mathbf{r} = \mathcal{Q}^\top \begin{bmatrix} \mathbf{y}(\mathbf{u}, \mathbf{d}) \\ \mathbf{u} \end{bmatrix} = \mathbf{r}^{\text{sp}}, && \forall \mathbf{d} \in \mathcal{D} \\
& && \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^{\text{U}}, && \forall \mathbf{d} \in \mathcal{D} \\
& && \mathbf{u}^{\text{L}} \leq \mathbf{u} \leq \mathbf{u}^{\text{U}},
\end{aligned} \tag{3.9}$$

donde $\mathbf{r} \in \mathbb{R}^{n_u}$ son las variables controladas, $\mathcal{Q} \in \mathbb{R}^{(n_u+n_y) \times n_u}$ es la matriz de combinaciones lineales, y $\mathbf{r}^{\text{sp}} \in \mathbb{R}^{n_u}$ son los valores de consigna. Tras la discretización del conjunto de perturbaciones \mathcal{D} , el Problema (3.9) puede aproximarse como un NLP:

$$\underset{\mathcal{Q}, \mathbf{r}^{\text{sp}}, \mathbf{u}}{\text{mín}} \sum_{k=1}^N p_k J(\mathbf{u}_k, \mathbf{d}_k) \tag{3.10a}$$

$$\text{s.t.} \quad \mathcal{Q}^\top \begin{bmatrix} \mathbf{y}(\mathbf{u}_k, \mathbf{d}_k) \\ \mathbf{u}_k \end{bmatrix} = \mathbf{r}^{\text{sp}}, \quad \forall k \in \mathcal{K}, \tag{3.10b}$$

$$\mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}_k, \mathbf{d}_k) \leq \mathbf{y}^{\text{U}}, \quad \forall k \in \mathcal{K}, \tag{3.10c}$$

$$\mathbf{u}^{\text{L}} \leq \mathbf{u}_k \leq \mathbf{u}^{\text{U}}, \quad \forall k \in \mathcal{K}. \tag{3.10d}$$

La selección de la matriz \mathcal{Q} conducirá en general a infinitas soluciones posibles. Esto se puede evitar imponiendo una estructura predefinida y parametrizando la matriz \mathcal{Q} (se verán ejemplos de esto en las subsecciones 3.5.1 y 3.5.2).

Observación 3. En el contexto de SOC, Yelchuru y Skogestad [35] formularon un problema de optimización mixto entero que encuentra el mejor subconjunto de variables medibles para las combinaciones lineales. Nótese que se podría implementar un enfoque similar en el Problema (3.10), que de este modo se convertiría en un MINLP.

Observación 4. A nivel práctico, al resolver el problema NLP en busca de la estructura de control con combinaciones lineales se utilizan variables continuas y no se descarta en principio que la combinación lineal óptima resulte ser finalmente una estructura de control de tipo clásica (ver definición 1).

3.3.1 Restricciones Adicionales

Opcionalmente, se puede agregar al Problema (3.10) la siguiente restricción:

$$\mathcal{Q}^T \begin{bmatrix} \mathbf{G} \\ \mathbf{I} \end{bmatrix} = \mathbf{T}, \quad (3.11)$$

donde $\mathbf{G} \in \mathbb{R}^{n_y \times n_u}$ es la matriz de ganancia de entrada-salida del sistema a lazo abierto, evaluada en el óptimo nominal, y $\mathbf{T} \in \mathbb{R}^{n_u \times n_u}$ es la matriz de ganancia impuesta al proceso controlado. En la literatura de SOC, \mathbf{T} ha sido seleccionada como la matriz de identidad [30, 34], o como la raíz cuadrada de la matriz Hessiana de la función objetivo (costo) [35, 37] (dicha selección responde a simplificar una formulación SOC local propuesta por los autores). También puede seleccionarse \mathbf{T} como cualquier permutación de la matriz identidad o, en general, como cualquier matriz bien condicionada. El uso de la restricción (3.11) se ilustrará en la Subsección 3.5.3. Tenga en cuenta que la elección de \mathbf{T} también determinará el emparejamiento de las variables manipuladas con las variables controladas.

La inclusión de la restricción (3.11) es ventajosa por tres razones: (i) restringe considerablemente el espacio de búsqueda, (ii) impone la invertibilidad del proceso controlado (al menos en condiciones nominales), lo cual está vinculado con el concepto de controlabilidad de estado estacionario, y (iii) no afecta el costo óptimo del Problema (3.10). Esta última afirmación se puede probar como sigue. Supongamos que \mathcal{Q}_1 con valores de consigna \mathbf{r}_1^{sp} es una solución óptima del Problema (3.10). Entonces, para cualquier matriz no singular \mathbf{C} , $\mathcal{Q}^T = \mathbf{C}\mathcal{Q}_1^T$ con valores de consigna $\mathbf{r}^{\text{sp}} = \mathbf{C}\mathbf{r}_1^{\text{sp}}$ también es una solución óptima con el mismo costo óptimo. La matriz de ganancia del proceso controlado $\mathbf{B} = \mathcal{Q}_1^T[\mathbf{G}; \mathbf{I}]$ debe ser invertible para la controlabilidad en estado estacionario. Por lo tanto, si \mathbf{T} también es no singular, se puede seleccionar la matriz no singular $\mathbf{C} = \mathbf{T}(\mathbf{B})^{-1}$ tal que $\mathbf{CB} = \mathbf{T}$.

3.4 Aproximación como un Problema MIQP

Una de las principales limitaciones del Problema MINLP (3.4) es la dificultad para encontrar el óptimo global para problemas no convexos. La fuente de no linealidades y no convexidades en este problema proviene de las funciones $J(\mathbf{u}, \mathbf{d})$ y $y_i(\mathbf{u}, \mathbf{d})$, $i = 1, \dots, n_y$. En esta sección, evitamos esta dificultad al aproximar el Problema (3.4) como un problema MIQP convexo. A continuación se propondrán dos aproximaciones MIQP alternativas. La primera, que denotamos lugd-MIQP, es local en las entradas \mathbf{u} y global en las perturbaciones \mathbf{d} , mientras que la segunda, denotada luld-MIQP, es local tanto en las entradas como en las perturbaciones.

3.4.1 MIQP Global en Perturbaciones

Para una perturbación dada $\mathbf{d}_k \in \mathcal{D}$, consideramos la aproximación QP utilizada en los métodos de programación cuadrática sucesiva (SQP) para resolver problemas NLP [10, 64]. La aproximación QP

del Problema NLP (2.5) en el punto óptimo $\mathbf{u}^*(\mathbf{d}_k)$ está dada por:

$$\min_{\delta \mathbf{u}_k} \left. \frac{\partial J}{\partial \mathbf{u}} \right|_k \delta \mathbf{u}_k + \frac{1}{2} \delta \mathbf{u}_k^\top \mathbf{H}_k \delta \mathbf{u}_k \quad (3.12a)$$

$$\text{s.t. } y_i^L \leq y_{i,k} + \left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_k \delta \mathbf{u}_k \leq y_i^U, \quad i = 1, \dots, n_y, \quad (3.12b)$$

$$\mathbf{u}_k = \mathbf{u}^*(\mathbf{d}_k) + \delta \mathbf{u}_k, \quad (3.12c)$$

$$\mathbf{u}^L \leq \mathbf{u}_k \leq \mathbf{u}^U, \quad (3.12d)$$

donde $\left. \frac{\partial J}{\partial \mathbf{u}} \right|_k$ y $\left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_k$ son evaluadas en el punto óptimo nominal $\mathbf{u}^*(\mathbf{d}_k)$, $y_{i,k} = y_i(\mathbf{u}^*(\mathbf{d}_k), \mathbf{d}_k)$, y \mathbf{H}_k es una aproximación definida positiva del Hessiano de la función Lagrangiana $\left. \frac{\partial^2 L}{\partial \mathbf{u}^2} \right|_k$, evaluada en $\mathbf{u}^*(\mathbf{d}_k)$.

En esta formulación, las restricciones de las variables de salida se linealizan en $\mathbf{u}^*(\mathbf{d}_k)$, y la curvatura de las restricciones se toma en cuenta utilizando el Hessiano de la función Lagrangiana en el término cuadrático de la función objetivo. Dado que el Hessiano del Lagrangiano puede ser indefinido en un punto de Karush-Kuhn-Tucker (KKT), lo reemplazamos por una aproximación semidefinida positiva \mathbf{H}_k para tener un problema QP convexo.

Hay varios métodos disponibles para aproximar el Hessiano del Lagrangiano. El método utilizado en este trabajo se basa en la descomposición de valores propios de la matriz simétrica $\left. \frac{\partial^2 L}{\partial \mathbf{u}^2} \right|_k$:

$$\left. \frac{\partial^2 L}{\partial \mathbf{u}^2} \right|_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^\top,$$

con una matriz diagonal \mathbf{D}_k y una matriz ortogonal normalizada \mathbf{U}_k . Sea \mathbf{D}_k^+ la matriz obtenida de \mathbf{D}_k reemplazando los valores negativos de \mathbf{D}_k con ceros. La aproximación del Hessiano \mathbf{H}_k se obtiene como:

$$\mathbf{H}_k = \mathbf{U}_k \mathbf{D}_k^+ \mathbf{U}_k^\top.$$

En base a la aproximación QP convexa (3.12), usamos las siguientes aproximaciones para las funciones $J(\mathbf{u}, \mathbf{d})$ y $y_i(\mathbf{u}, \mathbf{d})$:

$$\tilde{J}(\mathbf{u}_k, \mathbf{d}_k) = \left. \frac{\partial J}{\partial \mathbf{u}} \right|_k \delta \mathbf{u}_k + \frac{1}{2} \delta \mathbf{u}_k^\top \mathbf{H}_k \delta \mathbf{u}_k \quad (3.13a)$$

$$\tilde{y}_i(\mathbf{u}_k, \mathbf{d}_k) = y_{i,k} + \left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_k \delta \mathbf{u}_k, \quad i = 1, \dots, n_y \quad (3.13b)$$

Nótese que el término de orden cero en la función de costo en (3.13a) se ha omitido ya que no afecta la búsqueda de la solución óptima. Usando estas aproximaciones de funciones, las siguientes restricciones deben agregarse al Problema (3.4):

$$\mathbf{u}_n = \mathbf{u}^*(\mathbf{d}_n) + \delta \mathbf{u}_n \quad (3.14a)$$

$$\mathbf{u}_k = \mathbf{u}^*(\mathbf{d}_k) + \delta \mathbf{u}_k, \quad k = 1, \dots, N. \quad (3.14b)$$

De esta manera, la aproximación lugd-MIQP del Problema (3.4) se obtiene implementando los siguientes pasos:

- I) Para todos los escenarios de perturbación \mathbf{d}_k , $k = 1, \dots, N$, resuelva el Problema NLP (2.5) para obtener los puntos óptimos $\mathbf{u}^*(\mathbf{d}_k)$, $k = 1, \dots, N$.
- II) Evalúe $y_{i,k}$ y $\left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_k$ para $i = 1, \dots, n_y$, $\left. \frac{\partial J}{\partial \mathbf{u}} \right|_k$, y la aproximación del Hessiano \mathbf{H}_k , en los valores óptimos de las variables de entrada $\mathbf{u}^*(\mathbf{d}_k)$, para todos los escenarios de perturbación \mathbf{d}_k , $k = 1, \dots, N$.
- III) Reemplace $J(\mathbf{u}_k, \mathbf{d}_k)$ y $y_i(\mathbf{u}_k, \mathbf{d}_k)$ en el Problema (3.4) por las aproximaciones (3.13a) y (3.13b), respectivamente. Agregue las restricciones (3.14) al Problema (3.4).

3.4.2 MIQP Local en Perturbaciones

Para evitar la carga computacional de tener que resolver el Problema NLP (2.5) para todos los escenarios de perturbaciones, y tener que evaluar el gradiente y Hessiano requerido en todos los puntos óptimos obtenidos, proponemos simplificar aún más la aproximación MIQP del Problema (3.4) considerando la siguiente aproximación del Problema NLP (2.5):

$$\min_{\delta \mathbf{u}} \quad \frac{\partial J}{\partial \mathbf{u}} \Big|_n \delta \mathbf{u} + \delta \mathbf{d}^\top \frac{\partial^2 J}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n \delta \mathbf{u} + \frac{1}{2} \delta \mathbf{u}^\top \mathbf{H}_n \delta \mathbf{u} \quad (3.15a)$$

$$\text{s.t.} \quad y_i^L \leq y_{i,n} + \frac{\partial y_i}{\partial \mathbf{u}} \Big|_n \delta \mathbf{u} + \frac{\partial y_i}{\partial \mathbf{d}} \Big|_n \delta \mathbf{d} + \delta \mathbf{d}^\top \frac{\partial^2 y_i}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n \delta \mathbf{u} + \frac{1}{2} \delta \mathbf{d}^\top \frac{\partial^2 y_i}{\partial \mathbf{d}^2} \Big|_n \delta \mathbf{d} \leq y_i^U, \quad (3.15b)$$

$$i = 1, \dots, n_y,$$

$$\mathbf{u} = \mathbf{u}^*(\mathbf{d}_n) + \delta \mathbf{u}, \quad \mathbf{d} = \mathbf{d}_n + \delta \mathbf{d}, \quad (3.15c)$$

$$\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U, \quad (3.15d)$$

donde las derivadas parciales $\frac{\partial J}{\partial \mathbf{u}} \Big|_n$, $\frac{\partial^2 J}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n$, y $\frac{\partial y_i}{\partial \mathbf{u}} \Big|_n$, $\frac{\partial y_i}{\partial \mathbf{d}} \Big|_n$, $\frac{\partial^2 y_i}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n$, $\frac{\partial^2 y_i}{\partial \mathbf{d}^2} \Big|_n$, para $i = 1, \dots, n_y$, son todas evaluadas en el óptimo nominal del Problema (2.5), $\mathbf{u}^*(\mathbf{d}_n)$, y \mathbf{H}_n es una aproximación positiva definida del Hessiano de la función Lagrangiana $\frac{\partial^2 L}{\partial \mathbf{u}^2} \Big|_n$, evaluada en $\mathbf{u}^*(\mathbf{d}_n)$. Note que el Problema (3.15) es cuadrático y convexo en función de las entradas \mathbf{u} (es decir, para $\delta \mathbf{d}$ fijo), y representa una aproximación local del Problema (2.5) en término de las variables de entrada y las perturbaciones.

Basados en el Problema (3.15), $\tilde{J}(\mathbf{u}_k, \mathbf{d}_k)$ y $\tilde{y}_i(\mathbf{u}_k, \mathbf{d}_k)$ están dadas por:

$$\tilde{J}(\mathbf{u}_k, \mathbf{d}_k) = \frac{\partial J}{\partial \mathbf{u}} \Big|_n \delta \mathbf{u}_k + \delta \mathbf{d}_k^\top \frac{\partial^2 J}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n \delta \mathbf{u}_k + \frac{1}{2} \delta \mathbf{u}_k^\top \mathbf{H}_n \delta \mathbf{u}_k \quad (3.16a)$$

$$\tilde{y}_i(\mathbf{u}_k, \mathbf{d}_k) = y_{i,n} + \frac{\partial y_i}{\partial \mathbf{u}} \Big|_n \delta \mathbf{u}_k + \frac{\partial y_i}{\partial \mathbf{d}} \Big|_n \delta \mathbf{d}_k + \delta \mathbf{d}_k^\top \frac{\partial^2 y_i}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n \delta \mathbf{u}_k + \frac{1}{2} \delta \mathbf{d}_k^\top \frac{\partial^2 y_i}{\partial \mathbf{d}^2} \Big|_n \delta \mathbf{d}_k, \quad (3.16b)$$

$$i = 1, \dots, n_y$$

La aproximación luld-MIQP del Problema (3.4) se obtiene implementando los siguientes pasos:

- I) Resuelva el Problema NLP (2.5) para la perturbación nominal \mathbf{d}_n para obtener el óptimo nominal $\mathbf{u}^*(\mathbf{d}_n)$.
- II) Evalúe las derivadas parciales $\frac{\partial J}{\partial \mathbf{u}} \Big|_n$, $\frac{\partial^2 J}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n$, $\frac{\partial y_i}{\partial \mathbf{u}} \Big|_n$, $\frac{\partial y_i}{\partial \mathbf{d}} \Big|_n$, $\frac{\partial^2 y_i}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n$, $\frac{\partial^2 y_i}{\partial \mathbf{d}^2} \Big|_n$, para $i = 1, \dots, n_y$, y la aproximación del Hessiano \mathbf{H}_n , en $\mathbf{u}^*(\mathbf{d}_n)$.
- III) Reemplace $J(\mathbf{u}_k, \mathbf{d}_k)$ y $y_i(\mathbf{u}_k, \mathbf{d}_k)$ en el Problema (3.4) por las aproximaciones (3.16a) y (3.16b), respectivamente. Agregue las restricciones (3.14) al Problema (3.4).

Notar que las restricciones $\tilde{y}_i(\mathbf{u}_k, \mathbf{d}_k)$ introducidas en 3.16b son lineales dado que el vector $\delta \mathbf{d}_k$ ingresa como parámetro en el problema de optimización. Es decir, en el término $\delta \mathbf{d}_k^\top \frac{\partial^2 y_i}{\partial \mathbf{d} \partial \mathbf{u}} \Big|_n \delta \mathbf{u}_k$ y en $\frac{1}{2} \delta \mathbf{d}_k^\top \frac{\partial^2 y_i}{\partial \mathbf{d}^2} \Big|_n \delta \mathbf{d}_k$, no se presenta una multiplicación entre variables. La corrección introducida en $\tilde{y}_i(\mathbf{u}_k, \mathbf{d}_k)$ por estos términos resultará en un aproximación de mayor calidad.

3.4.3 Calidad de la Aproximación y Aspectos a Tener en Cuenta en la Implementación

Las aproximaciones MIQP presentadas en las subsecciones 3.4.1 y 3.4.2 pueden no conducir a la estructura de control óptima de la formulación MINLP original (3.4). Sin embargo, en los casos en

que el Problema (3.4) se vuelve computacionalmente intratable, las aproximaciones MIQP son mucho más fáciles de resolver y permiten obtener una solución (posiblemente subóptima). La calidad de la aproximación dependerá de la aplicación particular en cuestión, pero en general se puede esperar que las aproximaciones sean localmente buenas, para pequeñas variaciones de las perturbaciones.

En los casos en que el conjunto de restricciones activas cambia con los valores de las perturbaciones, y las restricciones que se vuelven activas o inactivas no sean lineales, la aproximación MIQP local en las perturbaciones (luld-MIQP) puede ser inadecuada ya que, en este enfoque local, el Hessiano de la función Lagrangiana calculado en el valor nominal óptimo se utiliza para todos los escenarios de perturbación. Si para cualquier escenario de perturbación dado el conjunto de restricciones activas cambia con respecto al conjunto activo en el óptimo nominal, el Hessiano del Lagrangiano puede cambiar significativamente con respecto al Hessiano nominal (este no es el caso si las restricciones que se vuelven activas o inactivas son los límites de las variables de entrada o restricciones lineales, ya que su Hessiano es igual a cero). Por lo tanto, en presencia de cambios en el conjunto de restricciones activas, se recomienda utilizar la aproximación MIQP global en las perturbaciones (lugd-MIQP).

3.5 Simulaciones y Resultados

3.5.1 Sistema Lineal

Consideraremos un sistema lineal similar al propuesto en Marchetti et al. [7]. El comportamiento en estado estacionario de la planta es descrito por el siguiente sistema:

$$\mathbf{y}(\mathbf{u}, \mathbf{d}) = \mathbf{y}_s - \mathbf{A}_s \mathbf{u} - \mathbf{D}_s \mathbf{d},$$

con,

$$\mathbf{A}_s = \begin{bmatrix} -1.00 & 1.00 \\ -1.00 & -1.00 \\ -1.60 & 0.24 \end{bmatrix}, \quad \mathbf{D}_s = \begin{bmatrix} -0.2702 & 0 & 0 \\ 0 & 0.3003 & 0 \\ 0 & 0 & -0.3809 \end{bmatrix}, \quad \mathbf{y}_s = \begin{bmatrix} -0.35 \\ -1.00 \\ -1.10 \end{bmatrix},$$

que consiste en dos variables de entrada $\mathbf{u} = [u_1, u_2]^\top$, tres variables de salida medibles $\mathbf{y} = [y_1, y_2, y_3]^\top$, y tres perturbaciones no medibles $\mathbf{d} = [d_1, d_2, d_3]^\top$. El objetivo es maximizar la función lineal $J(\mathbf{u}) = 6.85u_1 + 2.95u_2$ sujeto a restricciones en las variables de salida. Se considerarán dos problemas de optimización. El Problema A incluye restricciones en las salidas y_1 y y_2 , mientras que el Problema B incluye restricciones en las tres variables de salida.

Problema A

$$\begin{aligned} \max_{\mathbf{u}} \quad & 6.85u_1 + 2.95u_2 \\ \text{s.t.} \quad & y_1(\mathbf{u}, d_1) \leq 0, \\ & y_2(\mathbf{u}, d_2) \leq 0, \\ & 0 \leq u_1, u_2 \leq 1. \end{aligned}$$

Problema B

$$\begin{aligned} \max_{\mathbf{u}} \quad & 6.85u_1 + 2.95u_2 \\ \text{s.t.} \quad & \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{0}, \\ & 0 \leq u_1, u_2 \leq 1. \end{aligned}$$

En este ejemplo, cada perturbación afecta independientemente a cada variable de salida. Los valores nominales de las perturbaciones son iguales a cero, es decir, $\mathbf{d}_n = \mathbf{0}$. Se supone que las perturbaciones se distribuyen uniformemente en los siguientes rangos: $-1 \leq d_1 \leq 1$, $-1 \leq d_2 \leq 1$, $-0.1 \leq d_3 \leq 0.1$. Los mapas de contorno para los problemas A y B se muestran en la Figura 3.1. El área coloreada corresponde a la región factible para los valores nominales de las perturbaciones para ambos problemas.

El Punto \mathbf{a} es el óptimo nominal para ambos problemas, es decir, $\mathbf{a} = \mathbf{u}^*(\mathbf{d}_n)$. Note que las restricciones y_1 e y_2 están activas en el óptimo nominal. El Punto \mathbf{b} es la solución del problema de los márgenes de seguridad de estado estacionario a lazo abierto. Este punto corresponde al óptimo del peor caso de los escenarios de perturbación $\mathbf{d} = [1, -1]^T$.

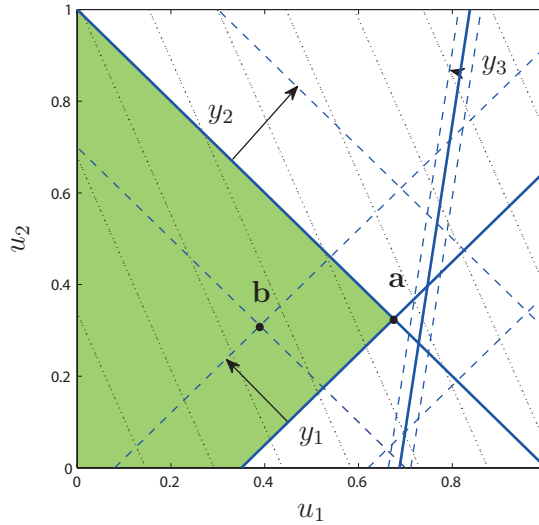


Figura 3.1: Mapas de contorno para los problemas A y B. Área coloreada: región factible para las perturbaciones nominales; Líneas sólidas gruesas: límites de las restricciones de salida para las perturbaciones nominales (y_3 solamente para el Problema B); Líneas discontinuas: límites de las restricciones de salida para los valores extremos de las perturbaciones (y_3 solamente para el Problema B); Líneas punteadas: contornos de la función objetivo; Punto \mathbf{a} : óptimo nominal; Punto \mathbf{b} : solución al problema de los márgenes de seguridad en estado estacionario a lazo abierto. Las flechas indican los movimientos de las restricciones producidos al cambiar los valores de las perturbaciones d_i , $i = 1, 2$, de 0 a 1 y d_3 de 0 a 0.1.

Para este sistema lineal, el Problema (3.4) y su correspondiente problema con minimización (maximización en este caso) del costo nominal, se convierten en problemas MILP, y como tales, los resolvemos en GAMS usando el solver CPLEX. Cada solución se obtiene en menos de un segundo de tiempo de CPU. Los N escenarios de perturbación se obtienen mediante una discretización tipo grilla del conjunto de perturbaciones con $d_{1,k} \in \{-1 : 0.5 : 1\}$, $d_{2,k} \in \{-1 : 0.5 : 1\}$, y $d_{3,k} \in \{-0.1 : 0.05 : 0.1\}$. De hecho, para un sistema lineal sería suficiente evaluar las restricciones en los valores extremos de las perturbaciones para garantizar la factibilidad. Para este ejemplo, como los rangos de las perturbaciones están centrados en los valores nominales y el problema es lineal, el problema de los márgenes de seguridad con minimización del costo promedio nos da los mismos resultados que el problema de los márgenes de seguridad con minimización del costo nominal.

3.5.1.1 Selección de Estructura de Control para el Problema A

En el caso del Problema A, las restricciones sobre y_1 y y_2 están activas para los valores nominales de las perturbaciones y para todas las posibles realizaciones de las perturbaciones d_1 y d_2 . Por lo tanto, en este problema, el conjunto de restricciones activas no cambia con los valores de las perturbaciones. La estructura de control óptima obtenida tanto para el problema de los márgenes de seguridad en estado estacionario con minimización del costo nominal como para el problema con minimización del costo promedio, es controlar ambas restricciones activas (las salidas y_1 y y_2) sin márgenes de seguridad, es

decir, con los valores de consigna $y_1^{\text{sp}} = y_2^{\text{sp}} = 0$.

3.5.1.2 Selección de Estructura de Control para el Problema B

El Problema B incluye restricciones en las tres variables de salida, y a partir de la observación de la Figura 3.1 queda claro que hay muchos escenarios de perturbación para los cuales la restricción en y_3 se vuelve activa mientras que la restricción en y_1 se vuelve inactiva. Por lo tanto, en este problema, el conjunto de restricciones activas cambia dependiendo de los valores de las perturbaciones. La estructura de control óptima obtenida tanto para el problema de los márgenes de seguridad en estado estacionario con minimización del costo nominal como para el problema con costo promedio, es controlar la salida y_2 sin margen de seguridad ($y_2^{\text{sp}} = 0$) y la salida y_1 con un margen de seguridad significativo de 0,427 ($y_1^{\text{sp}} = -0,427$). Esta estructura la denotamos como CS1. El punto de operación nominal y la región de operación para esta estructura de control se muestran en la Figura 3.2. Nótese que el margen de seguridad en y_1 se selecciona como el margen más pequeño tal que la restricción sobre y_3 nunca se viole al controlar y_1 y y_2 .

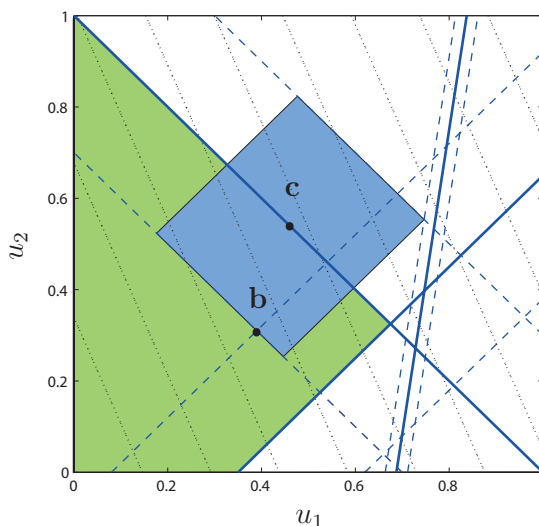


Figura 3.2: Mapa de contornos para Problema B. Área azul: región operativa para la estructura de control óptima CS1; Punto **b**: solución al problema de márgenes de seguridad en estado estacionario a lazo abierto; Punto **c**: punto de operación nominal para la estructura de control CS1.

En este problema de baja dimensión, solo hay diez posibles estructuras de control clásicas, que se ordenan en la Tabla 3.1 de acuerdo con sus valores óptimos. La estructura de control CS5 corresponde a la solución a lazo abierto. Nótese que las estructuras de control CS6 a CS9 tienen un peor rendimiento que la solución a lazo abierto, mientras que CS10 no es factible. Los valores de consigna óptimos obtenidos mediante la optimización para cada estructura de control se detallan en la Tabla 3.2.

Dado que muchos métodos basados en el cálculo de márgenes de seguridad dinámicos fijan las variables de entrada que no se seleccionan como variables manipuladas en sus valores nominales óptimos [18, 19, 21, 25], es interesante ver el efecto que esto tiene desde la perspectiva del estado estacionario. Consideremos, por ejemplo, la estructura de control CS3 de la Tabla 3.1, para la cual la línea de operación y el punto nominal s se muestran en la Figura 3.3. Si en lugar de fijar u_2 en el valor óptimo de 0.61, lo fijamos en el valor nominal de 0.325, se requiere un considerable margen de seguridad de valor igual a 0.5705 para la variable controlada y_2 con el fin de mantener la factibilidad para todos los escenarios de perturbación. Por lo tanto, el punto nominal s se mueve al punto s' en

Tabla 3.1: Estructuras de control para el sistema lineal

CS	\mathbf{z}^O			\mathbf{z}^I		J_n	J_{av}
	y_1	y_2	y_3	u_1	u_2	($N = 125$)	($N = 125$)
CS1	1	1	0	0	0	4.750	4.750
CS2	0	1	1	0	0	4.542	4.542
CS3	0	1	0	0	1	4.470	4.470
CS4	0	1	0	1	0	4.470	4.470
CS5	0	0	0	1	1	3.584	3.584
CS6	0	0	1	0	1	3.421	3.421
CS7	1	0	0	0	1	2.787	2.787
CS8	0	0	1	1	0	2.497	2.497
CS9	1	0	0	1	0	1.888	1.888
CS10	1	0	1	0	0	infactible	infactible

Tabla 3.2: Valores de consigna obtenidos para el sistema lineal

CS	\mathbf{r}^{SP}				
	y_1	y_2	y_3	u_1	u_2
CS1	-0.427	0.000	-	-	-
CS2	-	0.000	-0.589	-	-
CS3	-	0.000	-	-	0.610
CS4	-	0.000	-	0.397	-
CS5	-	-	-	0.389	0.310
CS6	-	-	-0.589	-	0.310
CS7	0.000	-	-	-	0.039
CS8	-	-	-0.805	0.231	-
CS9	-0.461	-	-	0.159	-

la Figura 3.3, que es muy subóptimo. Si bien esta estructura de control sería factible en la práctica, tenga en cuenta que la entrada u_1 se saturaría para muchos valores de perturbación, lo que significa que perdemos el control sobre y_2 . Dado que la posibilidad de saturación de entradas no se contempla en el Problema (3.4), esta estructura de control no será factible para el Problema (3.4).

3.5.1.3 Estructura de Control de Combinaciones Lineales para el Problema B

La estructura de control clásica óptima consiste en controlar y_1 e y_2 con los valores consigna -0.427 y 0 , respectivamente, lo que da una ganancia promedio $J_{av} = 4.750$. Ahora mostraremos cómo este beneficio se puede aumentar aún más seleccionando una estructura de control de combinaciones lineales. En particular, consideremos la estructura de control en la que se controla la salida y_2 y se fija una combinación lineal de las entradas, es decir, $r_1 = y_2$ y $r_2 = a_1 u_1 + a_2 u_2$. A continuación, fijamos arbitrariamente $a_1 = 1$ y encontramos los valores óptimos de a_2 y los valores de consigna r_1^{SP} y r_2^{SP} resolviendo el Problema (3.10). Los valores de consigna óptimos y la matriz de combinaciones lineales correspondiente \mathcal{Q} se muestran en la Tabla 3.3. La ganancia promedio aumenta a $J_{av} = 5.056$. La Figura 3.4 muestra la línea operativa y el punto operativo nominal \mathbf{d} para esta estructura de control, que requiere un solo lazo de control, y aumenta la ganancia con respecto al punto \mathbf{c} , obtenido para la estructura de control clásica óptima.

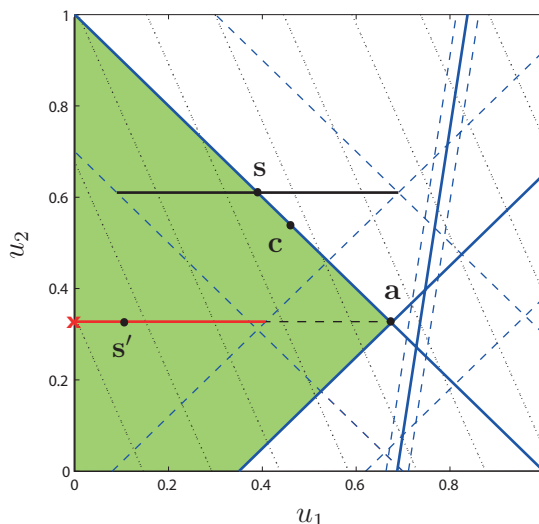


Figura 3.3: Mapas de contorno para Problema B. Segmento negro grueso: Línea de operación para la estructura de control CS3; Segmento rojo grueso: línea de operación para CS3 con u_2 fijo en su valor nominal óptimo; Punto **a**: óptimo nominal; Punto **c**: punto de funcionamiento nominal para CS1; Punto **s**: punto de funcionamiento nominal para CS3; Punto **s'**: punto de funcionamiento nominal para CS3 con $u_2 = u_2^*(\mathbf{d}_n)$.

Tabla 3.3: Estructura de control de combinaciones lineales obtenida para el Problema B.

\mathbf{r}	Q^T					\mathbf{r}^{sp}
	y_1	y_2	y_3	u_1	u_2	
r_1	0	1	0	0	0	0
r_2	0	0	0	1	-1	0.079

3.5.1.4 Conclusiones del Caso del Estudio 3.5.1

- Este caso de estudio muestra una situación para la cual el problema de los márgenes de seguridad en estado estacionario con minimización del costo nominal es equivalente al mismo problema usando la minimización del costo promedio. De hecho, la equivalencia se da ya que estamos considerando un sistema lineal para el cual los valores nominales de las perturbaciones son iguales a los valores esperados (o promedio) de las perturbaciones.
- En el Problema A, para el cual el conjunto de restricciones activas no cambia, los problemas de los márgenes de seguridad en estado estacionario determinan que es óptimo controlar ambas restricciones activas sin márgenes de seguridad. Esto corresponde a la conocida estrategia de control de restricciones [40].
- En el Problema B, para el cual el conjunto de restricciones activas cambia con las perturbaciones, el problema de los márgenes de seguridad de estado estacionario determina qué variables controlar y qué valores de consigna usar. Nótese que incluso para un sistema lineal simple, esta elección no es trivial.
- Se ha ilustrado la importancia de calcular valores óptimos para las entradas que no se seleccionan como variables manipuladas.
- Finalmente, se mostró que es posible incrementar aún más el rendimiento económico de la operación seleccionando una estructura de control de combinaciones lineales de variables. Tenga

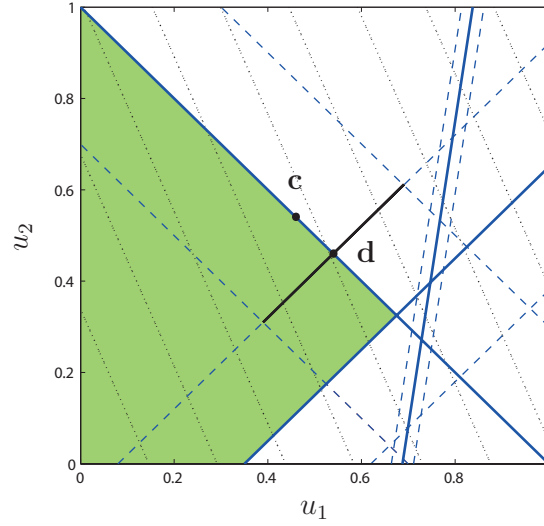
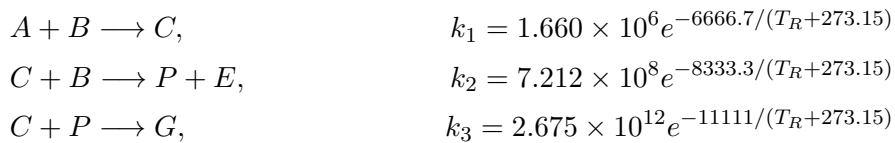


Figura 3.4: Mapas de contorno para el Problema B. Segmento negro grueso: Línea operativa para la estructura de control de la Tabla 3.3; Punto **c**: punto operativo nominal para la estructura de control CS1; Punto **d**: punto operativo nominal para la estructura de control en la Tabla 3.3.

en cuenta que, en la literatura de SOC, se han utilizado estructuras de control de combinaciones lineales para llevar el gradiente reducido de la función de costo a cero [30, 65, 66]. En este caso de estudio, mostramos que las estructuras de control de combinaciones lineales también pueden ser útiles en los casos en que la solución está completamente determinada por restricciones activas y el conjunto de restricciones activas cambia con los valores de las perturbaciones.

3.5.2 Reactor Tanque Agitado Continuo

Nuestro siguiente caso de estudio considera el reactor de Williams-Otto [67], que consiste en un reactor tanque agitado continuo (CSTR) ideal donde tienen lugar las siguientes reacciones:



Los reactivos A y B se alimentan con caudales másicos F_A y F_B , respectivamente. Los productos deseados son P y E , C es un producto intermedio y G es un subproducto no deseado. El reactor funciona isotérmicamente a temperatura T_R , con una retención de masa constante de 2105 kg. La dependencia de las constantes de velocidad de reacción k_1 , k_2 y k_3 con la temperatura absoluta viene dada por las ecuaciones de Arrhenius. En el apéndice podrá encontrar las ecuaciones del modelo.

Para este problema, las variables de entrada son F_B y T_R , es decir, $\mathbf{u} = [F_B, T_R]^T$, las variables de salida medidas son las fracciones de masa de todos los componentes a la salida del reactor, $\mathbf{y} = [X_A, X_B, X_C, X_P, X_G, X_E]^T$, y la única perturbación es $d = F_A$. El objetivo es maximizar la ganancia en estado estacionario sujeto a límites en las entradas y un límite superior de 0.25 en la

fracción másica de B a la salida del reactor. El problema de optimización es el siguiente:

$$\begin{aligned} \underset{F_B, T_R}{\text{máx}} \quad & 1200X_P F + 80X_E F - 76F_A - 114F_B \\ \text{s.t.} \quad & \text{ecuaciones del modelo en estado estacionario,} \\ & X_B \leq 0.25, \\ & F_B \in [2, 4], \quad T_R \in [70, 100]. \end{aligned} \tag{3.17}$$

donde $F = F_A + F_B$. El valor nominal de la perturbación es $F_A = 1.4$ kg/s, y el rango de variación es $F_A \in [1.2, 1.83]$.

En este problema, la restricción sobre X_B está activa en la solución óptima para todos los valores de la perturbación. El efecto de la perturbación sobre la solución óptima se muestra en la Figura 3.5. En azul sólido se representa la línea de contorno correspondiente a $X_B = 0.25$ para la perturbación nominal, mientras que las líneas discontinuas azules inferior y superior muestran cómo se mueve esta línea de contorno para los valores extremos de la perturbación. El punto **a** es el óptimo nominal, es decir, $\mathbf{a} = \mathbf{u}^*(1.4)$, y la línea roja representa la línea de operación óptima $\mathbf{u}^*(F_A)$. Además, la figura muestra las líneas de operación de la planta controlada usando dos estructuras de control particulares (línea vertical negra y línea horizontal negra).

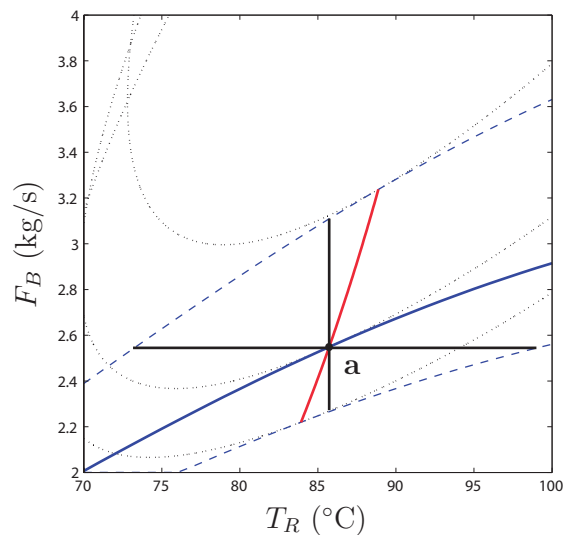


Figura 3.5: Mapa de contornos para el Problema (3.17). Línea azul gruesa: límite de la restricción sobre X_B para la perturbación nominal ($F_A = 1.4$); Líneas discontinuas azules: límites de la restricción X_B para los valores extremos de la perturbación F_A ; Punto **a**: óptimo nominal; Línea roja gruesa: línea de soluciones óptimas; Líneas punteadas: contornos de la ganancia económica; Línea vertical gruesa: línea de operación para la estructura de control CS4; Línea horizontal gruesa: línea de operación para la estructura de control CS7.

3.5.2.1 Estructura de Control Clásica para el CSTR

Para este problema, hay 28 estructuras de control posibles. Las siete mejores estructuras de control se enumeran en la Tabla 3.4 según su valor óptimo. Para cada estructura de control, se reporta la ganancia nominal y la ganancia promedio correspondiente a los problemas NLP de los márgenes de seguridad de estado estacionario, así como las ganancias obtenidas usando la aproximación QP (local

en las perturbaciones) de ambos problemas, descrita en la Subsección 3.4.2. En todos los casos, usamos una discretización uniforme del rango de la perturbación $F_A \in [1.2, 1.83]$, con un paso de 0.01, lo que da $N = 64$ escenarios de perturbación. En la Tabla 3.5 se dan los valores de consigna óptimos obtenidos para el problema NLP de los márgenes de seguridad con maximización de la ganancia promedio para cada estructura de control.

Tabla 3.4: Mejores estructuras de control para el problema CSTR. Clasificación basada en J_{av}^{NLP} .

CS	\mathbf{z}^O						\mathbf{z}^I		J_n^{QP}	J_{av}^{QP}	J_n^{NLP}	J_{av}^{NLP}
	X_A	X_B	X_C	X_P	X_G	X_E	F_B	T_R	($N = 64$)	($N = 64$)	($N = 64$)	($N = 64$)
CS1	0	1	0	0	1	0	0	0	215.79	228.92	215.79	225.16
CS2	0	1	1	0	0	0	0	0	215.79	228.90	215.79	225.15
CS3	1	1	0	0	0	0	0	0	215.79	228.90	215.79	225.04
CS4	0	1	0	0	0	0	0	1	215.79	228.54	215.79	224.88
CS5	0	1	0	0	0	1	0	0	215.79	228.80	215.79	224.78
CS6	0	0	1	0	1	0	0	0	215.15	227.56	212.15	222.50
CS7	0	1	0	0	0	0	1	0	215.79	218.29	215.79	212.26

Tabla 3.5: Problema CSTR: valores de consigna calculados para el problema de los márgenes de seguridad con ganancia promedio.

CS	\mathbf{r}^{sp}							
	X_A	X_B	X_C	X_P	X_G	X_E	F_B	T_R
CS1	-	0.250	-	-	0.170	-	-	-
CS2	-	0.250	0.021	-	-	-	-	-
CS3	0.126	0.250	-	-	-	-	-	-
CS4	-	0.250	-	-	-	-	-	86.723
CS5	-	0.250	-	-	-	0.326	-	-
CS6	-	-	0.022	-	0.170	-	-	-
CS7	-	0.250	-	-	-	-	2.561	-

Las últimas dos columnas de la Tabla 3.4 se construyeron fijando la estructura de control en el Problema (3.4), y resolviendo (3.4) como un problema NLP. En el caso del problema de márgenes de seguridad con funcional nominal, se puede ver en la Tabla 3.4 que todas las estructuras de control para las cuales se selecciona la restricción activa (la fracción másica X_B) como variable controlada son soluciones óptimas que comparten la misma ganancia nominal óptima. La ganancia J_n^{NLP} asociada a estas seis estructuras de control corresponde a la ganancia en el punto óptimo nominal \mathbf{a} en la Figura 3.5. La Figura 3.5 también muestra las líneas de operación de la planta controlada usando la estructura de control CS4 (línea vertical negra) y CS7 (línea horizontal negra). Note que la línea de operación de CS4 está más cerca de la línea de solución óptima (línea roja) que la de CS7, y está claro que para distintos valores de perturbación, la pérdida asociada con CS7 será significativamente mayor que la pérdida de CS4. Sin embargo, el problema de los márgenes de seguridad en estado estacionario con el funcional evaluado en el valor de perturbación nominal no puede distinguir entre ambas estructuras de control. En cambio, el problema de los márgenes de seguridad de estado estacionario con la ganancia promediada permite clasificar correctamente las estructuras de control de acuerdo a qué tan bien su línea de operación se aproxima a la línea de soluciones óptimas. Esto solo requiere considerar maximizar la ganancia promedio (o minimizar el costo promedio) en lugar de la ganancia nominal (o costo nominal). La (única) solución óptima al problema de los márgenes de seguridad en estado estacionario considerando la ganancia económica promedio (3.4) es la estructura de control CS1, que consiste en controlar las fracciones másicas X_B y X_G . La línea de operación para CS1 se observa en la Figura 3.6.

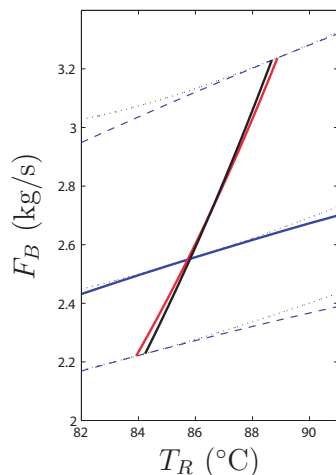


Figura 3.6: Problema CSTR. Línea negra gruesa: Línea operativa para la estructura de control clásica óptima CS1. Línea roja gruesa: mapa de soluciones óptimas.

3.5.2.2 Estrategias de Solución Alternativas

Como se menciona en la Sección 3.2.3, la estrategia de solución seleccionada depende de la complejidad de las ecuaciones del modelo de proceso. En el caso del problema CSTR, el solver global BARON llegó a la estructura de control óptima. Sin embargo, cuando aumenta el número de escenarios de perturbación (en este caso hasta $N = 64$) el solver BARON experimentó dificultades para reducir la brecha (gap en inglés) establecida como criterio de corte en un tiempo aceptable. Por ello, el solver fue interrumpido cuando alcanzó un tiempo máximo, sin demostrar la convergencia al óptimo global. Por otro lado, los solvers DICOPT y BONMIN mostraron un buen desempeño, llegando a la estructura de control óptima en un tiempo aceptable. Finalmente, la aproximación luld-MIQP convergió a la estructura de control óptima, CS1, en solo 0.4 segundos.

Tabla 3.6: Problema CSTR. Resultados obtenidos utilizando diferentes solvers (Ordenador utilizado: Intel (R) Core (TM), CPU i5-4440, 3,10 GHz, 8 GB RAM). (*) Interrumpido.

Problema	Solver/subsolvers	CS (r)	J_{av}^{NLP}	Tiempo(s)
Búsqueda Exhaustiva	CONOPT	X_B, X_G	225.158	343.3
MINLP	DICOPT/MINOS,CPLEX	X_B, X_C	225.154	5.2
MINLP	DICOPT/CONOPT,CPLEX	X_B, X_G	225.158	4.1
MINLP	BONMIN/IPOPT,B-B	X_B, X_G	225.158	56.8
MINLP	BARON/BARON,CPLEX	X_B, X_G	225.158	1400(*)
luld-MIQP	CPLEX/CONOPT,CPLEX	X_B, X_G	225.158	0.4

3.5.2.3 Estructuras de Control de Combinaciones Lineales para el CSTR

Marchetti y Zumoffen [36] mostraron que el método de espacio nulo (que es un método SOC local) se puede aplicar a este problema CSTR controlando la salida restringida activa X_B y fijando una combinación lineal de las entradas. La estructura de control correspondiente (tomada de Marchetti y Zumoffen [36]) está dada en la Tabla 3.7. Los valores de consigna son iguales a los valores de las variables controladas evaluadas en el óptimo nominal $\mathbf{u}^*(1.4)$. La línea de operación correspondiente se muestra en el gráfico de la izquierda de la Figura 3.7a. Se puede ver que la línea de operación (línea negra) es tangente a la curva de soluciones óptimas (línea roja) en el punto óptimo nominal \mathbf{a} .

A continuación, aplicamos el enfoque de SOC global propuesto en el Problema NLP (3.10).

Nuevamente, consideramos la estructura de control en la que se controla la restricción activa X_B y se fija una combinación lineal de las entradas, es decir, $r_1 = X_B$ y $r_2 = aF_B + T_R$. Al resolver el Problema (3.10) encontramos los valores óptimos del coeficiente a y los valores óptimos de las consignas r_1^{sp} y r_2^{sp} . La solución se da en la Tabla 3.7, y la línea de operación correspondiente se muestra en el gráfico de la derecha de la Figura 3.7a. Al comparar ambos gráficos, se puede ver cómo la línea operativa para el enfoque global de SOC se aproxima globalmente a la línea de soluciones óptimas. Sin embargo, la mejora en la ganancia económica para el enfoque global es insignificante, ya que para este ejemplo en particular, el método del espacio nulo funciona muy bien.

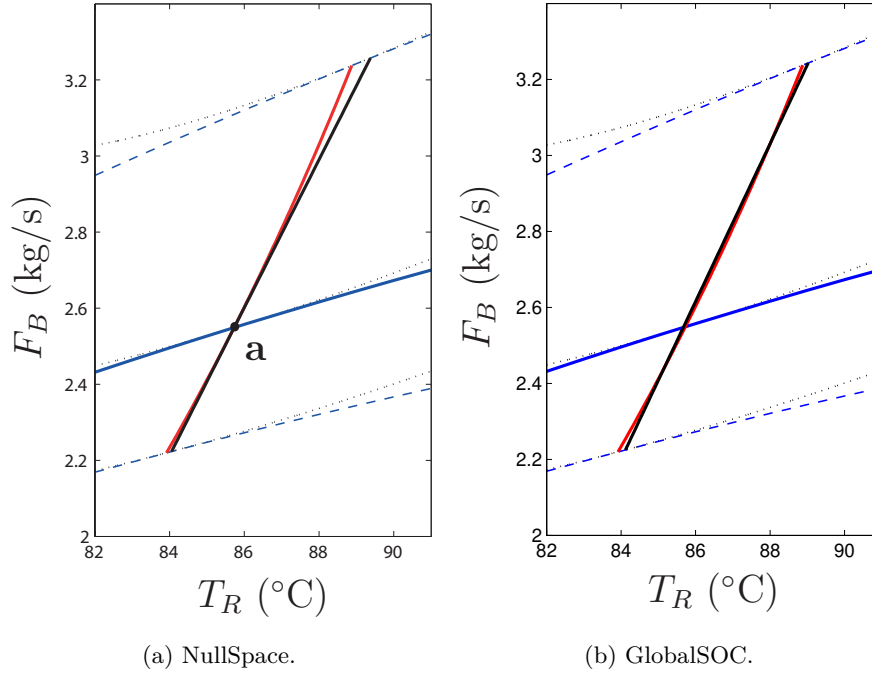


Figura 3.7: Problema CSTR. Línea roja gruesa: curva de soluciones óptimas. Punto **a**: óptimo nominal. Líneas negras gruesas: líneas operativas para las estructuras de control de combinaciones lineales. Gráfico de la izquierda: método del espacio nulo. Gráfico de la derecha: SOC global.

Tabla 3.7: Estructura de control de combinaciones lineales para el problema CSTR.

	\mathbf{r}	\mathcal{Q}^T								\mathbf{r}^{sp}	J_{av} ($N = 64$)
		X_A	X_B	X_C	X_P	X_G	X_E	F_B	T_R		
SOC	r_1	0	1	0	0	0	0	0	0	0.250	225.156
Global	r_2	0	0	0	0	0	0	-0.321	1	0.434	
Espacio nulo	r_1	0	1	0	0	0	0	0	0	0.250	225.155
	r_2	0	0	0	0	0	0	-0.324	0.946	0.407	

3.5.2.4 Conclusiones del Caso de Estudio 3.5.2

- Este caso de estudio muestra claramente las limitaciones de minimizar el costo nominal en los métodos basados en márgenes de seguridad (backoff approaches). Como se ve en la Tabla 3.4 y en la Figura 3.5, el uso del costo nominal no permite distinguir la diferencia de rendimiento entre las estructuras de control CS4 y CS7.
- Se ha ilustrado la idoneidad para resolver el problema de la aproximación MIQP local en las

perturbaciones. La aproximación luld-MIQP propuesta convergió a la estructura de control óptima en un tiempo reducido.

- En todos los casos (MINLP o aproximación MIQP con costo nominal o promedio), el enfoque de los márgenes de seguridad en estado estacionario determina que es óptimo controlar la restricción activa X_B con margen de seguridad nulo (es decir, con un valor de consigna de 0.25).
- Este caso de estudio también ilustra que los problemas de márgenes de seguridad de estado estacionario con minimización de costo promedio son métodos del tipo SOC globales. La comparación con el método del espacio nulo ilustra la diferencia entre un método SOC global y uno local.

3.5.3 Evaporador

Nuestro tercer caso de estudio considera el evaporador de circulación forzada descrito por Newell y Lee [68]. La alimentación líquida mezclada con licor recirculante hierve dentro del evaporador calentado con vapor. La mezcla vapor-líquido generada se divide en el separador. La mayor parte del líquido separado se recircula, mientras que una pequeña fracción se extrae como producto. El evaporador se ilustra en la Figura 3.8 y sus principales variables se enumeran en la Tabla 3.8 junto con los valores nominales.

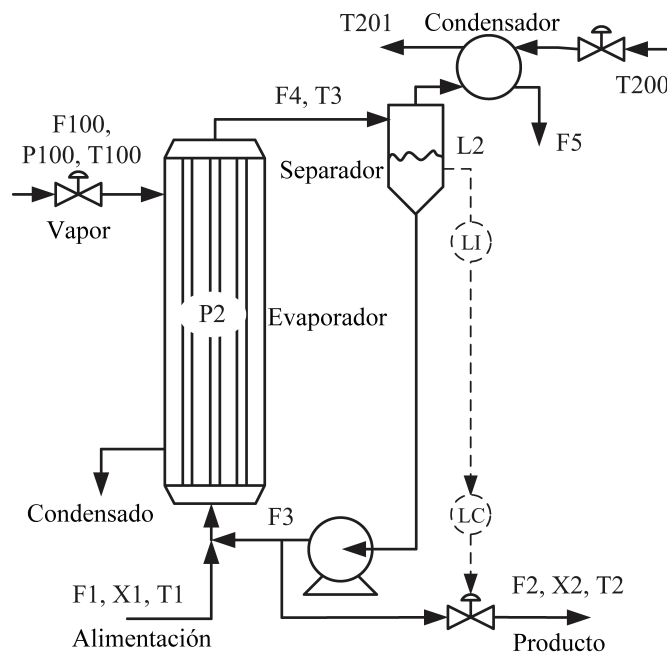


Figura 3.8: Sistema de Evaporador.

Las ecuaciones del modelo se encuentran detalladas en el apéndice de la tesis. El modelo del evaporador tiene tres variables de estado, L_2 , X_2 , P_2 , y ocho grados de libertad: las cinco variables de entrada P_{100} , F_{200} , F_1 , F_2 , F_3 , y las tres perturbaciones, X_1 , T_1 , T_{200} . Para estabilizar el proceso, se incluye un controlador PI, que controla el nivel del separador L_2 manipulando el caudal de producto F_2 . Por lo tanto, vamos a considerar las siguientes variables para el problema de selección de la estructura

Tabla 3.8: Evaporador: Principales variables y sus valores en el óptimo nominal.

Variable	Descripción	Valor	Unidad
F_{200}	Caudal del agua refrigerante	217.74	[kg/min]
P_{100}	Presión del vapor de entrada	400.00	[kPa]
F_3	Caudal de recirculación	24.72	[kg/min]
F_1	Caudal de alimentación	9.469	[kg/min]
X_2	Composición del producto	35.50	[%]
F_4	Caudal de la mezcla de salida	8.135	[kg/min]
F_5	Caudal del condensado	8.135	[kg/min]
T_2	Temperatura del producto	88.4	[°C]
T_3	Temperatura de la mezcla de salida	81.1	[°C]
P_2	Presión operativa	51.41	[kPa]
F_{100}	Caudal del vapor de entrada	9.434	[kg/min]
T_{100}	Temperatura del vapor de entrada	151.5	[°C]
T_{201}	Temperatura de salida del agua de refrigeración	45.55	[°C]
X_1	Composición de la alimentación	5.00	[%]
T_1	Temperatura de la alimentación	40.0	[°C]
T_{200}	Temperatura de la entrada del agua de refrigeración	25.0	[°C]
J	Ganancia económica	582.23	[\$/h]

de control:

$$\begin{aligned}\mathbf{u} &= [F_{200} \quad P_{100} \quad F_3 \quad F_1]^\top \\ \mathbf{y} &= [X_2 \quad F_4 \quad F_5 \quad T_2 \quad T_3 \quad P_2 \quad F_{100} \quad T_{100} \quad T_{201}]^\top \\ \mathbf{d} &= [X_1 \quad T_1 \quad T_{200}]^\top\end{aligned}$$

Los valores nominales de las perturbaciones son: $X_1 = 5$, $T_1 = 40$, $T_{200} = 25$, y los rangos de variación son $X_1 \in [4.75, 5.25]$, $T_1 \in [32, 48]$, y $T_{200} \in [20, 30]$. El objetivo es maximizar el beneficio económico en estado estacionario, que puede expresarse como:

$$J = 4800F_2 - 0.2F_1 - 600F_{100} - 0.6F_{200} - 1.009(F_2 + F_3),$$

donde el primer término es el valor del producto, mientras que los últimos cuatro términos son los costos operativos. El problema de optimización se puede escribir como:

$$\begin{aligned}\max_{\mathbf{u}} \quad & J \\ \text{s.t.} \quad & \text{ecuaciones del modelo en estado estacionario,} \\ & X_2 \geq 35.5 \\ & 40 \leq P_2 \leq 80 \\ & P_{100} \leq 400 \\ & 0 \leq F_{200} \leq 400 \\ & 0 \leq F_1 \leq 20 \\ & 0 \leq F_3 \leq 100\end{aligned}\tag{3.18}$$

En este problema, las restricciones en X_2 y P_{100} están activas para todos los escenarios de perturbaciones, mientras que la restricción en P_2 puede volverse activa en su límite inferior o superior dependiendo de los valores de las perturbaciones.

3.5.3.1 Estructuras de Control Clásicas para el Evaporador

El problema de los márgenes de seguridad en estado estacionario (3.4) con minimización del costo promedio se resolverá usando una discretización de tipo grilla del conjunto de perturbaciones con $d_{1,k} \in \{4.75 : 0.125 : 5.25\}$, $d_{2,k} \in \{32 : 4 : 48\}$, y $d_{3,k} \in \{20 : 2.5 : 30\}$, lo que da un total de $N = 125$ escenarios de perturbación. Se considerarán dos casos. En el caso A, el Problema (3.4) se resuelve sin restricciones adicionales, mientras que en el caso B, las restricciones (3.7) se incluyen con los límites de la matriz RGA $\delta^L = 0.5$ y $\delta^U = 1.5$. La estructura de control óptima y los valores de consigna correspondientes obtenidos en ambos casos se dan en la Tabla 3.9. El inconveniente con la solución óptima obtenida para el Caso A es que las variables controladas T_{100} y P_{100} son linealmente dependientes (ver las ecuaciones del modelo en el apéndice) y, por lo tanto, hay un grado de libertad restante que es utilizado por el problema MINLP para encontrar entradas óptimas para cada escenario de perturbación, lo que aumenta la ganancia promedio. Sin embargo, esto no representará solamente el comportamiento del sistema de control regulatorio en retroalimentación instalado sino además un sistema supervisor que puede medir el valor de las perturbaciones en tiempo real y realizar cambios en las variables manipuladas. Esta estructura de control se evita en el Caso B, dado que el cálculo de la matriz RGA requiere invertir el subproceso controlado.

Como se ve en la Tabla 3.9 para el Caso B, la estructura de control óptima consiste en controlar con nulo margen de seguridad las variables restringidas que están activas para todos los escenarios de perturbación. Es decir, la composición del producto X_2 se controla con un valor de consigna de 35,5 y la entrada P_{100} se fija en 400 kPa. Además, las entradas F_{200} y F_3 se fijan en sus valores óptimos calculados.

Tabla 3.9: Estructura de Control Óptima para el Evaporador

		X_2	F_4	F_5	T_2	T_3	P_2	F_{100}	T_{100}	T_{201}	F_{200}	P_{100}	F_3	F_1
Caso	\mathbf{z}	1	0	0	0	0	0	0	1	1	0	1	0	0
A	\mathbf{z}^{sp}	35.5	0	0	0	0	0	0	151.52	46.95	0	400	0	0
Caso	\mathbf{z}	1	0	0	0	0	0	0	0	0	1	1	1	0
B	\mathbf{z}^{sp}	35.5	0	0	0	0	0	0	0	0	218.24	400	24.75	0

3.5.3.2 Estrategias de Solución Alternativas para el Evaporador

El problema MINLP se resolvió en GAMS utilizando diferentes solvers locales y globales. Los resultados se muestran en la Tabla 3.10. El solver local DICOPT convergió rápidamente a una solución subóptima factible que está próxima a la óptima. Esta estructura de control está definida por las variables de salida controladas X_2, T_{201} , las variables de entrada fijas P_{100}, F_3 y los emparejamientos: $X_2 - F_1$ y $T_{201} - F_{200}$. Por otro lado, el solver local BONMIN, que implementa una estrategia branch and bound, convergió a la estructura de control óptima pero en un tiempo computacional mucho mayor. El solver global BARON nos guió a la solución óptima en alrededor de 1896 segundos, pero el programa se interrumpió después de 50000 segundos ya que no fue posible cerrar la brecha (gap) del criterio de detención (29% restante). También se aplicó la aproximación MIQP global en las perturbaciones descrita en la Subsección 3.4.1, la cual convergió a la misma solución subóptima obtenida para el problema MINLP usando DICOPT, pero en un menor tiempo.

Tabla 3.10: Evaporador. Resultados utilizando diferentes estrategias de solución. (Computadora utilizada: Intel(R) Core(TM), CPU i5-4440, 3.10 GHz, 8 GB de RAM). (*)Interrumpido.

Problema	Solver/subsolvers	CS (r)	J_{av}	Tiempo (s)
Búsqueda Exhaustiva	CONOPT	$X_2, F_{200}, P_{100}, F_3$	584.89	12271.1
MINLP	BONMIN/IPOPT,B-B	$X_2, F_{200}, P_{100}, F_3$	584.89	3519.1
MINLP	BARON/BARON,CPLEX	$X_2, F_{200}, P_{100}, F_3$	584.89	50000(*)
MINLP	DICOPT/MINOS,CPLEX	$X_2, T_{201}, P_{100}, F_3$	581.16	9.8
MINLP	DICOPT/CONOPT,CPLEX	$X_2, T_{201}, P_{100}, F_3$	581.16	8.5
lugd-MIQP	CPLEX/CONOPT,CPLEX	$X_2, T_{201}, P_{100}, F_3$	581.16	5.6

3.5.3.3 Estructura de Control de Combinaciones Lineales para el Evaporador

A continuación, buscamos una estructura de control de combinaciones lineales óptima resolviendo el Problema NLP (3.10) utilizando la misma discretización del conjunto de perturbaciones que en la Subsección 3.5.3.1. Como las restricciones en X_2 y P_{100} están activas para todos los escenarios de perturbación, seleccionamos $r_1 = X_2$ y $r_2 = P_{100}$. Las variables controladas restantes r_3 y r_4 se seleccionarán como combinaciones lineales óptimas de las variables $X_2, F_5, P_2, F_{100}, T_{201}, F_{200}, P_{100}, F_3$ y F_1 . Además, la restricción (3.11) se incluye con la siguiente matriz T:

$$T = \begin{bmatrix} \frac{\partial X_2}{\partial F_{200}}|_n & \frac{\partial X_2}{\partial P_{100}}|_n & \frac{\partial X_2}{\partial F_3}|_n & \frac{\partial X_2}{\partial F_1}|_n \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Como consecuencia de seleccionar $r_1 = X_2$, la primera fila de T es igual a la primera fila de la matriz de ganancia G. De ser posible, se recomienda resolver el problema NLP utilizando diferentes solvers y puntos de partida, ya que se puede caer fácilmente en óptimos locales. Para este ejemplo, la mejor solución conocida se encontró con el solver CONOPT, partiendo del óptimo nominal. La matriz de combinaciones lineales correspondiente Q y los valores de consigna óptimos se dan en la Tabla 3.11. La ganancia promedio para los $N = 125$ escenarios de perturbación es $J_{av} = 609.18$. Recuerde que la ganancia promedio para la mejor estructura de control clásica fue $J_{av} = 584.89$.

Tabla 3.11: Estructura de control óptima de combinaciones lineales para el Evaporador.

r	Q^T													r^{sp}
	X_2	F_4	F_5	T_2	T_3	P_2	F_{100}	T_{100}	T_{201}	F_{200}	P_{100}	F_3	F_1	
r_1	1	0	0	0	0	0	0	0	0	0	0	0	0	35.5
r_2	0	0	0	0	0	0	0	0	0	0	1	0	0	400
r_3	-57.31	0	1131	0	0	14.13	87.67	0	-38.01	-0.816	1.080	13.38	-1240	-4171
r_4	-9.326	0	172.7	0	0	5.266	17.56	0	-10.94	-0.515	0.044	1.589	-209.6	-1028

3.5.3.4 Conclusiones del Caso de Estudio 3.5.3

- Este caso de estudio muestra que la formulación de márgenes de seguridad en estado estacionario (3.4) puede conducir a una estructura de control que no es controlable en estado estacionario. Al seleccionar variables controladas que son linealmente dependientes, el proceso controlado no es invertible y los valores de las variables manipuladas no se determinan de forma única. Estos grados de libertad adicionales son explotados por el problema MINLP para encontrar valores óptimos para las variables manipuladas para cada escenario de perturbación. Sin embargo, esto

no se puede realizar mediante lazos de control clásicos del tipo PID. A pesar de no haber sido diseñadas específicamente para evitar este problema, las restricciones adicionales (3.7) evitan esta estructura de control inadecuada ya que el cálculo de la matriz RGA requiere invertir el subproceso controlado.

- Se ha ilustrado la capacidad de la aproximación MIQP global en las perturbaciones para resolver el problema de las estructuras de control. El lugd-MIQP convergió a la misma solución (cerca a la óptima global) que se encontró para el problema MINLP usando DICOPT.
- El problema de encontrar una estructura de control de combinaciones lineales se ilustró usando una matriz de combinación Q con una estructura predefinida e incluyendo la restricción (3.11) con una elección adecuada de matriz T .

3.6 Conclusiones del Capítulo

Este capítulo estudió métodos de los márgenes de seguridad en estado estacionario para la selección de estructuras de control basándose en criterios económicos y su vínculo con el control auto optimizante (SOC). En la literatura, los métodos de márgenes de seguridad generalmente proceden minimizando el costo nominal, y en menor medida el costo promedio, para un conjunto dado de escenarios de perturbación. Se ha comparado el desempeño de ambas estrategias. Se han estudiado dos formulaciones diferentes del problema de los márgenes de seguridad con minimización del costo promedio: (I) La formulación dada por el Problema MINLP (3.4) que selecciona la estructura de control clásica óptima, y (II) la formulación dada por el Problema NLP (3.10) que selecciona una estructura de control de combinaciones lineales óptima. Ambas formulaciones son estrategias de control auto optimizante globales ya que se basan en minimizar el costo promedio y evitando no violar las restricciones en presencia de perturbaciones.

Con respecto a la formulación (I), se ha demostrado que sin consideraciones adicionales el problema MINLP (3.4) puede conducir a estructuras de control inadecuadas si las variables controladas seleccionadas son linealmente dependientes. Se han agregado restricciones adicionales a esta formulación para garantizar que exista un emparejamiento de entrada-salida que dé como resultado un grado aceptable de interacción cuando se implementan controladores descentralizados. Los ejemplos de aplicación muestran que, cuando hay restricciones que están activas para todos los escenarios de perturbación, el problema MINLP selecciona controlar las variables restringidas correspondientes sin márgenes de seguridad. Esto corresponde al enfoque estándar de control de restricciones.

Por otro lado, la formulación (II) dada por el Problema NLP (3.10) (junto con la restricción (3.11)) representa una formulación SOC global para seleccionar combinaciones lineales óptimas de variables de salida y/o entrada como variables controladas, y los valores de consiga óptimos correspondientes, independientemente de si el conjunto de restricciones activas cambia o no con los valores de las perturbaciones. De hecho, varios métodos de control auto optimizante propuestos recientemente [34, 37, 38] pueden verse como aproximaciones y casos especiales del Problema NLP (3.10). En estos métodos, el recurso a aproximaciones matemáticas locales de la pérdida económica es en parte el resultado de intentar minimizar la pérdida promedio en lugar del costo promedio.

Las aproximaciones MIQP local y global en las perturbaciones se aplicaron a los casos de estudio 3.5.2 y 3.5.3, respectivamente, mostrando en ambos casos su idoneidad como aproximaciones a las formulaciones MINLP originales.

Por último, este capítulo se ha centrado en métodos de márgenes de seguridad de estado estacionario. Los métodos de los márgenes de seguridad dinámicos son un problema diferente. Sin embargo, dado que

los métodos de los márgenes de seguridad dinámicos también deben prestar atención a la factibilidad y el desempeño económico en estado estacionario, creemos que muchas de las cosas que hemos aprendido para los métodos de los márgenes de seguridad en estado estacionario también se pueden aplicar a los métodos de los márgenes de seguridad dinámico. Por ejemplo, varios métodos de márgenes de seguridad dinámicos fijan las variables de entradas que no se seleccionan como variables manipuladas en sus valores nominales óptimos, lo cual, como hemos visto, puede conducir a problemas de suboptimalidad o inviabilidad.

Capítulo 4

Selección de Estructuras de Control para Sistemas de Optimización en Tiempo Real

4.1 Introducción

En este capítulo, consideramos el problema de la selección de estructuras de control para sistemas multinivel basados en la economía del proceso en presencia de perturbaciones y restricciones de operación. Como se discutió en la Sección 1.2, la economía de operación se ha incorporado al problema de selección de estructuras de control en base a diferentes criterios. El criterio empleado en esta tesis es seleccionar la estructura de control que minimice el costo operativo económico en estado estacionario garantizando al mismo tiempo la factibilidad del punto operativo en estado estacionario ante la presencia de perturbaciones en la planta. En este contexto, dos de las estrategias principales para la selección de estructuras de control son (i) los métodos basados en márgenes de seguridad, y (ii) los métodos de control auto optimizante. Los métodos basados en márgenes de seguridad buscan estructuras de control y valores de consigna que minimicen la pérdida económica asociada con los márgenes adoptados en las restricciones para evitar que dichas restricciones sean violadas cuando ocurren perturbaciones [8, 20, 22, 23]. Por otro lado, el propósito de los métodos de control auto optimizante es seleccionar variables controladas y valores de consigna constantes de manera que la pérdida económica con respecto al funcionamiento óptimo se minimice en presencia de variaciones en el valor de las perturbaciones [5, 28–30, 36, 69]. Los métodos de control auto optimizante y de los márgenes de seguridad se combinaron en el Capítulo 3 utilizando una formulación de programación no lineal mixta entera (MINLP) que selecciona las variables controladas y sus valores de consigna óptimos minimizando el costo económico promedio, al mismo tiempo que garantiza la satisfacción de las restricciones en presencia de perturbaciones.

El uso de valores de consigna fijos suele ser subóptimo en los casos en que los parámetros del modelo cambian o en los casos en que los valores de las perturbaciones modifican el punto de funcionamiento óptimo de la planta. Este problema generalmente se aborda mediante una estructura de control jerárquica que involucra dos capas [43, 70]. La capa inferior es la capa de control regulatorio del proceso, cuyo propósito es conducir las variables controladas a sus valores de consigna deseados por medio de controladores de retroalimentación. En la capa superior, un algoritmo de optimización en tiempo real (RTO) determina el punto de funcionamiento óptimo en condiciones cambiantes y corrige los valores de consigna de los controladores reguladores. La optimización en tiempo real normalmente procede mediante un enfoque iterativo de dos pasos, que consiste en un paso de actualización del

modelo seguido de un paso de optimización [43, 46, 71]. En el paso de actualización del modelo, se utilizan mediciones para estimar los parámetros inciertos del modelo y los valores de las perturbaciones. En el paso de optimización, el modelo actualizado se utiliza para determinar los valores óptimos de consigna resolviendo un problema de optimización basado en modelo.

Por lo general, los sistemas de optimización en tiempo real se implementan para alcanzar el óptimo una vez que ya se ha definido la estructura de control para la capa de control regulatorio del proceso. En este capítulo, argumentamos que el desempeño económico de la planta controlada puede mejorarse si la estructura de control implementada en la capa de control regulatorio del proceso se selecciona teniendo en cuenta las acciones correctivas de la capa de optimización en tiempo real. Existen algunos trabajos en la literatura en los que la estructura de control (o las variables controladas) se seleccionan teniendo en cuenta las correcciones a los valores de consigna. Por ejemplo, se han propuesto métodos de control auto optimizante para seleccionar variables controladas y sus valores de consigna óptimos en función de las perturbaciones medibles [72, 73]. En [8] se propuso un método de los márgenes de seguridad que integra la capa de control regulatorio y la capa de optimización en tiempo real. La idea en dicho trabajo es permitir que la capa de optimización en tiempo real seleccione la estructura de control y también calcule los valores de consigna como funciones polinómicas de las perturbaciones medidas para mejorar el desempeño económico entre las ejecuciones sucesivas del optimizador en tiempo real. Una estrategia de control auto optimizante basada en la aproximación local, llamada método del espacio nulo, [30] se amplió en [44] explorando la idea de regiones críticas utilizadas en control predictivo basado en modelo (MPC) explícito [74]. Mediante una discretización de las perturbaciones, es posible identificar regiones en el espacio de las perturbaciones en las que se mantienen activas un cierto grupo de restricciones [44]. La idea de identificar las regiones de restricciones activas fue investigada en mayor profundidad por Manum y Skogestad [9], quienes sugirieron como investigación futura una estrategia de control variable aplicando el método del espacio nulo a cada región. En este capítulo, el método de selección de estructuras de control auto optimizantes con márgenes de seguridad de estado estacionario definido en el Capítulo 3 se amplía para incluir la presencia de una capa de control supervisor.

La capa de optimización en tiempo real calcula los valores de consigna óptimos para los controladores de la capa inferior basándose en los valores de las perturbaciones medidas (o estimadas). En presencia de perturbaciones no medidas (o rápidas), se resuelve un problema de márgenes de seguridad en la capa de optimización en tiempo real. En la literatura, se ha propuesto cambiar la estructura de control y los valores de consigna cuando cambia el conjunto de restricciones activas [8, 72, 73]. Teniendo en cuenta estas consideraciones, la formulación que se propone para la selección de estructuras de control incorpora: (i) los cambios en los valores de consigna óptimos calculados por la capa superior de optimización en tiempo real, y (ii) la posibilidad de cambiar la estructura de control y los valores de consigna cuando cambia el conjunto de restricciones activas. Estas extensiones permiten evaluar el beneficio potencial de implementar una capa supervisora de optimización en tiempo real en comparación con el beneficio económico que se puede lograr mediante la implementación de una estructura de control fija con valores de consigna constantes. Además, al integrar las decisiones de la capa de supervisión de optimización en tiempo real en el problema de selección de estructura de control, se muestra en esta tesis que puede ser posible mejorar el rendimiento económico promedio de una planta con control jerárquico de dos capas, en comparación con el enfoque estándar en el que la capa de control regulatorio del proceso y la capa superior de optimización en tiempo real se diseñan secuencialmente.

4.1.1 Optimización en Tiempo Real en una Sistema de Control de Dos Capas

Es de conocimiento que el funcionamiento de una planta controlada se puede mejorar utilizando información en línea para optimizar los valores de consigna y/o cambiar la estructura de control. Estos

cambios suelen ser implementados por una capa de control supervisor, como por ejemplo un sistema de optimización en tiempo real (RTO). En esta sección, argumentamos que el desempeño económico promedio de una planta controlada puede mejorarse aún más si la estructura de control implementada se selecciona teniendo en cuenta las acciones correctivas de la capa de supervisión en el problema de selección de la estructura de control. Esta es la principal motivación para los problemas de selección de estructuras de control introducidos en las secciones 4.2 y 4.3.

4.1.2 Contribuciones de este Capítulo

En esta sección de la tesis se presentarán estrategias para considerar la presencia de una capa supervisora en el diseño de la estructura de control en base a criterios económicos. La selección de la estructura de control en base a criterios económicos según el método de los márgenes de seguridad se integra con la existencia de un control supervisor. Como resultado se presenta una nueva formulación para el problema de selección de estructuras de control en base a criterios económicos. Mediante casos de estudios se muestra la importancia de considerar de manera integral la existencia de las distintas capas del sistema de control. La formulación presentada nos provee una herramienta útil para analizar el potencial beneficio económico de instalar un sistema de optimización en tiempo real. Los aportes de este trabajo son los siguientes:

- Se consideran diferentes formulaciones para la aplicación de cambios de estructuras de control y de los valores de consigna en tiempo real para la formulación ya estudiada de los márgenes de seguridad en estado estacionario. Dichas propuestas se sintetizan en una nueva formulación del tipo MINLP para la selección de estructuras de control que integran la capa de control supervisor.
- Se presenta una aproximación a la formulación original como un problema MIQP local en las variables de entrada y global en las perturbaciones.
- Mediante diferentes casos de estudios, comparamos el desempeño obtenido por las distintas estrategias propuestas. Se muestra que la selección de la estructura de control que tiene en cuenta el nivel de control superior debe hacerse de modo integral si se desea obtener el mayor beneficio económico posible. Se observa la necesidad de estudiar las diversas alternativas de control supervisor para la identificación del mejor sistema de control a instalar en base a criterios económicos.

4.2 Selección de la Estructura de Control para un Sistema de Optimización en Tiempo Real

En plantas químicas de operación continua altamente automatizadas, la operación óptima generalmente se aborda mediante una jerarquía de dos capas de decisión que incluye optimización en tiempo real (RTO) en la capa superior y control regulatorio en la capa inferior (ver Figura 4.1) [71, 75]. Esta estructura en cascada es apropiada en general debido a la separación en la escala de tiempo asociada con las perturbaciones y decisiones que se tratan en cada capa [43]. El papel de la capa de control regulatorio es conducir la planta a los puntos de consigna deseados por medio del control predictivo basado en modelo (MPC) o simplemente controladores descentralizados del tipo proporcional integral derivativos (PID). La capa de control regulador rechaza las perturbaciones de alta frecuencia y logra los objetivos de seguridad y un funcionamiento estable. Por otro lado, la capa RTO calcula el punto de funcionamiento óptimo de estado estacionario en condiciones cambiantes debido a perturbaciones de baja frecuencia y pasa el punto óptimo a la capa de control en forma de

valores de consigna para los controladores. Normalmente, un sistema RTO utiliza mediciones para estimar un conjunto de parámetros ajustables del modelo y los valores de las perturbaciones en el punto de operación actual. A continuación, el modelo del proceso se actualiza en función de estas estimaciones y se utiliza para calcular nuevos valores de consigna óptimos resolviendo un problema de optimización.

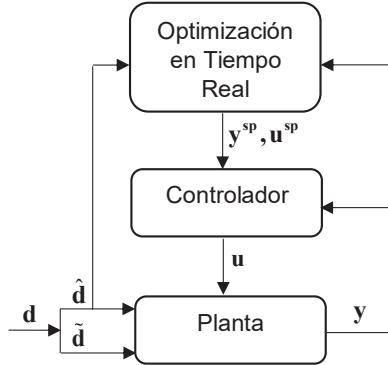


Figura 4.1: Sistema de dos capas con optimización en línea

En este capítulo, clasificamos explícitamente las perturbaciones que afectan el funcionamiento de la planta en las perturbaciones *medibles* $\hat{\mathbf{d}} \in \mathbb{R}^{n_a}$ y las perturbaciones *no medibles* $\tilde{\mathbf{d}} \in \mathbb{R}^{n_a}$. Más precisamente, $\hat{\mathbf{d}}$ incluye perturbaciones de baja frecuencia o de lenta variación que pueden medirse directamente o estimarse en función de mediciones y , por lo tanto, están disponibles en la capa RTO, mientras que $\tilde{\mathbf{d}}$ incluye perturbaciones de alta frecuencia o no medibles que no se actualizan en la capa RTO.

En consecuencia, definimos los conjuntos de perturbaciones $\hat{\mathcal{D}} = \{\hat{\mathbf{d}} \in \mathbb{R}^{n_a} : \hat{\mathbf{d}}^L \leq \hat{\mathbf{d}} \leq \hat{\mathbf{d}}^U\}$, y $\tilde{\mathcal{D}} = \{\tilde{\mathbf{d}} \in \mathbb{R}^{n_a} : \tilde{\mathbf{d}}^L \leq \tilde{\mathbf{d}} \leq \tilde{\mathbf{d}}^U\}$.

El propósito de la capa RTO es calcular los valores de consigna óptimos para los controladores de la capa inferior basándose en el conocimiento de las perturbaciones medidas (o estimadas) $\hat{\mathbf{d}}$, garantizando al mismo tiempo la factibilidad del estado estacionario para todos los posibles escenarios de las perturbaciones no medidas (o rápidas) $\tilde{\mathbf{d}}$. Comúnmente, los sistemas RTO se implementan para optimizar el punto de operación de la planta una vez que la estructura de control en la capa inferior ya ha sido definida. En esta sección, se argumenta que el rendimiento logrado por la capa superior de RTO está limitado por la estructura de control implementada en la capa inferior de control regulatorio. Por lo tanto, se propone un problema de selección de estructuras de control basado en el método de los márgenes de seguridad que incorpore los cambios en los valores de consigna óptimos producidos por la capa superior de RTO.

Basado en la partición de las perturbaciones $\mathbf{d} = [\hat{\mathbf{d}}; \tilde{\mathbf{d}}]$, redefinimos los mapeos de entrada-salida de estado estacionario como $\mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}})$ y $J(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}})$. Con esta nueva definición y considerando los cambios de los valores de consigna según el ingreso de perturbaciones medibles podemos escribir sin simplificaciones el problema de selección de estructura de control como:

$$\begin{aligned}
 \min_{\mathbf{z}, \mathbf{r}^{\text{sp}}} \quad & \mathbb{E}_{[\hat{\mathbf{d}}; \tilde{\mathbf{d}}] \in \mathcal{D}} \left\{ J(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}}) \right\} \\
 \text{s.t.} \quad & \mathbf{r}(\mathbf{z}, \mathbf{u}, \mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}})) = \mathbf{r}^{\text{sp}}(\mathbf{z}, \hat{\mathbf{d}}), \quad \forall [\hat{\mathbf{d}}; \tilde{\mathbf{d}}] \in \mathcal{D} \\
 & \mathbf{y}^L \leq \mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}}) \leq \mathbf{y}^U, \quad \forall [\hat{\mathbf{d}}; \tilde{\mathbf{d}}] \in \mathcal{D} \\
 & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U,
 \end{aligned} \tag{4.1}$$

Para poder resolver este problema procedemos, como en el capítulo anterior, aproximando el

problema mediante una discretización del espacio de perturbaciones. Sea $\hat{\mathcal{D}}$ una discretización uniforme de dimension \hat{N} del conjunto de perturbaciones medibles $\hat{\mathcal{D}}$, es decir, $\hat{\mathcal{D}} = \{\hat{\mathbf{d}}_1, \dots, \hat{\mathbf{d}}_{\hat{N}}\}$, y sea $\tilde{\mathcal{D}}$ una discretización uniforme de dimension \tilde{N} del conjunto de perturbaciones no medibles $\tilde{\mathcal{D}}$, es decir, $\tilde{\mathcal{D}} = \{\tilde{\mathbf{d}}_1, \dots, \tilde{\mathbf{d}}_{\tilde{N}}\}$. Tenga en cuenta que el número total de escenarios discretos viene dado por $N = \hat{N} \times \tilde{N}$. De esta manera, el problema de los márgenes de seguridad en estado estacionario propuesto con cambios óptimos de valores de consigna debido a las perturbaciones medibles se puede formular de la siguiente manera:

$$\begin{aligned} \min_{\mathbf{z}, \mathbf{r}^{\text{SP}}(\mathbf{z}, \hat{\mathbf{d}})} \quad & \sum_{k=1}^{\hat{N}} \sum_{l=1}^{\tilde{N}} p_{k,l} J(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \\ \text{s.t.} \quad & \left\{ \begin{array}{l} \mathbf{r}(\mathbf{z}, \mathbf{u}_{k,l}, \mathbf{y}(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)) = \mathbf{r}^{\text{SP}}(\mathbf{z}, \hat{\mathbf{d}}_k), \\ \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \leq \mathbf{y}^{\text{U}}, \\ \mathbf{u}^{\text{L}} \leq \mathbf{u}_{k,l} \leq \mathbf{u}^{\text{U}} \end{array} \right\} \quad \begin{array}{l} k = 1, \dots, \hat{N}, \\ l = 1, \dots, \tilde{N}. \end{array} \end{aligned} \quad (4.2)$$

Al igual que en el método de los márgenes de seguridad, la satisfacción de las restricciones se impone para todas las posibles realizaciones de las perturbaciones. La diferencia entre esta estrategia de variación de los valores de consigna y el problema de los márgenes de seguridad dado por la ecuación (3.1) es que los valores de consigna son ahora una función de las perturbaciones medidas $\hat{\mathbf{d}}$. De esta forma, la estructura de control obtenida mediante la optimización tiene en cuenta la posibilidad de cambios en los valores consigna óptimos debido al efecto de las perturbaciones medidas. Esto significa que la estructura de control obtenida es óptima para el sistema de dos capas que comprende la capa de optimización en tiempo real y la capa de control regulatorio del proceso. Nos referiremos al Problema (4.1) o al (4.2) como un problema de márgenes de seguridad más RTO (BO + RTO).

Una vez que se ha encontrado la estructura de control óptima \mathbf{z}^* resolviendo el Problema (4.2), el sistema de control regulatorio se implementa en la planta, por ejemplo, mediante el uso de controladores PID, y la capa RTO se implementa en la capa superior. En una ejecución de RTO dada, con las estimaciones de las perturbaciones $\hat{\mathbf{d}}$, se resuelve el siguiente problema en la capa de RTO:

$$\begin{aligned} \min_{\mathbf{r}^{\text{SP}}(\mathbf{z}^*, \hat{\mathbf{d}})} \quad & \sum_{l=1}^{\tilde{N}} p_l(\hat{\mathbf{d}}) J(\mathbf{u}_l, \hat{\mathbf{d}}, \tilde{\mathbf{d}}_l) \\ \text{s.t.} \quad & \left\{ \begin{array}{l} \mathbf{r}(\mathbf{z}^*, \mathbf{u}_l, \mathbf{y}(\mathbf{u}_l, \hat{\mathbf{d}}, \tilde{\mathbf{d}}_l)) = \mathbf{r}^{\text{SP}}(\mathbf{z}^*, \hat{\mathbf{d}}), \\ \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}_l, \hat{\mathbf{d}}, \tilde{\mathbf{d}}_l) \leq \mathbf{y}^{\text{U}}, \\ \mathbf{u}^{\text{L}} \leq \mathbf{u}_l \leq \mathbf{u}^{\text{U}} \end{array} \right\} \quad l = 1, \dots, \tilde{N}. \end{aligned} \quad (4.3)$$

donde $p_l(\hat{\mathbf{d}})$ son los pesos de las perturbaciones no medidas dadas las perturbaciones medidas $\hat{\mathbf{d}}$. Tenga en cuenta que el Problema (4.3) es un problema de márgenes de seguridad, ya que calcula valores de consigna óptimos mientras que presta atención a la satisfacción de las restricciones para el conjunto de escenarios de las perturbaciones no medidas.

Observación 5. La formulación a la que se llega en el Problema (4.2) comparte cierta similitud con la formulación propuesta por Kookos [8], donde los valores de consigna de las variables controladas varían en el tiempo y se calculan como funciones polinómicas de las perturbaciones medidas. Los coeficientes de las funciones polinomiales se incluyen como variables de decisión en el problema de optimización. El Problema (4.2), por otra parte, calcula los valores de consigna óptimos para cada realización de las perturbaciones medidas discretizadas. Por lo tanto, la formulación aquí propuesta imita lo que normalmente hace la capa RTO, que es calcular los valores de consigna óptimos para los controladores de la capa inferior en función de los valores de las perturbaciones medidas o las estimaciones de los parámetros del modelo.

4.3 Selección de Estructuras de Control Basada en Regiones de Restricciones Activas

En esta sección, el método de selección de estructuras de control dado por Problema (3.1) se amplía para incluir información sobre las regiones de restricciones activas. Se consideran dos formulaciones alternativas: (i) utilizar una estructura de control invariante con diferentes valores de consigna para cada región de restricciones activas, y (ii) usar diferentes estructuras de control y valores de consigna para cada región de restricciones activas. Se supone disponible una capa de control supervisor que identifica la región de restricciones activas actual y cambia la estructura de control y/o los valores de consigna, como se indica en la Figura 4.2.

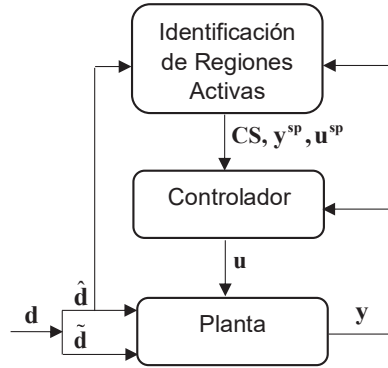


Figura 4.2: Sistema de dos capas con identificación de la región de restricciones activas.

4.3.1 Regiones de Restricciones Activas

Comenzaremos por definir los conceptos de *superficie de las restricciones activas* y *región de restricciones activas* en el espacio de las perturbaciones.

Definición 3 (Superficie de una restricción activa). Para una restricción dada $g(\mathbf{u}, \mathbf{d}) \leq 0$ la superficie de la restricción, denotada por \mathcal{S}_g , es la superficie en el espacio de las perturbaciones que separa las regiones donde la restricción se encuentra activa en el punto de operación óptimo de las regiones donde la restricción se encuentra inactiva en el punto de operación óptimo.

Matemáticamente, la superficie de una restricción activa se puede definir de la siguiente manera:

$$\mathcal{S}_g = \{\mathbf{d} \in \mathcal{D} : g(\mathbf{u}^*(\mathbf{d}), \mathbf{d}) = 0, \forall \epsilon > 0 \exists \boldsymbol{\delta} \in B_\epsilon(\mathbf{d}) \text{ tal que } g(\mathbf{u}^*(\boldsymbol{\delta}), \boldsymbol{\delta}) < 0\},$$

donde $B_\epsilon(\mathbf{d}) = \{\boldsymbol{\delta} \in \mathbb{R}^{n_d} : \|\boldsymbol{\delta} - \mathbf{d}\| \leq \epsilon\}$ es el *entorno- ϵ* de \mathbf{d} .

Definición 4 (Región de restricciones activas). Una región de restricciones activas $\mathcal{D}_a \subseteq \mathcal{D}$ es la región en el espacio de las perturbaciones para la cual la operación óptima involucra un conjunto de restricciones activas dado \mathcal{A}_a .

En términos matemáticos, la región de restricciones activas \mathcal{D}_a se puede definir de la siguiente manera:

$$\mathcal{D}_a = \{\mathbf{d} \in \mathcal{D} : \mathcal{A}(\mathbf{u}^*(\mathbf{d})) = \mathcal{A}_a\}$$

Las diferentes regiones de restricciones activas están separadas por las superficies de restricciones activas. Suponiendo que hay un total de n_r regiones de restricciones activas diferentes, el conjunto

de las perturbaciones \mathcal{D} se puede descomponer como $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \dots \cup \mathcal{D}_{n_r}$, con $\mathcal{D}_i \cap \mathcal{D}_j = \emptyset$, para $i \neq j$. Las regiones de restricciones activas se pueden encontrar aproximadamente resolviendo el Problema (2.5) sobre una cuadrícula definida de escenarios de las perturbaciones. El conjunto de perturbaciones discretizadas $\underline{\mathcal{D}}$ se puede descomponer como $\underline{\mathcal{D}} = \underline{\mathcal{D}}_1 \cup \underline{\mathcal{D}}_2 \cup \dots \cup \underline{\mathcal{D}}_{n_a}$, con $\underline{\mathcal{D}}_i \cap \underline{\mathcal{D}}_j = \emptyset$, para $i \neq j$, donde $\underline{\mathcal{D}}_a$ incluye los escenarios de perturbaciones discretas en $\underline{\mathcal{D}}$ correspondientes a la región de restricciones activas \mathcal{D}_a , para $a = 1, \dots, n_a$. Notar que $n_a \leq n_r$ ya que la discretización podría no capturar todas las regiones de restricciones activas existentes.

Ejemplo 6. Como ejemplo, consideremos el caso de un problema de optimización con 2 perturbaciones, $\mathbf{d} = [d_1, d_2]^\top$, y tres restricciones $\mathbf{g} = [g_1, g_2, g_3]^\top$. Consideremos un conjunto de perturbaciones de tipo caja para \mathcal{D} que se discretiza en $\underline{\mathcal{D}}$ usando una cuadrícula uniforme que comprende $N = 7 \times 7$ escenarios de perturbaciones. En la Figura 4.3 se ilustra un ejemplo con cuatro regiones de restricciones activas diferentes. Los conjuntos activos correspondientes se sintetizan en la Tabla 4.1. La superficie de restricción \mathcal{S}_{g_2} separa la región II de las regiones I y III, mientras que la superficie de restricción \mathcal{S}_{g_3} separa la región III de las regiones I y II. Ambas superficies de restricción coinciden en la línea que separa las regiones II y III. De hecho, esta línea que separa las regiones II y III corresponde a una región activa IV. Nótese también que la superficie de restricción entre las regiones I y II pertenece a la región II, mientras que la superficie de restricción entre las regiones I y III pertenece a la región III.

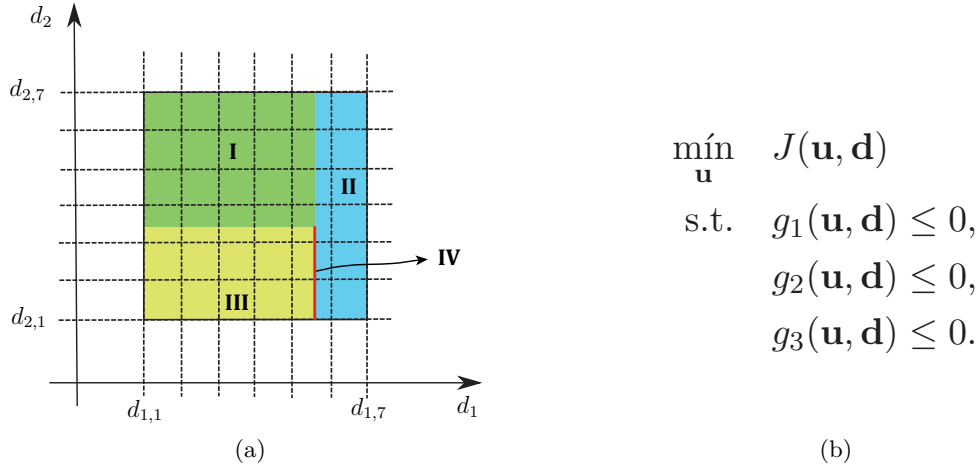


Figura 4.3: Regiones de restricciones activas en el Ejemplo 6. Las restricciones activas son g_1 en la región I, g_1 y g_2 en la región II, g_1 y g_3 en la región III, y g_1 , g_2 y g_3 en la región IV.

Tabla 4.1: Conjuntos de Restricciones Activas en Ejemplo 6.

	\mathcal{A}_I	\mathcal{A}_{II}	\mathcal{A}_{III}	\mathcal{A}_{IV}
g_1	1	1	1	1
g_2		2		2
g_3			3	3

4.3.2 Estructura de Control con Valores de Consigna Variables

La idea es seleccionar la estructura de control óptima que sea factible para todas las regiones de restricciones activas, con la ventaja de cambiar los valores de consigna cuando se detecta una región de restricciones activas diferente. Este problema introduce nuevas variables continuas (valores de consigna para cada región de restricciones activas) al problema MINLP original (3.1). El problema se puede

4.4 Formulación General para la Selección de Estructuras de Control en Sistemas con Optimización en Tiempo Real

En esta sección, presentamos una formulación generalizada del problema de selección de estructuras de control que incluye las diferentes estrategias discutidas en la Secciones 4.2 y 4.3, y permite su combinación. La formulación generalizada incluye las siguientes características:

- Contempla las diferentes regiones de restricciones activas.
- Clasificación de las perturbaciones como medibles o no medibles.
- Valores de consigna variables basados en los valores de las perturbaciones medidas y el conocimiento de las regiones de restricciones activas.
- Estructuras de control variables para cada región de restricciones activas.

El problema de selección de estructura de control generalizado se puede escribir como:

$$\begin{aligned}
 \min_{\mathbf{z}_a, \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}})} \quad & \mathbb{E}_{[\hat{\mathbf{d}}; \tilde{\mathbf{d}}] \in \mathcal{D}} \left\{ J(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}}) \right\} \\
 \text{s.t.} \quad & \mathbf{r}_a(\mathbf{z}_a, \mathbf{u}, \mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}})) = \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}}), \quad \forall [\hat{\mathbf{d}}; \tilde{\mathbf{d}}] \in \mathcal{D}_a, \quad a = 1, \dots, n_r \\
 & \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}, \tilde{\mathbf{d}}) \leq \mathbf{y}^{\text{U}}, \\
 & \mathbf{u}^{\text{L}} \leq \mathbf{u} \leq \mathbf{u}^{\text{U}},
 \end{aligned} \tag{4.8}$$

Al igual que en los problemas anteriores, se plantea la discretización del espacio de perturbaciones para obtener una aproximación:

$$\begin{aligned}
 \min_{\mathbf{z}_a, \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}})} \quad & \sum_{k=1}^{\hat{N}} \sum_{l=1}^{\tilde{N}} p_{k,l} J(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \\
 \text{s.t.} \quad & \left\{ \begin{array}{l} \mathbf{r}_a(\mathbf{z}_a, \mathbf{u}_{k,l}, \mathbf{y}(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)) = \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}}_k), \\ \mathbf{y}^{\text{L}} \leq \mathbf{y}(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \leq \mathbf{y}^{\text{U}}, \\ \mathbf{u}^{\text{L}} \leq \mathbf{u}_{k,l} \leq \mathbf{u}^{\text{U}} \end{array} \right\} \quad \forall [\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \mathcal{D}_a, \\
 & \quad \quad \quad a = 1, \dots, n_a.
 \end{aligned} \tag{4.9}$$

Las diferentes alternativas para la selección de estructuras de control se pueden obtener simplificando el Problema (4.9) o, alternativamente, agregando restricciones adicionales, como se describe a continuación:

- Una estructura de control invariante para todas las regiones de restricciones activas se impone mediante las restricciones::

$$\mathbf{z}_1 = \mathbf{z}_2 \dots = \mathbf{z}_{n_a} = \mathbf{z} \tag{4.10}$$

Las variables controladas y las variables manipuladas son invariantes para todas las regiones de restricciones activas, pero los valores de consigna pueden cambiar en función de las perturbaciones medidas para las diferentes regiones. El uso de (4.10) como única restricción adicional, implica una combinación del método BO + RTO de la Subsección 4.2 y la política de valores de consigna variantes para cada región de restricciones activas de la Subsección 4.3.2.

- Una estructura de control invariante y valores de consigna variantes en función de las perturbaciones medidas (método BO + RTO de la Subsección 4.2). Además de (4.10), se agregan las siguientes restricciones:

$$\mathbf{r}_1^{\text{SP}}(\mathbf{z}_1, \hat{\mathbf{d}}_k) = \mathbf{r}_2^{\text{SP}}(\mathbf{z}_2, \hat{\mathbf{d}}_k) = \dots = \mathbf{r}_{n_a}^{\text{SP}}(\mathbf{z}_{n_a}, \hat{\mathbf{d}}_k) = \mathbf{r}^{\text{SP}}(\mathbf{z}, \hat{\mathbf{d}}_k) \tag{4.11}$$

- Una estructura de control invariante con valores de consigna variantes para cada región de restricciones activas (la estrategia de la Subsección 4.3.2). Además de (4.10), se agregan las siguientes restricciones:

$$\mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}}_k) = \mathbf{r}_a^{\text{SP}}(\mathbf{z}), \quad a = 1, \dots, n_a. \quad (4.12)$$

- Estructura de control y valores de consigna variantes para cada región de restricciones activas (la estrategia de la Subsección 4.3.3).

$$\mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}}_k) = \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a), \quad a = 1, \dots, n_a. \quad (4.13)$$

4.4.1 Estrategia de Solución del Problema General de Selección de Estructuras de Control con Optimización en Tiempo Real

El Problema (4.9) debe reformularse de modo que se pueda resolver utilizando software y solvers comerciales. El desarrollo se presentará para las estructuras de control clásicas, para ello, se define el vector de variables binarias \mathbf{z}_a que determina la estructura de control para la región de restricciones activas a , y los valores de consigna correspondientes \mathbf{r}_a^{SP} para todas las variables de entrada y salida como:

$$\mathbf{z}_a = \begin{bmatrix} \mathbf{z}_a^{\text{O}} \\ \mathbf{z}_a^{\text{I}} \end{bmatrix}, \quad \mathbf{z}_a^{\text{O}} \in \mathbb{B}^{n_y}, \quad \mathbf{z}_a^{\text{I}} \in \mathbb{B}^{n_u}, \quad (4.14)$$

$$\mathbf{r}_a^{\text{SP}}(\hat{\mathbf{d}}) = \begin{bmatrix} \mathbf{y}_a^{\text{SP}}(\hat{\mathbf{d}}) \\ \mathbf{u}_a^{\text{SP}}(\hat{\mathbf{d}}) \end{bmatrix}, \quad \mathbf{y}_a^{\text{SP}} \in \mathbb{R}^{n_y}, \quad \mathbf{u}_a^{\text{SP}} \in \mathbb{R}^{n_u}, \quad (4.15)$$

donde \mathbf{z}_a^{O} y \mathbf{z}_a^{I} corresponden a las variables de entrada y salida, respectivamente. Se selecciona una variable de salida como variable controlada si el elemento binario correspondiente en \mathbf{z}_a^{O} es igual a uno, mientras que una variable de entrada se selecciona como variable manipulada si el elemento binario correspondiente en \mathbf{z}_a^{I} es igual a cero.

Como se mencionó anteriormente, el conjunto de perturbaciones \mathcal{D} esta determinado por restricciones de tipo caja $\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{n_d} : \mathbf{d}^{\text{L}} \leq \mathbf{d} \leq \mathbf{d}^{\text{U}}\}$. A su vez, el conjunto de perturbaciones discretizadas $\underline{\mathcal{D}}$ consiste en una grilla multidimensional generada a partir de la discretización uniforme de cada perturbación d_i en el intervalo $[d_i^{\text{L}}, d_i^{\text{U}}]$, para $i = 1, \dots, n_d$.

Sea N_i la dimensión de la discretización de la perturbación d_i , entonces la cantidad total de escenarios en $\underline{\mathcal{D}}$ esta dado por $N = N_1 \times N_2 \times \dots \times N_{n_d}$. Recuerde también que las perturbaciones se dividen en perturbaciones medidas y no medidas como $\mathbf{d} = [\hat{\mathbf{d}}; \tilde{\mathbf{d}}] = [d_1, d_2, \dots, d_{n_{\hat{\mathbf{d}}}}, d_1, \dots, d_{n_{\tilde{\mathbf{d}}}}]^{\text{T}}$, con $n_d = n_{\hat{\mathbf{d}}} + n_{\tilde{\mathbf{d}}}$.

Para cada variable de entrada y salida, la formulación general debe calcular valores de consigna óptimos independientes para cada región de restricciones activas. Para ello, introduciremos las matrices de identificación de regiones $\mathbf{R}_a \in \mathbb{B}^{N_1 \times N_2 \times \dots \times N_{n_d}}$, $a = 1, \dots, n_a$. Las entradas de \mathbf{R}_a son todos los escenarios discretos de las perturbaciones que se encuentran en $\underline{\mathcal{D}}$. La matriz \mathbf{R}_a informa qué valores de las perturbaciones están asociados con una región de restricciones activas determinada. La entrada de \mathbf{R}_a correspondiente a una perturbación $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \underline{\mathcal{D}}_a$ es igual a uno, mientras que la entrada de \mathbf{R}_a correspondiente a una perturbación $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \notin \underline{\mathcal{D}}_a$ es igual a cero. La Figura 4.4 ilustra la construcción de la matriz \mathbf{R}_a para la Región I del Ejemplo 6, en un espacio de perturbaciones bidimensional.

Usamos la notación $R_a(\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l) \in \{0, 1\}$ para denotar la entrada de \mathbf{R}_a correspondiente a la perturbación $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l]$. De esta manera, los valores de consigna para un escenario de perturbación

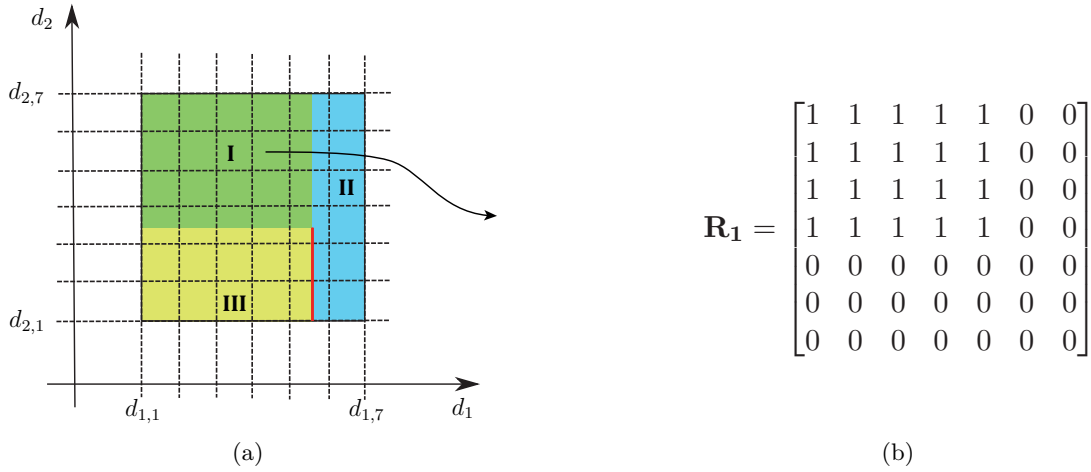


Figura 4.4: Construcción de la matriz de regiones activas

particular pueden obtenerse como:

$$\mathbf{r}^{\text{SP}}(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) = \sum_{a=1}^{n_a} R_a(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}}_k), \quad (4.16)$$

lo cual permite calcular los valores de consigna en función de las perturbaciones medidas para cada una de las regiones de restricciones activas identificadas.

Ejemplo 7. Para ilustrar la forma de la ecuación (4.16), consideremos un problema con dos variables de entrada, $\mathbf{u} = [u_1, u_2]^T$, tres variables de salida medibles, $\mathbf{y} = [y_1, y_2, y_3]^T$, y dos perturbaciones no medibles, $\mathbf{d} = \tilde{\mathbf{d}} = [d_1, d_2]^T$. Para un problema de optimización dado, supongamos que la discretización de las perturbaciones y las regiones de restricciones activas son las indicadas en la Figura 4.4. Por lo tanto, tenemos $n_a = 3$, ya que la región IV no es capturada por la discretización. Suponga que para la región I tenemos $\mathbf{z}_1 = [1, 0, 0, 0, 1]^T$, lo cual significa que para la región I estamos controlando la salida y_1 usando la variable manipulada u_1 , mientras que u_2 se fija a un valor óptimo. Por lo tanto, tenemos $\mathbf{r}_1 = [y_1, u_2]^T$ y $\mathbf{r}_1^{\text{SP}} = [y_1^{\text{SP}}, u_2^{\text{SP}}]^T$. Ahora supongamos que queremos evaluar los valores de consigna que se aplicarán para los valores de perturbación $(d_{1,4}, d_{2,4})$. Para este punto, tendremos $R_1(d_{1,4}, d_{2,4}) = \mathbf{R}_1(4, 4) = 1$, mientras que $R_2(d_{1,4}, d_{2,4}) = R_3(d_{1,4}, d_{2,4}) = 0$. Por lo tanto, la ecuación (4.16) nos dice que, para estos valores de las perturbaciones, debemos aplicar los valores de consigna correspondientes a la región I, es decir $\mathbf{r}^{\text{SP}} = \mathbf{r}_1^{\text{SP}}$, y no los correspondientes a las regiones II o III.

De manera similar, la estructura de control para un escenario de perturbación dado viene dada por:

$$\mathbf{z}(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) = \sum_{a=1}^{n_a} R_a(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \mathbf{z}_a. \quad (4.17)$$

Ahora estamos en condiciones de proporcionar la reformulación del Problema (4.9), que se puede

escribir como:

$$\min_{\mathbf{z}_a, \mathbf{r}_a^{\text{SP}}(\mathbf{z}_a, \hat{\mathbf{d}})} \sum_{k=1}^{\hat{N}} \sum_{l=1}^{\tilde{N}} p_{k,l} J(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \quad (4.18a)$$

$$\text{s.t.} \quad \sum_{i=1}^{n_y} z_{i,a}^{\text{O}} + \sum_{j=1}^{n_u} z_{j,a}^{\text{I}} = n_u, \quad a = 1, \dots, n_a, \quad (4.18b)$$

$$y_i^{\text{L}} z_{i,a}^{\text{O}} \leq y_{i,a,k}^{\text{SP}} \leq y_i^{\text{U}} z_{i,a}^{\text{O}}, \quad i = 1, \dots, n_y, \quad a = 1, \dots, n_a, \quad (4.18c)$$

$$u_j^{\text{L}} z_{j,a}^{\text{I}} \leq u_{j,a,k}^{\text{SP}} \leq u_j^{\text{U}} z_{j,a}^{\text{I}}, \quad j = 1, \dots, n_u, \quad a = 1, \dots, n_a, \quad (4.18d)$$

$$y_{i,k,l}^{\text{SP}} = \sum_{a=1}^{n_a} R_{a,k,l} y_{i,a,k}^{\text{SP}}, \quad i = 1, \dots, n_y, \quad (4.18e)$$

$$u_{j,k,l}^{\text{SP}} = \sum_{a=1}^{n_a} R_{a,k,l} u_{j,a,k}^{\text{SP}}, \quad j = 1, \dots, n_u, \quad (4.18f)$$

$$z_{i,k,l}^{\text{O}} = \sum_{a=1}^{n_a} R_{a,k,l} z_{i,a}^{\text{O}}, \quad i = 1, \dots, n_y, \quad (4.18g)$$

$$z_{j,k,l}^{\text{I}} = \sum_{a=1}^{n_a} R_{a,k,l} z_{j,a}^{\text{I}}, \quad j = 1, \dots, n_u, \quad (4.18h)$$

$$y_i^{\text{L}} (1 - z_{i,k,l}^{\text{O}}) + y_{i,k,l}^{\text{SP}} \leq y_i(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) \leq y_i^{\text{U}} (1 - z_{i,k,l}^{\text{O}}) + y_{i,k,l}^{\text{SP}}, \quad i = 1, \dots, n_y, \quad (4.18i)$$

$$u_j^{\text{L}} (1 - z_{j,k,l}^{\text{I}}) + u_{j,k,l}^{\text{SP}} \leq u_{j,k,l} \leq u_j^{\text{U}} (1 - z_{j,k,l}^{\text{I}}) + u_{j,k,l}^{\text{SP}}, \quad j = 1, \dots, n_u, \quad (4.18j)$$

$$\forall [\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \underline{\mathcal{D}}.$$

La función objetivo en (4.18a) consiste en minimizar un costo promedio ponderado para los escenarios de perturbación considerados, donde $p_{k,l}$ son los pesos para cada escenario de perturbación. La ecuación (4.18b) restringe el número total de variables de salida controladas y variables de entrada fijas para que coincida con el número total de variables de entrada. Por medio de (4.18c) y (4.18d), los valores de consigna de las variables para cada región a están restringidos por sus valores límite cuando las variables son seleccionadas, o en caso contrario se fijan arbitrariamente en cero. Las ecuaciones (4.18e) a (4.18h) hacen uso de la información de la región de restricciones activas identificada y definen la estructura de control y los valores de consigna correspondientes para el proceso, de acuerdo con los valores de las perturbaciones. Las ecuaciones (4.18i) y (4.18j) obligan a las variables a coincidir con sus valores de consigna cuando se seleccionan, o a respetar sus valores límites cuando no se seleccionan.

Observación 6. Si bien el cálculo de los valores de consigna óptimos en función de las perturbaciones medidas se puede aplicar resolviendo el Problema de RTO (4.3), las formulaciones basadas en regiones de restricciones activas requieren el bloque de identificación esquematizado en la Figura 4.2. La correcta identificación en línea de las regiones de restricciones activas no es trivial sin el conocimiento de las perturbaciones no medidas. Sin embargo, al considerar la posibilidad de variar la estructura de control para cada región de restricciones activas, es posible evaluar cuánto se puede ganar en términos de optimalidad al cambiar la estructura de control. Si para una aplicación en particular la ganancia es significativa, se pueden diseñar controladores tipo override y bloques lógicos basados en el conocimiento de las posibles regiones de restricciones activas, como se propone en Reyes-Lúa y Skogestad [42]. El diseño del bloque de identificación de la región activa en la Figura 4.2 es un tema de investigación en sí mismo y cae fuera del alcance de esta tesis. En la Subsección 4.6.4 se proporciona un compendio

de la literatura sobre métodos para cambiar la estructura de control cuando cambian las restricciones activas.

4.5 Aproximación como un Problema MIQP

En general, los problemas MINLP son difíciles de resolver en presencia de no linealidades y no convexidades. Similar al Problema (3.4), la fuente de no linealidades y no convexidades en el Problema (4.9), y en el Problema reformulado (4.18), proviene de las funciones $J(\mathbf{u}, \mathbf{d})$ y $y_i(\mathbf{u}, \mathbf{d})$, $i = 1, \dots, n_y$. Para tratar esta dificultad, el Problema (3.4) fue aproximado como un MIQP en la Sección 4.2 y se puede aplicar el mismo enfoque para aproximar el Problema (4.18) como un MIQP. Para este propósito, las funciones $J(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ y $y_i(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ en el Problema (4.18) se reemplazan por las siguientes aproximaciones:

$$\tilde{J}(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) = \left. \frac{\partial J}{\partial \mathbf{u}} \right|_{k,l} \delta \mathbf{u}_{k,l} + \frac{1}{2} \delta \mathbf{u}_{k,l}^\top \mathbf{H}_{k,l} \delta \mathbf{u}_{k,l} \quad (4.19a)$$

$$\tilde{y}_i(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) = y_{i,k,l} + \left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_{k,l} \delta \mathbf{u}_{k,l}, \quad i = 1, \dots, n_y, \quad (4.19b)$$

donde $\left. \frac{\partial J}{\partial \mathbf{u}} \right|_{k,l}$ y $\left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_{k,l}$ son evaluadas en la solución óptima $\mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ del Problema (2.5), y $y_{i,k,l} = y_i(\mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l), \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$. La matriz $\mathbf{H}_{k,l}$ es una aproximación positiva definida de la matriz Hessiana de la función Lagrangiana evaluada en $\mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$. Además, las siguientes restricciones deben agregarse al Problema (4.18):

$$\mathbf{u}_{k,l} = \mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l) + \delta \mathbf{u}_{k,l}, \quad \forall [\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \underline{\mathcal{D}}. \quad (4.20a)$$

Los pasos necesarios para aproximar el Problema (4.18) como un MIQP pueden resumirse de la siguiente manera:

- (I) Para todos los escenarios de perturbación $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \underline{\mathcal{D}}$, resolver el Problema NLP (2.5) a fin de obtener los valores de entrada óptimos $\mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ e identificar las restricciones activas asociadas a las perturbaciones $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l]$.
- (II) Evaluar $y_{i,k,l}$ y $\left. \frac{\partial y_i}{\partial \mathbf{u}} \right|_{k,l}$ para $i = 1, \dots, n_y$, $\left. \frac{\partial J}{\partial \mathbf{u}} \right|_{k,l}$, y la aproximación de la matriz Hessiana $\mathbf{H}_{k,l}$, en $\mathbf{u}^*(\hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$, para todos los escenarios de perturbación $[\hat{\mathbf{d}}_k; \tilde{\mathbf{d}}_l] \in \underline{\mathcal{D}}$.
- (III) Identificar las distintas regiones de restricciones activas y construir las matrices \mathbf{R}_a , para $a = 1, \dots, n_a$.
- (IV) Reemplazar $J(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ y $y_i(\mathbf{u}_{k,l}, \hat{\mathbf{d}}_k, \tilde{\mathbf{d}}_l)$ en el Problema (4.18) por las aproximaciones (4.19a) y (4.19b), respectivamente. Agregar las restricciones (4.20) al Problema (4.18).

Una vez que se obtiene la estructura de control \mathbf{z}^* como la solución a la aproximación MIQP del Problema (4.18), es importante reevaluar los valores óptimos correspondientes de los valores de consigna, lo cual se hace fijando \mathbf{z}^* en el Problema (4.18), y resolviendo el Problema (4.18) como un problema NLP a fin de reevaluar los valores de consigna (o verificar los valores encontrados por la aproximación). Para una mejor precisión, este NLP se puede resolver utilizando una discretización más fina del espacio de perturbaciones que la utilizada en el problema de selección de estructura de control.

4.6 Consideraciones Adicionales

4.6.1 Límites en la Matriz de Ganancias Relativas

En el capítulo anterior se agregaron variables adicionales y restricciones lineales al Problema (3.4) para definir un límite superior y un límite inferior en los elementos de la matriz de ganancias relativas (RGA). Esto asegura que, para la estructura de control óptima seleccionada, exista un emparejamiento entre las variables manipuladas y las variables controladas con un grado aceptable de interacción para la implementación de controladores descentralizados. Las mismas variables y restricciones adicionales también se pueden agregar al Problema (4.18).

4.6.2 Múltiples Soluciones

Cualquiera de los problemas de selección de estructura de control formulados podría tener más de una estructura de control con el mismo costo óptimo. Para detectar la presencia de múltiples soluciones óptimas sugerimos eliminar la estructura de control óptima dada por el algoritmo de optimización agregando un corte entero, y re-optimizando hasta que la solución óptima obtenida tenga un costo mayor. A continuación, se puede seleccionar una estructura de control entre las múltiples soluciones en función de otros criterios.

4.6.3 Control Indirecto de Restricciones Activas

Asumimos que todas las variables restringidas son variables medidas (ya sea porque son directamente medibles o porque son funciones conocidas de las variables medidas). Por lo tanto, en nuestra formulación, cualquiera de las variables de salida restringidas se puede seleccionar como una variable controlada. Uno esperaría que si una restricción de salida dada, por ejemplo $y_A \leq y_A^U$, está activa (para todo el conjunto de perturbaciones o en una región activa determinada), entonces la variable de salida correspondiente y_A se seleccionará como variable controlada con su valor de consigna igual al valor límite óptimo, es decir, con $y_A^{sp} = y_A^U$. Sin embargo, según el modelo, podría ser equivalente que el solver de optimización elija controlar otra variable de modo que y_A esté indirectamente activa. Esto es lo que llamamos control indirecto de restricciones activas. Por ejemplo, suponga que el modelo incluye la ecuación $f(y_A, y_B) = 0$. Entonces se podría fijar y_B en un valor de consigna dado y_B^{sp} de manera que a través de la ecuación f tengamos $y_A = y_A^U$. Para el solver, es equivalente en términos de costo económico controlar y_A o y_B . Sin embargo, en la práctica, donde el error de modelado, las perturbaciones y los transitorios dinámicos siempre están presentes, controlar y_B podría significar que la restricción sobre y_A esté siendo violada todo el tiempo, para todo el rango de perturbaciones. Debido a ello, por lo general buscamos evitar controlar en forma indirecta las restricciones activas.

Para verificar si hay restricciones activas que están siendo controladas indirectamente, sugerimos simular la planta controlada para todos los escenarios de perturbaciones discretas utilizando la estructura de control óptima obtenida y observar los valores que alcanzan las distintas variables, o verificar los valores de los multiplicadores de Lagrange (en caso de que los proporcione el solver) en las soluciones óptimas.

4.6.4 Identificación de las Regiones de Restricciones Activas

El bloque de identificación de las regiones activas en la Figura 4.2 es una capa de control supervisor que identifica la región de restricciones activas actual y cambia la estructura de control y/o los valores

de consigna en consecuencia. La idea de cambiar la estructura de control o los valores de consigna cuando las restricciones activas cambian bajo el efecto de perturbaciones no es nueva (ver por ejemplo Maarleveld y Rijnsdorp [40], Arkun y Stephanopoulos [41]). Sin embargo, como se señaló recientemente en Reyes-Lúa y Skogestad [42], todavía hay escasas investigaciones realizadas sobre procedimientos de diseño sistemático de una capa de control supervisor utilizando control avanzado para este propósito. En Reyes-Lúa y Skogestad [42] se identifican tres tipos de casos de conmutación: cambio de variable controlada por otra variable controlada; cambio de variable manipulada por otra variable manipulada; y cambio de variable de entrada fija por una variable de salida controlada. Para cada caso, los autores proponen pautas para construir la capa de control de supervisor basada en controles override y bloques lógicos. Las ventajas de este enfoque son que no requiere ejecutar problemas de optimización en línea y que no identifica explícitamente la región activa actual.

La conmutación según las restricciones activas que utilizan las estrategias clásicas de control avanzado, tales como los controladores override, el control de rango dividido y los selectores lógicos máximo/mínimo, puede ser una solución útil para procesos simples, con solo un número reducido de regiones de restricciones activas diferentes. Sin embargo, el diseño requiere un estudio sistemático y calificado del proceso (alto conocimiento heurístico), y para procesos de mediana a gran escala, la complejidad final de la estrategia de cambio de restricciones controladas puede hacer que el método sea difícil de implementar.

En Manum y Skogestad [9] se presenta una estrategia alternativa para la identificación en línea de regiones de restricciones activas. La idea en dicho trabajo es identificar cambios en la región de las restricciones activas basándose en los valores de las variables controladas y manipuladas, y seleccionar las variables controladas en cada región mediante el método del espacio nulo [30], como combinaciones lineales de las variables de entrada y salida. Los autores hacen un paralelismo con el concepto de función afín por partes (PWA, del inglés *piecewise affine*) utilizado en MPC explícito [76]. Para problemas cuadráticos, muestran que es posible obtener una función descriptiva PWA para monitorear (en estado estacionario) los cambios de la región de restricciones activas.

Hay muy poca investigación en la literatura sobre la identificación en línea de regiones de restricciones activas, y no es intención resolver este desafiante problema en esta tesis.

4.6.5 Control de Restricciones Activas

Cuando hay restricciones activas para todos los escenarios de perturbación, los problemas MINLP (3.4) y (4.2) normalmente seleccionan controlar estas restricciones activas en sus valores límite óptimos. Como se menciona en la Subsección 4.6.3, se debe tener la precaución de evitar el control indirecto de las restricciones activas. Por supuesto, si se sabe a priori que una restricción está activa para todo el conjunto de perturbaciones \mathcal{D} , entonces dicha restricción se puede fijar directamente, reduciendo así el tamaño combinatorio del problema.

Por otro lado, las formulaciones para la selección de estructuras de control basadas en regiones de restricciones activas, estudiadas en la Sección 4.3, ya usan información de las regiones de restricciones activas, que se obtiene al resolver el Problema (2.5) sobre la grilla seleccionada de escenarios de perturbación. Esta información se puede utilizar para controlar directamente las restricciones activas:

- En el caso del Problema 4.5, para el cual la estructura de control es invariante, proponemos controlar directamente las restricciones que están activas en todas las regiones.
- En el caso del Problema 4.7, para el cual se permite que la estructura de control cambie en cada región, proponemos controlar las restricciones activas en cada región.

4.6.6 Dificultad Dimensional

El limitante de la dimensionalidad es un desafío común para todos los problemas de selección de estructuras de control basados en optimización. Los problemas formulados en el presente capítulo pueden resultar difíciles de resolver para problemas de gran escala con un número significativo de perturbaciones y variables manipuladas/controladas. Afortunadamente, a menudo es posible revelar la estructura de control óptima utilizando una discretización gruesa de las perturbaciones. Por ejemplo, en el caso de un modelo lineal del proceso, incluir solo los valores extremos de las perturbaciones es suficiente para garantizar la factibilidad para todo el rango de perturbaciones. Si se utiliza una discretización gruesa en una formulación MINLP, recomendamos volver a calcular los valores de consigna resolviendo el problema de los márgenes de seguridad para la estructura de control óptima obtenida.

Los métodos basados en la identificación en línea de las regiones de restricciones activas están claramente limitados a los casos en los que el número de regiones de restricciones activas es bajo (manejable).

La aproximación propuesta por el MIQP de la Sección 4.5 también ayuda en los casos en que la formulación MINLP es difícil de resolver debido al tamaño del problema o debido a la presencia de no linealidades y no convexidades.

4.6.7 Estructuras de Control de Combinaciones Lineales de Variables

En general, sería posible disminuir aún más la pérdida de optimalidad seleccionando como variables controladas combinaciones lineales de las variables de entrada y salida, como se hace en la literatura sobre control auto optimizante y como se presentó en el Capítulo 3. Sin embargo, el enfoque en este capítulo es el de seleccionar la estructura de control clásica óptima. El problema de seleccionar una estructura de control de combinaciones lineales óptimas permitiendo cambios en la estructura de control y en los valores de consigna puede representar un tema de investigación futura.

4.7 Resultados de Simulación

4.7.1 Sistema Lineal

Consideramos el sistema lineal propuesto en la Sección 3.5.1 con algunos cambios. El comportamiento de la planta en estado estacionario se modela como:

$$\mathbf{y}(\mathbf{u}, \mathbf{d}) = \mathbf{y}_s - \mathbf{A}_s \mathbf{u} - \mathbf{D}_s \mathbf{d},$$

con,

$$\mathbf{A}_s = \begin{bmatrix} -1.00 & 1.00 \\ -1.00 & -1.00 \\ -1.60 & 0.24 \end{bmatrix}, \quad \mathbf{y}_s = \begin{bmatrix} -0.35 \\ -1.00 \\ -1.10 \end{bmatrix},$$

$$\mathbf{D}_s = \begin{bmatrix} -0.2702 & 0 & 0 \\ 0 & 0.3003 & 0 \\ 0 & 0 & -0.3809 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} \tilde{\mathbf{d}} \\ \hat{d} \end{bmatrix},$$

Este sistema posee dos variables de entrada $\mathbf{u} = [u_1, u_2]^T$, tres variables medibles de salida $\mathbf{y} = [y_1, y_2, y_3]^T$, dos perturbaciones no medibles $\tilde{\mathbf{d}} = [\tilde{d}_1, \tilde{d}_2]^T$, y una perturbación medible \hat{d} .

La función objetivo a ser maximizada es $J(\mathbf{u}) = 6.85u_1 + 2.95u_2$, y se deben respetar restricciones en las entradas y en las salidas. El problema de optimización se puede escribir como:

$$\begin{aligned} \underset{\mathbf{u}}{\text{máx}} \quad & 6.85u_1 + 2.95u_2 \\ \text{s.t.} \quad & \mathbf{y}(\mathbf{u}, \mathbf{d}) \leq \mathbf{0}, \\ & 0 \leq u_1, u_2 \leq 1. \end{aligned} \tag{4.21}$$

Tenga en cuenta que D_s es una matriz diagonal, lo que significa que las salidas se ven afectadas independientemente por cada perturbación. El valor nominal de las perturbaciones es cero, es decir, $\tilde{\mathbf{d}}_n = \mathbf{0}$ y $\hat{d}_n = 0$. Asumimos que las perturbaciones se encuentran uniformemente distribuidas en los siguientes rangos: $\tilde{d}_1 \in [-1, 1]$, $\tilde{d}_2 \in [-1, 1]$, $\hat{d} \in [-0.75, 0.75]$. Los mapas de contorno del Problema (4.21) se muestran en la Figura 4.5. La región operativa factible para los valores nominales de las perturbaciones está sombreada en verde. Los límites de las restricciones sobre las variables de salida se representan con líneas azules continuas. El punto **a** corresponde al óptimo nominal, es decir, $\mathbf{a} = \mathbf{u}^*(\tilde{\mathbf{d}}_n, \hat{d}_n)$. En este punto, las restricciones sobre y_1 y y_2 están activas. Sin embargo, cuando las perturbaciones ingresan al sistema (flechas en la Figura 4.5), el conjunto de restricciones activas puede cambiar según como se muevan los límites de las restricciones de salida (las líneas discontinuas azules). El punto **b** es la solución óptima para el problema de los márgenes de seguridad a lazo abierto. Este punto operativo conservador representa la mejor solución a lazo abierto para el peor escenario de perturbación $\tilde{\mathbf{d}} = [1, -1]^T$.

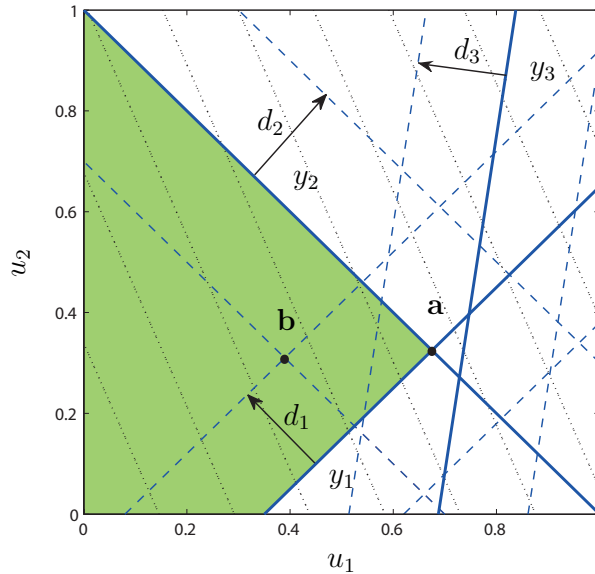


Figura 4.5: Mapa de contorno del sistema lineal. Área verde: región factible para los valores nominales de las perturbaciones; Líneas sólidas azules gruesas: ubicación de los límites de las restricciones de salida para los valores nominales de las perturbaciones; Líneas discontinuas: límites de las restricciones sobre las variables de salida para los valores extremos de las perturbaciones; Líneas punteadas: líneas de nivel de la función objetivo; Punto **a**: óptimo nominal; Punto **b**: solución al problema de los márgenes de seguridad en estado estacionario a lazo abierto. Las flechas muestran cómo se desplazan los límites de las restricciones al aumentar los valores de las perturbaciones d_i , $i = 1, 2, 3$, desde el valor nominal hasta el valor extremo.

Para este sistema lineal, el Problema (4.18) es un MILP, que resolvemos en GAMS usando el solver CPLEX. Se adopta una discretización de grilla uniforme del conjunto de perturbaciones con $\tilde{d}_{1,l} \in \{-1 : 0.5 : 1\}$, $\tilde{d}_{2,l} \in \{-1 : 0.5 : 1\}$, y $\hat{d}_k \in \{-0.75 : 0.375 : 0.75\}$. Como en este ejemplo las restricciones son lineales en \mathbf{u} y \mathbf{d} , sería suficiente considerar solo los valores extremos de las

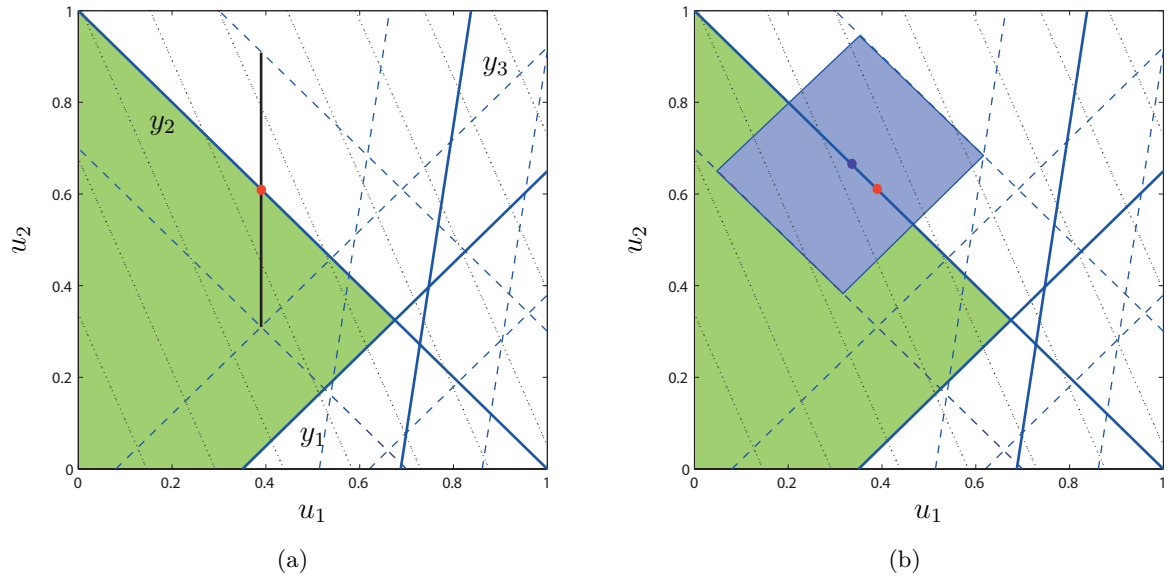


Figura 4.6: Formulación de los márgenes de seguridad aplicada al sistema lineal. **Imagen Izquierda.** Línea negra: Línea operativa para CS3; Punto rojo: punto de funcionamiento nominal para CS3. **Imagen Derecha.** Área azul: región operativa para CS1; Punto Azul: punto de funcionamiento nominal para CS1.

perturbaciones para garantizar la operación factible en todo el conjunto de perturbaciones. El problema de selección de la estructura de control se resolverá utilizando el enfoque de los márgenes de seguridad con funcional económico promedio, dado por el Problema (3.4), y el problema de los márgenes de seguridad con optimización en tiempo real, dado por el Problema (4.2).

4.7.1.1 Solución Utilizando el Método de los Márgenes de Seguridad con Costo Promedio

El enfoque de los márgenes de seguridad de estado estacionario determina la estructura de control óptima con valores de consigna invariantes. El Problema (3.4) se resuelve considerando la discretización previamente definida de las tres perturbaciones \tilde{d}_1 , \tilde{d}_2 , and \hat{d} . No se hace distinción entre las perturbaciones medidas y no medidas, que están incluidas en el vector de perturbaciones $\mathbf{d} = [\tilde{\mathbf{d}}; \hat{d}]$.

La solución obtenida vía optimización consiste en controlar la variable de salida y_2 sin margen de seguridad, es decir, con el valor de consigna $y_2^{\text{sp}} = 0$, y fijando la variable de entrada u_1 con un valor de consigna de $u_1^{\text{sp}} = 0.38975$. Esto corresponde a la estructura de control CS3 de la Tabla 4.2, para la cual el beneficio promedio correspondiente es $J_{\text{av}} = 4.470$. La línea de operación para CS3 y el punto de operación nominal se muestran en el gráfico de la izquierda de la Figura 4.6. Para fines de comparación, el gráfico de la derecha de la Figura 4.6 muestra la región operativa y el punto operativo nominal obtenido para la estructura de control CS1. Esta solución subóptima consiste en controlar y_1 con el el valor de consigna $y_1^{\text{sp}} = -0.696$, es decir, con un margen de seguridad de 0.696, y controlar y_2 con margen de seguridad nulo, es decir, $y_2^{\text{sp}} = 0$.

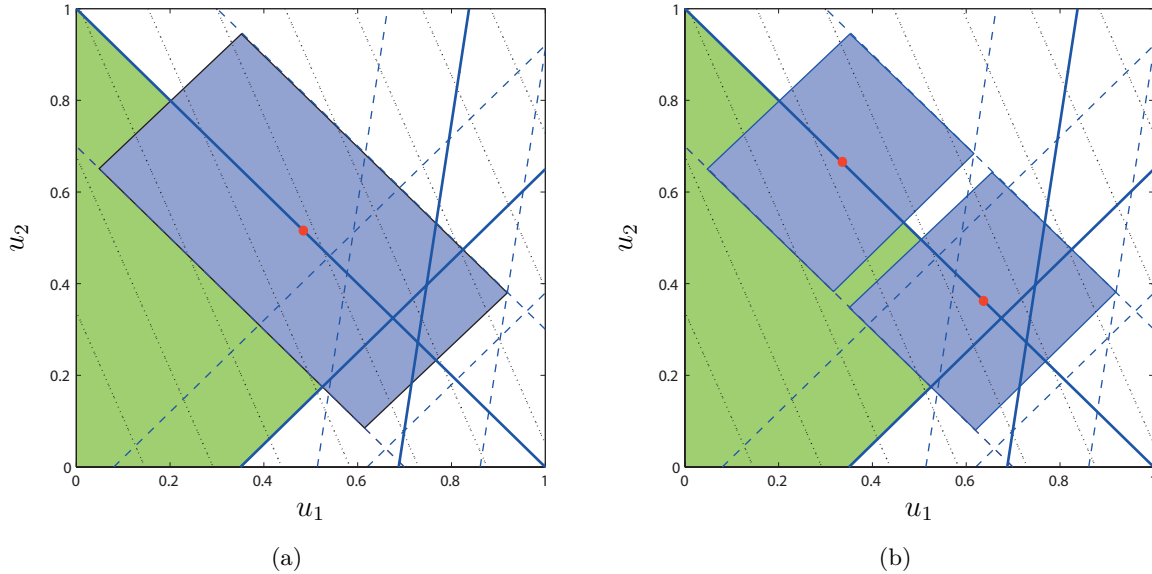


Figura 4.7: Formulación BO + RTO aplicada al sistema lineal. **Imagen Izquierda.** Área azul: región operativa para la estructura de control óptima CS1; Punto rojo: punto de funcionamiento nominal para CS1. **Imagen Derecha.** Áreas azules: regiones operativas de CS1 para los valores extremos de la perturbación medida \hat{d} , es decir, para $\hat{d} = -0,75$ y para $\hat{d} = 0,75$; Puntos rojos: punto de funcionamiento nominal de CS1 para $\hat{d} = -0,75$ y para $\hat{d} = 0,75$.

4.7.1.2 Solución utilizando el Método de los Márgenes de Seguridad con Costo Promedio y Cambios en los Valores de Consigna

A continuación, el método de los márgenes de seguridad con cambios en los valores de consigna óptimos se resuelve utilizando la discretización de las perturbaciones previamente definida y la formulación dada por el Problema (4.2). En este caso, la estructura de control óptima resultante consiste en controlar la variable de salida y_1 con un valor de consigna variable $y_{1,k}^{sp}$, y la variable de salida y_2 sin margen de seguridad, es decir, con $y_{2,k}^{sp} = 0$ para todos los escenarios discretos k de la perturbación medida \hat{d} . Esto corresponde a la estructura de control CS1 de la Tabla 4.2, para la cual el beneficio promedio correspondiente es $J_{av} = 4.8305$. La región operativa para CS1 y el punto operativo nominal se muestran en el gráfico de la izquierda de la Figura 4.7. El gráfico de la derecha de la Figura 4.7 muestra las regiones operativas correspondientes a los valores extremos de la perturbación medida. El valor de consigna óptimo $y_{1,k}^{sp}$ varía linealmente con la perturbación medida \hat{d} , como se ilustra en la Figura 4.8.

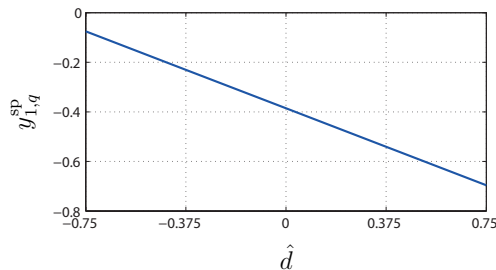


Figura 4.8: Variación de los valores de consigna en el método BO + RTO aplicado al sistema lineal.

4.7.1.3 Comparación de los Métodos de los Márgenes de Seguridad

En este problema de baja dimensión, solo hay diez posibles estructuras de control, que se ordenan en la Tabla 4.2 de acuerdo con sus valores de ganancia económica promedio J_{av} obtenidos mediante el método propuesto de los márgenes de seguridad con optimización en tiempo real (BO+RTO). Para cada estructura de control, los valores de ganancia promedio obtenidos con el método de los márgenes de seguridad con valores de consigna constantes también se muestran en la Tabla 4.2. La mayor ganancia promedio posible para este problema se puede calcular resolviendo el Problema (4.21) para todos los escenarios de perturbaciones y tomando el promedio de las ganancias óptimas, lo que da $J_{av}^* = 5.4112$. Esto nos permite evaluar el porcentaje de pérdida asociado con una estructura de control dada como: $Loss = (J_{av}^* - J_{av})/J_{av}^* \times 100$.

Tabla 4.2: Clasificación de estructuras de control para el sistema lineal.

CS	\mathbf{z}^O			\mathbf{z}^I		BO		BO+RTO	
	y_1	y_2	y_3	u_1	u_2	J_{av}	$\%Loss$	J_{av}	$\%Loss$
CS1	1	1	0	0	0	4.2250	21.9	4.8305	10.7
CS2	0	1	1	0	0	infactible	-	4.6228	14.6
CS3	0	1	0	1	0	4.4700	17.4	4.4700	17.4
CS4	0	1	0	0	1	4.1663	23.0	4.4091	18.5
CS5	0	0	1	0	1	2.3611	56.4	3.5841	33.8
CS6	0	0	1	1	0	infactible	-	3.5841	33.8
CS7	0	0	0	1	1	3.5841	33.8	3.5841	33.8
CS8	1	0	0	1	0	1.8883	65.1	1.8883	65.1
CS9	1	0	1	0	0	infactible	-	infactible	-
CS10	1	0	0	0	1	infactible	-	infactible	-

La estructura de control CS7 corresponde a la solución a lazo abierto, que tiene una pérdida de optimalidad de 33,8 %. Se puede ver en la Tabla 4.2 que al aplicar la estructura de control óptima se da una mejora significativa en las ganancias en comparación con la solución a lazo abierto. La pérdida de optimalidad de la CS3 para el método de los márgenes de seguridad es 17.4 %, mientras que la pérdida de optimalidad de CS1 para el método BO+RTO se reduce a 10.7 %.

Además, es interesante notar que la estructura de control CS2, que no es factible usando el método de los márgenes de seguridad sin considerar la presencia del nivel de optimización en tiempo real, es la segunda mejor estructura de control usando el método BO+RTO. Al comparar la ganancia promedio obtenida usando el problema BO+RTO para las diferentes estructuras de control, está claro que la selección de la estructura de control tiene una influencia importante en el rendimiento en estado estacionario que puede lograrse mediante la capa superior de optimización en tiempo real que calcula los valores de consigna óptimos en función de las perturbaciones medidas. La elección de la mejor estructura de control para un sistema de control jerárquico no es trivial y no se puede encontrar sin considerar explícitamente los cambios en los valores de consigna óptimos en el problema de selección de la estructura de control. Por ejemplo, si se implementa CS3, que corresponde a la estructura de control óptima obtenida sin considerar la capa superior de optimización en tiempo real, no hay nada que ganar al recalcular los valores de consigna óptimos para y_2 y u_1 , ya que para esta estructura de control los valores óptimos de las consignas no cambian con la perturbación medida \hat{d} .

4.7.2 Evaporador

Nuestro siguiente caso de estudio considera el evaporador de circulación forzada descrito por Newell y Lee [68]. El diseño del evaporador se muestra en la Figura 4.9 y sus principales variables con sus

valores nominales se enumeran en la Tabla 4.3. El evaporador se alimenta con la entrada líquida F_1 y el licor de recirculación F_3 . El suministro de vapor P_{100} se utiliza para calentar el evaporador y la mezcla de líquido-vapor generada F_4 se divide en el separador. El líquido separado se recircula en su mayor parte, mientras que una pequeña fracción F_2 se extrae como producto.

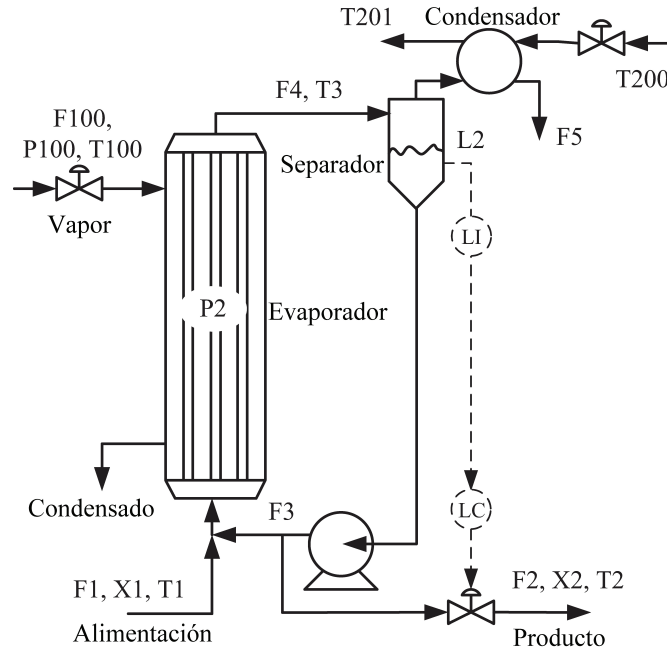


Figura 4.9: Sistema de Evaporador

Las ecuaciones del modelo presentado en Kariwala et al. [31] se encuentran reproducidas en el apéndice de la presente tesis. Tenga en cuenta que el nivel L_2 del separador se mantiene mediante un controlador PI que manipula el caudal de producto F_2 . El problema de selección de estructura de control involucra las siguientes variables de entradas, variables de salidas medidas y perturbaciones no medidas:

$$\begin{aligned} \mathbf{u} &= [F_{200} \quad P_{100} \quad F_3 \quad F_1]^\top \\ \mathbf{y} &= [X_2 \quad F_4 \quad F_5 \quad T_2 \quad T_3 \quad P_2 \quad F_{100} \quad T_{100} \quad T_{201}]^\top \\ \mathbf{d} &= [X_1 \quad T_1 \quad T_{200}]^\top \end{aligned}$$

Los valores nominales de las perturbaciones son: $X_1 = 5$, $T_1 = 40$, $T_{200} = 25$, y sus rangos de variación son $X_1 \in [4.75, 5.25]$, $T_1 \in [32, 48]$, y $T_{200} \in [20, 30]$.

El objetivo es maximizar la ganancia económica de la operación en estado estacionario, que se expresa como:

$$J = 4800F_2 - 0.2F_1 - 600F_{100} - 0.6F_{200} - 1.009(F_2 + F_3),$$

donde el primer término es el valor del producto, mientras que los últimos cuatro términos son los costos operativos. Deben respetarse varias restricciones, incluidos los límites superior e inferior en las entradas, un límite superior en la presión dentro del evaporador, y un límite inferior en la composición

Tabla 4.3: Evaporador: Principales variables y sus valores para el punto operativo nominal.

Variable	Descripción	Valor	Unidad
F_{200}	Caudal del agua refrigerante	217.74	[kg/min]
P_{100}	Presión del vapor de entrada	400.00	[kPa]
F_3	Caudal de circulación	24.72	[kg/min]
F_1	Caudal de alimentación	9.469	[kg/min]
X_2	Composición del producto	35.50	[%]
F_4	Caudal de la mezcla de salida	8.135	[kg/min]
F_5	Caudal del condensado	8.135	[kg/min]
T_2	Temperatura del producto	88.4	[°C]
T_3	Temperatura de la mezcla de salida	81.1	[°C]
P_2	Presión operativa	51.41	[kPa]
F_{100}	Caudal del vapor de entrada	9.434	[kg/min]
T_{100}	Temperatura del vapor de entrada	151.5	[°C]
T_{201}	Temperatura de salida del agua refrigeración	45.55	[°C]
X_1	Composición de la alimentación	5.00	[%]
T_1	Temperatura de la alimentación	40.0	[°C]
T_{200}	Temperatura de la entrada del agua refrigeración	25.0	[°C]
J	Ganancia económica	582.23	[\$/h]

del producto. El problema de optimización se puede escribir como:

$$\begin{aligned}
& \underset{\mathbf{u}}{\text{máx}} && J \\
& \text{s.t.} && \text{ecuaciones del modelo en estado estacionario,} \\
& && X_2 \geq 35.5 \\
& && 40 \leq P_2 \leq 80 \\
& && P_{100} \leq 400 \\
& && 0 \leq F_{200} \leq 400 \\
& && 0 \leq F_1 \leq 20 \\
& && 0 \leq F_3 \leq 100
\end{aligned} \tag{4.22}$$

4.7.2.1 Regiones de Restricciones Activas para el Evaporador

Para visualizar las diferentes regiones de restricciones activas usamos los datos obtenidos al resolver el Problema (4.22) para una cuadrícula definida de escenarios de perturbación. En las soluciones óptimas obtenidas, las restricciones en X_2 y P_{100} siempre están activas para todos los escenarios de perturbación, mientras que la restricción en la variable de salida P_2 cambia de estar activa en su límite superior a estar inactiva, o a estar activa en su límite inferior, dependiendo de los valores de las perturbaciones.

En la Figura 4.10 ilustramos las regiones de restricciones activas identificadas cortando el espacio de las perturbaciones en tres dimensiones por medio de planos para una mejor visualización. Las restricciones activas para cada región son las siguientes: en la región I (azul) las restricciones activas son $X_2 = 35.5$ y $P_{100} = 400$. En la región II (rojo), las restricciones activas son $X_2 = 35.5$, $P_{100} = 400$ y $P_2 = 80$. En la región III (magenta), las restricciones activas son $X_2 = 35.5$, $P_{100} = 400$ y $P_2 = 40$.

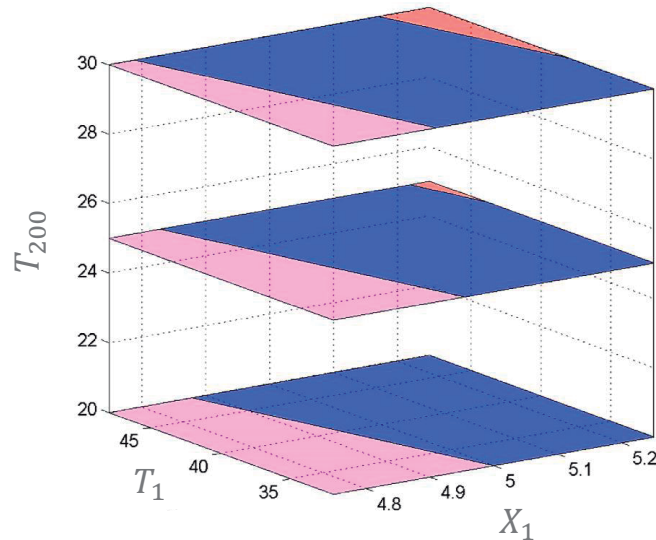


Figura 4.10: Regiones de restricciones activas para el evaporador. Región I: azul; Región II: rojo; Región III: magenta.

4.7.2.2 Comparación de las Estrategias de Selección de Estructuras de Control

El problema de los márgenes de seguridad en estado estacionario con maximización de la ganancia promedio se resuelve utilizando la siguiente discretización de las perturbaciones: $X_1 \in \{4.75 : 0.125 : 5.25\}$, $T_1 \in \{32 : 4 : 48\}$, y $T_{200} \in \{20 : 2.5 : 30\}$, lo que da un total de $N = 125$ escenarios de perturbación. Para mostrar la mejora económica de las estrategias planteadas, consideramos tres formulaciones de selección de estructura de control alternativas: (a) primero resolvemos el Problema (3.4), usando una estructura de control invariante con valores de consigna invariantes, como en el Capítulo 3. A continuación, utilizando la información de las regiones de restricciones activas, buscamos (b) una estructura de control invariante óptima con valores de consigna variantes para cada región activa, y (c) una estructura de control y valores de consigna óptimos para cada una de las tres regiones I, II y III.

Los problemas se resolvieron en GAMS utilizando el solver MINLP local DICOPT, el solver MINLP global BARON y la aproximación MIQP convexa descrita en la Subsección 4.5. Los resultados obtenidos se muestran en las Tablas 4.4 y 4.5. En todos los casos, la restricción sobre la variable de salida $X_2 \geq 35.5$ y el límite sobre la variable de entrada $P_{100} \leq 400$ están activas en las soluciones y las variables X_2 y P_{100} se controlan sin márgenes de seguridad, es decir, con $X_2^{\text{sp}} = 35.5$ y $P_{100}^{\text{sp}} = 400$ respectivamente. Por lo tanto, las diferencias entre las distintas estrategias dependerán de cómo se fijen los dos grados de libertad restantes.

La Tabla 4.4 presenta los resultados obtenidos sin considerar cambios en las estructuras de control y/o en los valores de consigna para diferentes regiones de restricciones activas. Esto corresponde a la implementación de una estructura de control invariante con valores de consigna invariantes, según lo determinado por el Problema (3.4). Observe que la mejor solución se obtuvo con BARON, aunque no puede asegurarse que el solver convergió a la solución óptima ya que no se alcanzó el criterio de corte por brecha relativa y el solver se detuvo después de 50.000 segundos (la solución de DICOPT se utilizó como punto inicial para BARON).

La Tabla 4.5 presenta la estructura de control obtenida usando información de las regiones de restricciones activas. Como se propone en la subsección 4.6.5, las restricciones que están activas en todas las regiones (es decir, X_2 y P_{100}) se controlan directamente en el Problema (4.5), mientras

que las restricciones activas en cada región se controlan directamente en el Problema (4.7). Para ambos problemas, el solver MINLP local DICOPT y la aproximación convexa MIQP obtuvieron la misma solución. El solver local DICOPT convergió en menos de 8 seg., mientras que la aproximación MIQP, resuelta mediante CPLEX, convergió en menos de 2 seg. (Computadora utilizada: Intel (R) Core (TM), CPU i5-4440, 3.19 GHz, 8 GB de RAM)).

En el caso del Problema (4.7), en lugar de controlar las restricciones activas sobre P_2 en las regiones II y III, una solución alternativa con la misma ganancia consiste en controlar $T_2 = 104.5$ en la región II y $T_2 = 82$ en la región III. Esta solución controla indirectamente las restricciones activas en P_2 en las regiones II y III. La posibilidad de obtener soluciones con el mismo beneficio que controlen indirectamente algunas restricciones activas está ligada al modelo utilizado.

Los valores de consigna informados en la Tabla 4.5 para las formulaciones MIQP se reevaluaron fijando la estructura de control correspondiente en el Problema (4.5) o (4.7), y resolviendo el problema para recalcular los valores de consigna como un problema NLP, utilizando la misma discretización de las perturbaciones.

El beneficio económico promedio obtenido para cada estrategia de solución se informa en las Tablas 4.4 y 4.5, junto con la pérdida económica asociada. Puede verse que la política de control de cambiar los valores de consigna para cada región de restricciones activa muestra un beneficio económico considerable en comparación con el enfoque estándar de usar una estructura de control invariante con valores de consigna invariantes. Por otro lado, en esta aplicación particular no hay necesidad de cambiar la estructura de control cuando cambian las regiones de restricciones activas.

Tabla 4.4: Estructura de control y valores de consigna seleccionados para el evaporador sin identificación de regiones activas. *Solver detenido por tiempo.

Método CS Selección		MINLP solver: DICOPT CS				MINLP solver: BARON* CS				MIQP CS			
	\mathbf{r}	X_2	F_3	P_{100}	T_{201}	X_2	F_3	P_{100}	F_{200}	X_2	F_3	P_{100}	T_{201}
Invariante CS & SPs (Problema (3.4))	\mathbf{r}^{SP}	35.5	24.9	400	46.4	35.5	24.7	400	218.2	35.5	24.9	400	46.4
	J_{av}			581.2				584.9				581.2	
	%Loss			4.7				4.1				4.7	

Tabla 4.5: Estructura de control y valores de consigna seleccionados para el evaporador con identificación de regiones activas

Método CS Selección		CS				J_{av}	%Perdida
Invariante CS & Variante SPs (Problema (4.5))	\mathbf{r}	X_2	P_2	P_{100}	F_{200}	602.4	1.2
	\mathbf{r}_I^{SP}	35.5	60.4	400	241.5		
	\mathbf{r}_{II}^{SP}	35.5	80	400	291.8		
	\mathbf{r}_{III}^{SP}	35.5	40	400	159.6		
Variante CSs & SPs (Problema (4.7))	\mathbf{r}_I	X_2	P_2	P_{100}	F_{200}	602.4	1.2
	\mathbf{r}_I^{SP}	35.5	60.4	400	241.5		
	\mathbf{r}_{II}	X_2	P_2	P_{100}	F_{200}		
	\mathbf{r}_{II}^{SP}	35.5	80.0	400	291.8		
	\mathbf{r}_{III}	X_2	P_2	P_{100}	F_{200}		
	\mathbf{r}_{III}^{SP}	35.5	40.0	400	159.6		

4.7.3 Sistema de Reacción-Separación con Reciclo

4.7.3.1 Descripción del Problema

Nuestro caso de estudio final considera el proceso de reacción-separación con reciclo que se utilizó como ejemplo para el control auto optimizante y la identificación de las regiones de restricciones activas

en Jacobsen y Skogestad [44]. Aquí se presenta una breve descripción del sistema [44], y se pueden encontrar más detalles en el apéndice. El diagrama de flujo del proceso se muestra en la Figura 4.11 y las principales variables se enumeran en la Tabla 4.6, junto con sus valores de estado estacionario en el óptimo nominal.

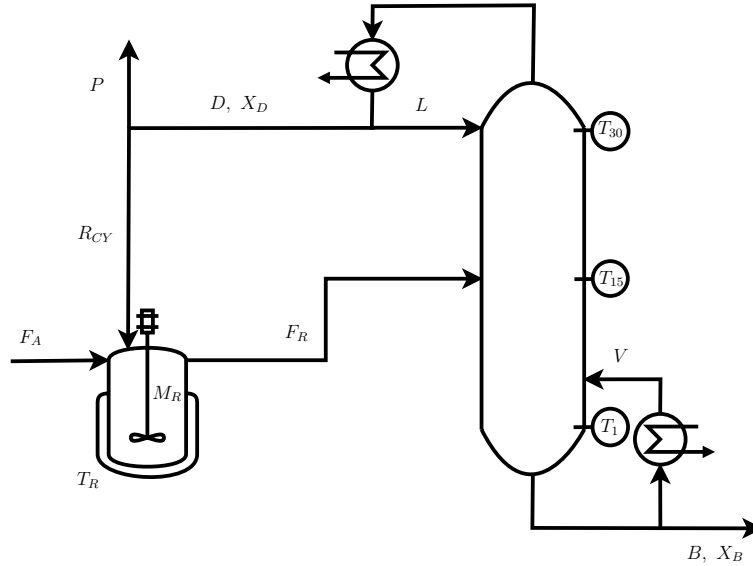
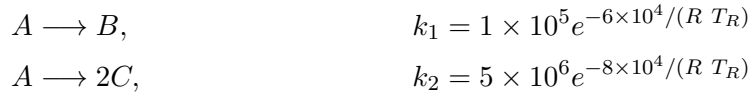


Figura 4.11: Procesos de Reacción-Separación con Reciclo.

El reactor tanque agitado continuo (CSTR) se alimenta con el flujo F_A de producto crudo A y con el flujo de reciclo R_{CY} . Las reacciones que tienen lugar en el CSTR se modelan con cinéticas de primer orden y se muestran a continuación:



donde B es el producto deseado y C un subproducto, R es la constante universal de los gases y T_R la temperatura absoluta del reactor. El flujo de salida del reactor se envía a la columna de destilación. El producto deseado B se separa luego como el producto del fondo, y el subproducto C y el reactivo A sin reaccionar se destilan. El flujo de destilado F_D luego se separa entre el flujo de purga y el de reciclado. El modelo de la columna de destilación se basa en el modelo de Columna 'A' descrito en Skogestad y Morari [77], que considera las siguientes suposiciones: volatilidades relativas constantes, desbordamiento molar constante, presión constante sobre toda la columna, equilibrio líquido-vapor en cada plato, y retención de vapor insignificante. Los parámetros de la columna de destilación se enumeran en la Tabla 4.7.

En este trabajo, asumimos que la retención en el reactor se controla manipulando el flujo de salida del reactor, F_R , mientras que se supone que el caudal de alimentación F_A es una perturbación no medible. Esto deja cuatro grados de libertad como posibles variables manipuladas para mejorar la economía de la operación: L , V , R_{CY} y T_R .

El punto operativo óptimo en estado estacionario para este proceso se obtiene maximizando la siguiente ganancia económica:

$$J = 2B + 0.5P - F_A - p_V V$$

4.7. Resultados de Simulación

Tabla 4.6: Proceso Reacción-Separación con Reciclo: Principales variables y sus valores en el óptimo nominal.

Variable	Descripción	Valor	Unidad
L	Caudal de reflujo en la columna	1.929	[mol/s]
V	Caudal de Vapor en la columna	2.121	[mol/s]
R_{CY}	Caudal de reciclo	0.000	[mol/s]
T_R	Temperatura en el reactor	390.000	[K]
F_R	Caudal de salida del reactor	1.121	[mol/s]
M_R	Volumen de Reactor	11000.000	[mol]
P	Caudal de purga	0.192	[mol/s]
B	Caudal de fondo de Columna	0.927	[mol/s]
$X_{B,A}$	Fracción molar de A en el fondo de columna	0.087	
$X_{B,B}$	Fracción molar de B en el fondo de columna	0.900	
$X_{B,C}$	Fracción molar de C en el fondo de columna	0.013	
D	Caudal de destilado en la columna	0.192	[mol/s]
$X_{D,A}$	Fracción molar de A en el destilado de columna	0.068	
$X_{D,B}$	Fracción molar de B en el destilado de columna	0.071	
$X_{D,C}$	Fracción molar de C en el destilado de columna	0.861	
F_A	Caudal de alimentación de A en el reactor	1.030	[mol/s]
p_V	Precio de la energía	0.060	[\$/mol]
J	Ganancia económica	0.789	[\$/s]

Tabla 4.7: Columna de destilación: parámetros principales

Parámetros	Valor
α_{AC}	0.70
α_{BC}	0.60
NT: número de platos	30
NF: ubicación del plato de alimentación	15
fracción líquida en alimentación	1
V_{max}	30

El problema de optimización con la especificación del producto y las restricciones operativas se puede escribir como:

$$\begin{aligned}
 & \underset{\mathbf{u}}{\text{máx}} && J \\
 & \text{s.t.} && \text{ecuaciones del modelo de estado estacionario,} \\
 & && g_1 = 0.9 - X_{B,B} \leq 0 \\
 & && g_2 = -R_{CY} \leq 0 \\
 & && g_3 = T_R - 390 \leq 0 \\
 & && g_4 = V - 30 \leq 0
 \end{aligned} \tag{4.23}$$

El precio de la energía p_V puede variar en el tiempo y se considerará una perturbación medible. En resumen, las variables del problema de selección de estructura de control son las siguientes:

$$\begin{aligned}
 \mathbf{u} &= [L, V, R_{CY}, T_R]^\top, \\
 \mathbf{y} &= [T_1, T_5, T_9, T_{13}, T_{17}, T_{21}, T_{25}, T_{30}, X_{B,A}, X_{B,B}, X_{B,C}]^\top, \\
 \mathbf{d} &= [\hat{d}, \tilde{d}]^\top = [p_V, F_A]^\top,
 \end{aligned}$$

donde T_i , para $i = 1, \dots, 30$, corresponde a la temperatura en el i -ésimo plato de la columna de destilación.

4.7.3.2 Regiones de Restricciones Activas para el Sistema de Reacción-Separación con Reciclo

El espacio de las perturbaciones se discretiza como una cuadrícula de 20×20 puntos y se resuelve el Problema (4.23) para cada punto. Como resultado, se construye el mapa de las regiones de restricciones activas que se muestra en la Figura 4.12, el cual concuerda con las regiones de restricciones activas identificadas en Jacobsen y Skogestad [44]. Se pueden encontrar cinco regiones de restricciones activas diferentes, que están separadas entre sí por tres curvas de restricciones (en un espacio de perturbación bidimensional, una superficie de restricción se convierte en una curva de restricción). Las variables restringidas que están activas en cada región son las siguientes:

- Region I: $X_{B,B}, R_{CY}$
- Region II: $X_{B,B}, R_{CY}, T_R$
- Region III: $X_{B,B}, T_R$
- Region IV: $X_{B,B}, T_R, V$
- Region V: $X_{B,B}, R_{CY}, T_R, V$

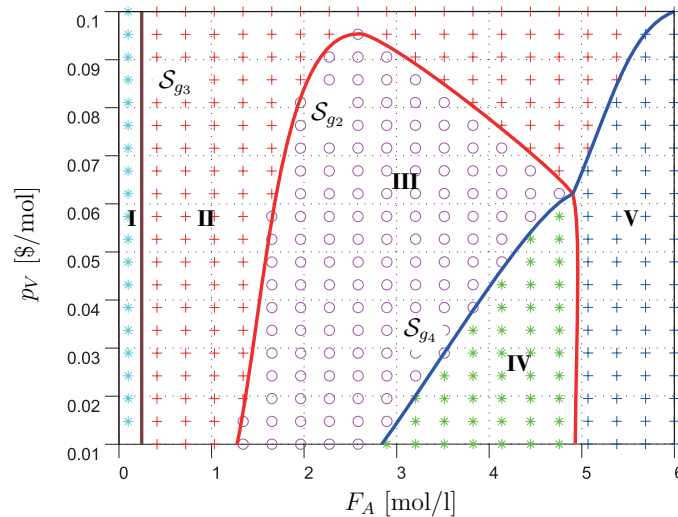


Figura 4.12: Regiones de restricciones activas para el proceso de reactor-separador-reciclo.

4.7.3.3 Estrategias de Solución para el Sistema de Reacción-Separación con Reciclo

El modelo del proceso de reacción-separación con reciclo [44] tiene 98 ecuaciones (lineales y no-lineales) independientes que deben resolverse para los 400 escenarios de perturbación. El problema resultó difícil de resolver con los solvers MINLP disponibles en GAMS. El uso de solvers MINLP locales arroja una variedad de soluciones subóptimas dependiendo del solver y subsolver seleccionados, el punto inicial y los criterios de detención. Por otro lado, el solver global BARON mostró dificultades para disminuir la brecha del criterio de detención. Por lo tanto, se decide resolver las diferentes estrategias de selección de estructura de control utilizando la aproximación MIQP descrita en la subsección 4.5.

Los resultados obtenidos empleando las estrategias que involucran una estructura de control invariante se enumeran en la Tabla 4.8.

Tabla 4.8: Resultados utilizando las estrategias de selección de estructura de control con un estructura invariante. La notación $f(p_V)$ denota que el valor de consigna varía con p_V , mientras que $f(ar)$ denota que el valor de consigna varía para cada región de restricciones activas.

Método Selección de CS	CS					J_{av}	%Loss
CS & SPs invariantes (Problema (3.4))	\mathbf{z} \mathbf{r}^{SP}	$X_{B,b}$ 0.9	L 17.84	R_{CY} 0	T_R 390	-0,054	112,8
CS invariante & SPs variantes (Problema (4.2))	\mathbf{z} \mathbf{r}^{SP}	$X_{B,b}$ 0.9	L $f(p_V)$	R_{CY} $f(p_V)$	T_R 390	0,037	91,2
CS invariante & SPs variantes (Problema (4.5))	\mathbf{z} \mathbf{r}^{SP}	$X_{B,b}$ 0.9	L $f(ar)$	R_{CY} $f(ar)$	T_R $f(ar)$	0,225	46,9
CS invariante & SPs variantes ((4.2) y (4.5) Combinados)	\mathbf{z} \mathbf{r}^{SP}	$X_{B,b}$ 0.9	L $f(ar, p_V)$	R_{CY} $f(ar, p_V)$	T_R $f(ar, p_V)$	0,255	39,9

El método estándar de seleccionar una estructura de control invariante con valores de consigna invariantes ni siquiera es rentable, ya que da como resultado una pérdida promedio de 0.054. Al calcular los valores de consigna variables como funciones de la perturbación medida p_V , la operación del proceso se vuelve económicamente rentable (segundo método en la Tabla 4.8) con una ganancia promedio de 0.037. Sin embargo, la ganancia aumenta significativamente a 0.225 al calcular los valores de consigna óptimos para cada región de restricciones activas (tercer método en la Tabla 4.8). Esto puede explicarse por el hecho de que, para los rangos de perturbaciones estudiados, la perturbación no medida F_A influye más en la ganancia económica que la perturbación medida, y el hecho de que las regiones de restricciones activas cambian principalmente con el valor de dicha perturbación no medida, como se puede inferir de las Figuras 4.12 y 4.13. El cuarto método de la Tabla 4.8 consiste en calcular los valores de consigna óptimos en función tanto de la perturbación medida como del conocimiento de las regiones de restricciones activas, lo que aporta solo una pequeña mejora adicional, con una ganancia de 0.255.

La Figura 4.13 muestra en rojo el perfil óptimo de la relación de reflujo de la columna L en función de la perturbación no medida F_A , para dos valores particulares de la perturbación medida p_v . Los perfiles del valor de consigna correspondientes L^{SP} , calculados utilizando el cuarto método en la Tabla 4.8, se muestran en la misma figura.

Vale la pena notar que las cuatro estrategias que se muestran en la Tabla 4.8 dieron como resultado la misma estructura de control óptima. Lo que cambia de un enfoque a otro es la política a emplear para establecer los valores de consigna.

Finalmente, aplicamos la formulación generalizada completa dada por el Problema (4.9), que selecciona estructura de control y valores de consigna en función de la perturbación medida, para cada región de restricciones activas. Los resultados se dan en la Tabla 4.9. Permitir que la estructura de control varíe no da como resultado ningún cambio en selección de la estructura de control óptima para las regiones I, II y III, mientras que para las regiones IV y V la variable de entrada V se fija en su valor límite superior de 30. La ganancia promedio aumenta ligeramente hasta 0,259.

Observación 7. La ganancia promedio óptima obtenida al resolver el Problema (4.23) para todos los escenarios de perturbación es 0.424, lo que significa que la pérdida de optimización de la estrategia de control de la Tabla 4.9 es 38.9%! Por supuesto, una solución práctica para eliminar esta pérdida de optimalidad sería medir F_A . Sin embargo, en este trabajo asumimos que F_A no se puede medir.

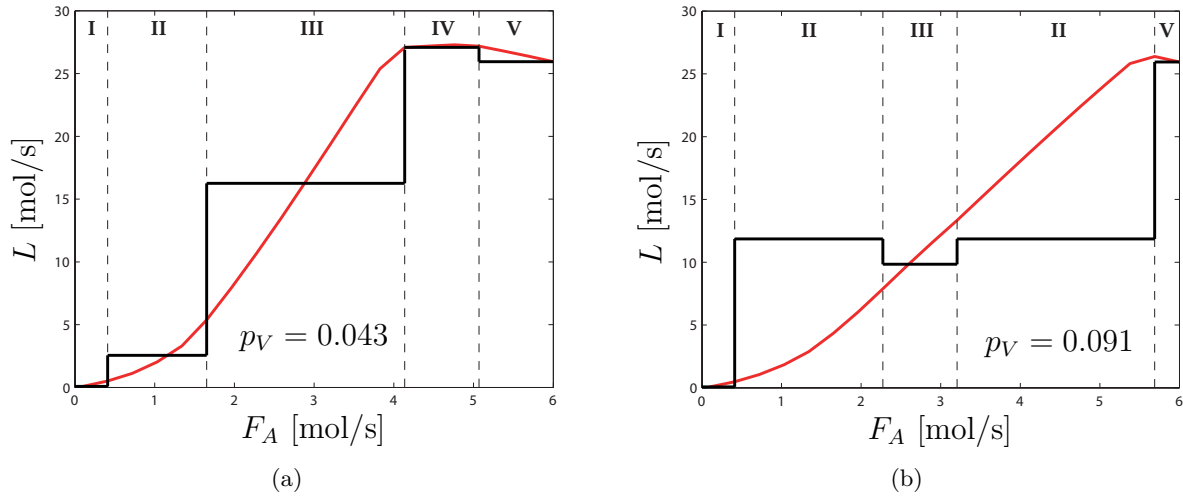


Figura 4.13: Gráfico de la izquierda: $p_V = 0.043$. Gráfico de la derecha: $p_V = 0.091$. Línea roja gruesa: perfil óptimo de L en función de F_A . Línea negra gruesa: perfil del valor de consigna L^{sp} , calculado usando la estrategia con una estructura de control fija y valores de consigna variables (Problemas (4.2) y (4.5) combinados).

Tabla 4.9: Resultados utilizando la formulación generalizada con estructura de control y valores de consigna variantes. La notación $f(p_V)$ denota que el valor de consigna varía en función de p_V .

Región	CS					J_{av}	%Loss
I	\mathbf{z}	$X_{B,b}$	L	R_{CY}	T_R	0,005	0,0
	\mathbf{r}^{sp}	0.9	$f(p_V)$	0	$f(p_V)$		
II	\mathbf{z}	$X_{B,b}$	L	R_{CY}	T_R	-0,033	28,8
	\mathbf{r}^{sp}	0.9	$f(p_V)$	0	390		
III	\mathbf{z}	$X_{B,b}$	L	R_{CY}	T_R	0,239	9,7
	\mathbf{r}^{sp}	0.9	$f(p_V)$	$f(p_V)$	390		
IV	\mathbf{z}	$X_{B,b}$	V	R_{CY}	T_R	0,122	0,4
	\mathbf{r}^{sp}	0.9	30	$f(p_V)$	390		
V	\mathbf{z}	$X_{B,b}$	V	R_{CY}	T_R	-0,074	0,0
	\mathbf{r}^{sp}	0.9	30	0	390		
Total						0,259	38,9

4.8 Discusiones y Conclusiones del Capítulo

En este capítulo se propuso una formulación de selección de estructura de control que optimiza el rendimiento económico que se puede lograr mediante la implementación de un sistema control de dos capas, en presencia de perturbaciones medidas (o lentas) que están disponibles para la capa RTO, y perturbaciones no medidas (o rápidas) que no están disponibles para la capa RTO. La formulación también permite estudiar la ventaja de variar la estructura de control y los valores de consigna utilizando información sobre las regiones de restricciones activas. El problema se definió como un problema de márgenes de seguridad en estado estacionario que encuentra las variables controladas y manipuladas óptimas, y los valores de consigna para las variables controladas y para las variables de entrada restantes que no se seleccionan como variables manipuladas, mientras se presta atención a la satisfacción de las restricciones en presencia de las perturbaciones no medidas. El problema se formuló como un MINLP y se aproximó como un MIQP convexo.

Las diferentes características de las formulaciones propuestas se probaron en tres casos de estudio para los cuales el conjunto de restricciones activas cambia con los valores de las perturbaciones. En el diseño de una estrategia de control con el propósito de la operación óptima, surgen las siguientes preguntas: ¿Es suficiente seleccionar una estructura de control fija con valores de consigna fijos? ¿Cuánto se puede ganar al incluir una capa de supervisión RTO? En caso de que se incluya una capa RTO, ¿es suficiente variar los valores de consigna en función de las perturbaciones medidas o estimadas y los valores de los parámetros? ¿O debería uno también intentar identificar cambios en las restricciones activas y variar la estructura de control y/o los valores de consigna para diferentes regiones de restricciones activas?

Los resultados muestran que la formulación propuesta es una herramienta útil para el análisis y la toma de decisiones en respuesta a estas preguntas. Además, los resultados del sistema lineal demostraron claramente que se pueden lograr mejoras importantes en el desempeño económico promedio de la planta controlada por dos capas si se tienen en cuenta las acciones correctivas de la capa supervisora RTO en el problema de selección de estructura de control, en comparación con el método estándar en el que la capa de control del proceso y la capa RTO se diseñan secuencialmente.

Pueden preverse varias direcciones de investigación futura. La desafiante tarea de identificar en línea los cambios de regiones de restricciones activas en presencia de perturbaciones no medidas sigue sin una estrategia sistemática. En este sentido, sería interesante investigar la identificación de regiones de restricciones activas utilizando redes neuronales artificiales o herramientas de aprendizaje automático. Como se menciona en la Subsección 4.6.7, el método de este artículo fue la selección de estructura de control clásicas óptimas. Los procedimientos para la selección de combinaciones lineales de variables de entrada y salida como variables controladas podrían investigarse para disminuir aún más la pérdida de optimalidad. Además, en esta tesis, la técnica de los márgenes de seguridad se limitó a una formulación de estado estacionario. Sin embargo, sería interesante ampliar el método a un problema de márgenes de seguridad dinámicos para prestar atención también a la satisfacción de las restricciones en la respuesta transitoria de la planta controlada.

Capítulo 5

Optimización en Línea para Modelos Imperfectos

5.1 Introducción

La formulación de un programa de optimización matemática en base a un modelo de proceso es una de las estrategias más utilizadas para mejorar el desempeño operativo de una instalación industrial. Tanto los modelos basados en datos como los basados en primeros principios se utilizan actualmente para calcular las condiciones operativas óptimas:

- Los modelos de primeros principios se basan en balances de masa y energía, propiedades termodinámicas y cinética de reacciones químicas. Se requiere la posterior experimentación para estimar y ajustar los parámetros del modelo.
- Los modelos basados en datos se construyen sobre la información recopilada experimentalmente.

Las técnicas de optimización basadas en modelos son intuitivas y están muy extendidas, pero sufren el efecto del desajuste de modelo de planta. Por ejemplo, el punto de operación óptimo calculado a partir de un modelo de planta inexacto puede ser subóptimo para la planta y podría violar las restricciones de la planta. Además, en presencia de perturbaciones, los valores de consigna óptimos de la planta podrían cambiar y ser necesario modificar los mismos para mantener la operación de la planta en su punto óptimo.

Los Capítulos 3 y 4 de esta tesis se desarrollaron según el supuesto de un modelo de planta perfecto (o que solo posee parámetros inciertos acotados y que se los considera dentro del vector \mathbf{d}) y en base al cual se selecciona la estructura de control. Se han propuesto distintas estrategias de selección incluida la estrategia descrita en la Sección 4.2 que tiene en cuenta la capa de optimización en tiempo real. Sin embargo, si el modelo no se condice con la realidad de la planta la ejecución del problema de RTO de la Sección 4.3 no conducirá al óptimo real de la planta.

En la práctica sucede frecuentemente que los modelos teóricos propuestos presentan errores paramétricos y/o estructurales. En consecuencia, existen esquemas de optimización en tiempo real que se utilizan para alcanzar el punto operativo óptimo de la planta mediante la actualización del modelo basado en mediciones en tiempo real. Estos esquemas varían según las mediciones que se requieren, cómo se actualiza el modelo y los experimentos que se necesitan. Por ejemplo, el enfoque de dos pasos utiliza (i) mediciones para actualizar los parámetros del modelo y (ii) el modelo actualizado para calcular las entradas óptimas ([43]). En lugar de ajustar los parámetros del modelo, el enfoque de adaptación por modificadores (MA) ([6] y [78]) consiste en agregar términos correctivos de primer

orden a la función de costo y a las restricciones del problema de optimización para que las condiciones necesarias de optimalidad de primer orden calculadas empleando el modelo modificado coincidan con las condiciones de optimalidad de la planta. La principal dificultad en el esquema de adaptación por modificadores proviene del hecho de que los gradientes de las restricciones y de la función costo de la planta deben estar disponibles en cada iteración de RTO. Una revisión de varias técnicas para estimar los gradientes y/o calcular los modificadores requeridos se puede encontrar en Marchetti et al. [54].

Ampliando la idea de adaptación por modificadores de primer orden, Faulwasser y Bonvin [53] propusieron agregar términos de corrección de segundo orden a las funciones de costo y restricciones para que las condiciones de optimalidad de segundo orden del modelo modificado coincidan con las de la planta.

Por otra parte, en los métodos donde el modelado de la planta está basado en datos es común llevar a cabo una optimización en una delimitada región de confianza (trust-region optimization). Estos esquemas generalmente proceden de la siguiente manera: (i) la construcción de modelos aproximados basados en mediciones locales (punto operativo y cercanía) y (ii) la optimización de este modelo en una región local de confianza ([79]).

En este capítulo nos adentramos en los llamados esquemas de optimización en tiempo real que tratan las dificultades de optimizar un modelo que difiere de la planta real mediante la utilización de modificadores de segundo orden. La formulación del problema de optimización con adaptación por modificadores se presenta en la Sección 5.2. En la Sección 5.3.1 se formulan los modificadores mediante aproximación cuadrática. En la Subsección 5.3.2 se explica la influencia de la selección de datos, y se discute brevemente el esquema seleccionado. El algoritmo para el esquema RTO que considera los modificadores de segundo orden obtenidos por aproximación cuadrática se describe en la Sección 5.3.3. Para ilustrar los resultados del método propuesto, en la Sección 5.4 el reactor William Otto se simulará en diversos escenarios. Finalmente, las conclusiones y las posibles direcciones de la investigación futura se pueden encontrar en la Sección 5.5.

5.1.1 Contribuciones de este Capítulo

En esta sección de la tesis se presenta un nuevo método para el cálculo y la obtención de los modificadores que se emplean en los esquemas de optimización en tiempo real por adaptación de modificadores. La disponibilidad de modelos imperfectos de los procesos industriales requiere que, para perseguir el óptimo real de la planta, sea necesario la utilización de una estrategia que corrija la diferencia planta-modelo existente. En consecuencia plantearemos una estrategia, que encuadra dentro de los métodos de optimización en tiempo real con adaptación por modificadores, mediante el uso de aproximaciones cuadráticas y regiones de confianza.

- Se considera el uso de métodos desarrollados en la literatura de optimización libre de derivadas (DFO del inglés derivative free optimization) en los esquemas de optimización en tiempo real de procesos industriales.
- Se propone el cálculo de los modificadores, utilizados en los métodos de optimización en tiempo real de adaptación por modificadores, mediante una aproximación cuadrática de la diferencia planta-modelo basada en la información disponible de puntos operativos pasados.
- Mediante casos de estudios se muestra la idoneidad del método propuesto para converger a puntos operativos óptimos en una diversidad de escenarios que pueden presentarse en la práctica industrial.

5.2 Problema de Optimización en Línea por Adaptación de Modificadores para Modelos Imperfectos

Retomamos ahora el Problema de optimización (2.5) planteado en la Sección 2 y supondremos que las perturbaciones son estimables, es decir, $\mathbf{d} = [\hat{\mathbf{d}}]$. La optimización del punto operativo de la planta entonces se puede escribir como:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \hat{\mathbf{d}}) \\ \text{s.t.} \quad & \mathbf{y}^L \leq \mathbf{y}(\mathbf{u}, \hat{\mathbf{d}}) \leq \mathbf{y}^U, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (5.1)$$

Dado que $\mathbf{d} = [\hat{\mathbf{d}}]$ ingresa como un parámetro conocido, se puede simplificar la expresión (5.1) para llegar a las formulaciones de los problemas de optimización comúnmente usadas en la bibliografía de RTO. La optimización de la planta real se escribe como:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J_p(\mathbf{u}) \\ \text{s.t.} \quad & \mathbf{G}_p(\mathbf{u}) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (5.2)$$

Mientras que la optimización del modelo de planta se escribe como:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (5.3)$$

Debido al error de modelado, la solución óptima del Problema (5.3), \mathbf{u}^* , generalmente no coincide con el óptimo de planta \mathbf{u}_p^* , es decir con la solución óptima del Problema (5.2).

En los esquemas de adaptación por modificadores, para manejar el desajuste del modelo de la planta, el Problema (5.3) se adapta iterativamente (recordar la Figura 2.6). En la i -ésima ejecución del RTO se resuelve el siguiente problema:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^J + (\boldsymbol{\lambda}_i^J)^T (\mathbf{u} - \mathbf{u}_i) + \frac{1}{2} (\mathbf{u} - \mathbf{u}_i)^T \mathbf{H}_i^J (\mathbf{u} - \mathbf{u}_i) \\ \text{s.t.} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^G + (\boldsymbol{\lambda}_i^G)^T (\mathbf{u} - \mathbf{u}_i) + \frac{1}{2} (\mathbf{u} - \mathbf{u}_i)^T \mathbf{H}_i^G (\mathbf{u} - \mathbf{u}_i) \leq \mathbf{0}, \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U. \end{aligned} \quad (5.4)$$

donde el modificador de orden cero, ε_i^J , es la diferencia entre las funciones de costo de la planta y del modelo, el modificador de primer orden, $(\boldsymbol{\lambda}_i^J)^T$, representa la diferencia entre los gradientes del costo de la planta y del modelo, y el modificador de segundo orden, \mathbf{H}_i^J , representa la diferencia entre las matrices hessianas de la planta y del modelo. Los modificadores de las restricciones, ε_i^G , $(\boldsymbol{\lambda}_i^G)^T$ y \mathbf{H}_i^G , se definen de manera similar (ver Subsección 2.5.2).

5.3 Adaptación por Modificadores de Segundo Orden Mediante Interpolación Cuadrática

5.3.1 Cálculo de los Modificadores

En lugar de evaluar los modificadores por medio de su definición en las ecuaciones (2.37)-(2.40) y (2.46)-(2.47), proponemos calcularlos directamente por interpolación cuadrática utilizando un conjunto de puntos operativos pasados. Describimos el método solo para la función de costo, ya que los modificadores de las restricciones se pueden calcular de manera similar utilizando el mismo conjunto de puntos.

Al definir $\Delta \mathbf{u}^{(i)} = (\mathbf{u} - \mathbf{u}_i)$, la función objetivo adaptada (5.4) se puede expresar como:

$$J(\mathbf{u}, \boldsymbol{\theta}) + \varepsilon_i^J + (\boldsymbol{\lambda}_i^J)^T (\Delta \mathbf{u}^{(i)}) + \frac{1}{2} (\Delta \mathbf{u}^{(i)})^T \mathbf{H}_i^J (\Delta \mathbf{u}^{(i)}) \quad (5.5)$$

o de manera compacta como:

$$J(\mathbf{u}, \boldsymbol{\theta}) + \mathbf{m}(\Delta \mathbf{u}^{(i)}) \boldsymbol{\Lambda}_i^J, \quad (5.6)$$

donde $\mathbf{m}(\Delta \mathbf{u}^{(i)})$ es un vector fila formado con los elementos del polinomio de base natural para \mathbb{P}_n^{dim} (cuando $dim = 2$ y $n = n_u$) evaluados en $\Delta \mathbf{u}^{(i)}$, y $\boldsymbol{\Lambda}_i^J$ es un vector que contiene todos los modificadores.

A saber, la base natural de \mathbb{P}_n^2 se puede escribir de la siguiente manera [79]:

$$\bar{\varphi} = \left\{ 1 \quad x_1 \quad x_2 \quad \dots \quad x_n \quad \frac{x_1^2}{2} \quad x_1 x_2 \quad \dots \quad x_{n-1} x_n \quad \frac{x_n^2}{2} \right\} \quad (5.7)$$

Entonces, el vector fila $\mathbf{m}(\Delta \mathbf{u}^{(i)})$ se interpreta como:

$$\mathbf{m}(\Delta \mathbf{u}^{(i)}) = \left[1 \quad \Delta u_1^{(i)} \quad \Delta u_2^{(i)} \quad \dots \quad \Delta u_{n_u}^{(i)} \quad \frac{(\Delta u_1^{(i)})^2}{2} \quad \Delta u_1^{(i)} \Delta u_2^{(i)} \quad \dots \quad \Delta u_{n_u-1}^{(i)} \Delta u_{n_u}^{(i)} \quad \frac{(\Delta u_{n_u}^{(i)})^2}{2} \right] \in \mathbb{R}^{1 \times p_1}. \quad (5.8)$$

La construcción del vector $\boldsymbol{\Lambda}_i^J$ es:

$$\boldsymbol{\Lambda}_i^J = \left[\varepsilon_i^J \quad \text{vec}(\boldsymbol{\lambda}_i^J) \quad \text{vec}(\mathbf{H}_i^J) \right]^T, \quad (5.9)$$

$$\boldsymbol{\Lambda}_i^J = \left[\varepsilon_i^J \quad \lambda_{1,i}^J \quad \lambda_{2,i}^J \quad \dots \quad \lambda_{n_u,i}^J \quad h_{11,i}^J \quad h_{12,i}^J \quad \dots \quad h_{(n_u-1)n_u,i}^J \quad h_{n_u n_u,i}^J \right]^T \in \mathbb{R}^{p_1}. \quad (5.10)$$

En el caso de un polinomio cuadrático en \mathbb{R}^{n_u} se necesitan $p_1 = (n_u + 1)(n_u + 2)/2$ puntos para la construcción de una aproximación única.

Asuma que tenemos el conjunto $U = \{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_p\}$, con $p = p_1 - 1$. Eligiendo $\mathbf{u}_0 = \mathbf{u}_i$, este conjunto puede escribirse como: $U_i = \{\mathbf{u}_i, \mathbf{u}_1^{(i)}, \dots, \mathbf{u}_p^{(i)}\}$. Ahora consideremos las siguientes condiciones de interpolación para estos puntos:

$$\mathbf{m}(\Delta \mathbf{u}_j^{(i)}) \boldsymbol{\Lambda}_i^J = J_p(\mathbf{u}_j) - J(\mathbf{u}_j, \boldsymbol{\theta}) = \varepsilon_j^J, \quad \text{para } j = i, 1, \dots, p. \quad (5.11)$$

Recolectando estas p_1 condiciones en un arreglo matricial podemos formar el siguiente sistema:

$$\mathbf{M}(U_i) \boldsymbol{\Lambda}_i^J = \mathbf{E}_i^J, \quad (5.12)$$

donde:

$$\mathbf{M}(U_i) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & \Delta u_{1,1}^{(i)} & \Delta u_{2,1}^{(i)} & \cdots & \Delta u_{n_u,1}^{(i)} & \frac{(\Delta u_{1,1}^{(i)})^2}{2} & \Delta u_{1,1}^{(i)} \Delta u_{2,1}^{(i)} & \cdots & \Delta u_{n_u-1,1}^{(i)} \Delta u_{n_u,1}^{(i)} & \frac{(\Delta u_{n_u,1}^{(i)})^2}{2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & \Delta u_{1,p}^{(i)} & \Delta u_{2,p}^{(i)} & \cdots & \Delta u_{n_u,p}^{(i)} & \frac{(\Delta u_{1,p}^{(i)})^2}{2} & \Delta u_{1,p}^{(i)} \Delta u_{2,p}^{(i)} & \cdots & \Delta u_{n_u-1,p}^{(i)} \Delta u_{n_u,p}^{(i)} & \frac{(\Delta u_{n_u,p}^{(i)})^2}{2} \end{bmatrix}, \quad (5.13)$$

$$\mathbf{E}_i^J = [\varepsilon_i^J \quad \varepsilon_1^J \quad \varepsilon_2^J \quad \cdots \quad \varepsilon_p^J]^T \in \mathbb{R}^{p+1}. \quad (5.14)$$

Proposición: Los modificadores de hasta el segundo orden para la función objetivo se pueden calcular (asumiendo que $\mathbf{M}(U_i)$ no es singular) resolviendo:

$$\mathbf{\Lambda}_i^J = (\mathbf{M}(U_i))^{-1} \mathbf{E}_i^J. \quad (5.15)$$

De la misma manera se procede para calcular los modificadores de las restricciones.

Observación 8. En el caso que la dimensión del polinomio considerado sea de primer grado, $dim = 1$, se llega a la expresión de los modificadores de primer orden obtenidos en Marchetti [56].

Por ejemplo, siendo $dim=1$ and $n=n_u=2$, tendremos:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & \Delta u_{1,1}^{(i)} & \Delta u_{2,1}^{(i)} \\ 1 & \Delta u_{1,2}^{(i)} & \Delta u_{2,2}^{(i)} \end{bmatrix} \begin{bmatrix} \varepsilon_i^J \\ \lambda_{1,i}^J \\ \lambda_{2,i}^J \end{bmatrix} = \begin{bmatrix} \varepsilon_i^J \\ \varepsilon_1^J \\ \varepsilon_2^J \end{bmatrix}$$

Observación 9. Si consideramos un conjunto U_i que tenga más de la mínima cantidad de puntos para una interpolación cuadrática única, entonces se puede proceder mediante una regresión. Por ejemplo, podemos obtener los modificadores $\mathbf{\Lambda}_i^J$ que minimizan el error cuadrático:

$$\min_{\mathbf{\Lambda}_i^J} \sum_{j=0}^p (\mathbf{m}(\Delta \mathbf{u}_j^{(i)}) \mathbf{\Lambda}_i^J - \varepsilon_j^J)^2 \quad (5.16)$$

Observación 10. Si planteamos un polinomio de segundo orden, $dim = 2$, y los modificadores se obtienen mediante una regresión (5.16), entonces podemos vincular la propuesta con los resultados obtenidos en [58]. Sin embargo, en la publicación mencionada, los autores primero construyen una aproximación cuadrática de la función de la planta para aproximar el gradiente real por diferenciación de dicha aproximación cuadrática. Luego, los modificadores de primer orden se obtienen a través de la diferencia entre el gradiente de planta estimado y el gradiente del modelo obtenido mediante diferenciación analítica o diferencias finitas. Aquí, en cambio, construimos directamente una aproximación cuadrática de la función diferencia planta-modelo (condiciones 5.11) para extraer directamente los modificadores de primer y segundo orden, basándonos en valores pasados.

Observación 11. En caso de que no se disponga de un modelo, es decir, $J(\mathbf{u}, \boldsymbol{\theta}) = 0$ y $\mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) = \mathbf{0}$, la formulación representa una optimización en base a los polinomios cuadráticos interpolados (o bien obtenidos mediante regresión) de la función objetivo de la planta y de las restricciones para el conjunto de puntos U .

Observación 12. En la literatura de optimización matemática el cálculo del gradiente de una función mediante interpolación lineal es conocido como el gradiente simplex [80]. La obtención del hessiano de una función mediante interpolación cuadrática ha sido denominado también hessiano simplex [79]. Los modificadores aquí calculados podrían entonces entenderse como modificadores simplex.

5.3.2 Selección de Puntos

La selección del conjunto de puntos para llevar a cabo la interpolación es crucial para asegurar una aproximación de buena calidad de la función real de la planta. Para reflejar qué tan bien el conjunto considerado abarca la región donde se realizará la interpolación se puede utilizar un criterio denominado Λ -poisedness [79]. Es posible relacionar este concepto de conjunto equilibrado (Λ -poisedness) con el número de condición de la matriz $\mathbf{M}(U_i)$ si se utiliza como base del polinomio interpolador la base natural y el conjunto de puntos se encuentra centrado en un punto y escalado de modo que la norma máxima de los puntos del conjunto sea igual a uno [79]. En el mencionado libro, Conn et al. [79] analizan los algoritmos para construir, mantener o mejorar el valor de referencia Λ -poisedness del conjunto de interpolación o regresión. Sin embargo, estos algoritmos pueden resultar poco prácticos en un contexto de optimización de un proceso industrial donde cada evaluación significa un punto de operación real de la planta.

Por otro lado, es necesario definir la región donde el polinomio de interpolación puede considerarse como una aproximación precisa de la función estudiada. Esta región se conoce como región de confianza y comúnmente se define por:

$$B(\mathbf{u}_i; \Delta_i) = \{\mathbf{u} \in \mathbb{R}^{n_u} : \|\mathbf{u} - \mathbf{u}_i\| \leq \Delta_i\} \quad (5.17)$$

donde Δ_i se denomina radio de la región de confianza y se puede ajustar de acuerdo con los resultados obtenidos en la optimización anterior (basados por ejemplo en la diferencia entre el valor de la función objetivo estimada frente al valor real obtenido de la función objetivo).

Recientemente, en Gao et al. [58] se desarrolló un algoritmo para seleccionar puntos operativos pasados a modo de formar el mejor conjunto de puntos disponibles, y a su vez definir la región de confianza, para la construcción de una aproximación cuadrática en un contexto de optimización en tiempo real de procesos industriales. El algoritmo de selección de puntos y definición de región de confianza de Gao et al. [58] se puede resumir en los siguientes pasos:

Paso 1: Algoritmo de selección: basado en todos los datos recopilados y disponibles U , el algoritmo construye el conjunto de regresión seleccionando puntos cercanos (U_i^n) y distantes (U_i^d) para la i -ésima iteración del optimizador.

El conjunto de puntos cercanos viene dado por:

$$U_i^n = \{\mathbf{u} \in \mathbb{R}^{n_u} : \|\mathbf{u} - \mathbf{u}_i\| \leq \Delta_u\}, \quad (5.18)$$

donde Δ_u es un valor seleccionado para garantizar una aproximación cuadrática robusta en caso de la existencia de ruido en los datos.

Los puntos distantes se determinan mediante la siguiente optimización:

$$\begin{aligned} \min_{U_i^d} \quad & \sum_{\mathbf{u} \in U_i^d} \|\mathbf{u} - \mathbf{u}_i\| / \alpha(U_i^d) \\ \text{s.t.} \quad & \text{size}(U_i^d) \geq p; \quad U_i^d \subset U \setminus U_i^n, \end{aligned} \quad (5.19)$$

donde $\alpha(U_i^d)$ es el ángulo mínimo entre todos los vectores posibles definidos por $(\mathbf{u} - \mathbf{u}_i)$, p_1 es el número de puntos mínimos necesarios para una interpolación cuadrática única y $p = p_1 - 1$.

Entonces, el conjunto de regresión viene dado por $U_i = U_i^n \cup U_i^d$.

Paso 2: Región de confianza basada en covarianza: a diferencia de la definición de región de confianza habitual dada por la ecuación (5.17), en Gao et al. [58] se propone determinar una región

restringida utilizando la matriz de covarianza del conjunto de regresión seleccionado:

$$\beta_i = \left\{ \mathbf{u} \in \mathbb{R}^{n_u} : (\mathbf{u} - \mathbf{u}_i)^T \mathbf{C}^{-1} (\mathbf{u} - \mathbf{u}_i) \leq \gamma^2 \right\}, \quad (5.20)$$

donde $\mathbf{C} = \text{cov}(U_i)$, y γ es un parámetro de escalado. Como en los métodos de la región de confianza, el parámetro γ también se puede ajustar a medida que avanzan las iteraciones y de acuerdo con la precisión de los valores predichos.

En esta tesis seguiremos tras los pasos de Gao et al. [58] para la selección de puntos y definición de la región de confianza. En caso de que optemos por hacer una interpolación, seleccionando el número mínimo de puntos necesarios para una aproximación cuadrática única, proponemos aquí modificar el algoritmo de selección rechazando los puntos cercanos, excepto \mathbf{u}_i , y usando solo los puntos externos necesarios.

5.3.3 Algoritmo para Adaptación por Modificadores de Segundo Orden Mediante Interpolación Cuadrática

En esta sección, se presenta una descripción general del algoritmo para el esquema de optimización en tiempo real con modificadores de segundo orden. Supongamos un punto inicial factible \mathbf{u}_0 , y los puntos operativos en $\mathbf{u}_0 + h\mathbf{e}_n$, donde h es el tamaño de paso inicial y $\mathbf{e}_n \in \mathbb{R}^{n_u}$, ($n = 1, \dots, n_u$), son vectores unitarios ortogonales, para así obtener una primera estimación de los modificadores de primer orden. Luego de las evaluaciones de planta iniciales y hasta que tengamos suficientes puntos para una aproximación cuadrática única, aplicamos algún algoritmo de adaptación por modificadores de primer orden, por ejemplo, podemos usar el método derivado de la observación 8. Posteriormente estaremos en condiciones de generar aproximaciones cuadráticas.

El algoritmo propuesto para adaptación por modificadores de segundo orden comprende los siguientes pasos:

- (I) Configuración de los parámetros iniciales. Defina los parámetros de la región de confianza (normalmente un valor de escalado γ , un parámetro de reducción r y un mínimo γ_{\min}). Elija un punto operativo de partida factible \mathbf{u}_0 . Evalúe la planta en los puntos $\mathbf{u}_0 + h\mathbf{e}_n$, donde h es el parámetro de paso inicial a definir y $\mathbf{e}_n \in \mathbb{R}^{n_u}$, ($n = 1, \dots, n_u$), son vectores unitarios ortogonales. Estime los modificadores de primer orden.
- (II) Inicialización mediante el método de adaptación de modificadores de primer orden. De ser necesario tunear los parámetros adicionales del método de adaptación de modificadores seleccionado. Aplicar el método de adaptación de modificadores hasta que p_1 puntos operativos distintos estén disponibles. Hallar:

$$\mathbf{u}^* = \underset{\mathbf{u}}{\text{arg mín}} \{J_p(\mathbf{u}_j)\}, \text{ para } j = 0, 1, \dots, p.$$

Identifique el punto de operación correspondiente \mathbf{u}^* , obtenido en dicha búsqueda, como el paso de mejora inicial, es decir $\mathbf{u}_i = \mathbf{u}_1 = \mathbf{u}^*$.

- (III) Ejecute el algoritmo de selección de datos para generar el conjunto U_i . Defina la región de confianza β_i .
- (IV) Calcule los modificadores resolviendo el sistema (5.15) o (5.16). Ejecute el problema de optimización adaptado (5.4) agregando la restricción: $\mathbf{u} \in \beta_i$. Obtenga la solución óptima \mathbf{u}_i^* .
- (V) Ejecute el siguiente paso crítico,

Paso Crítico: Si $\|\mathbf{u}_i^* - \mathbf{u}_i\| \leq \Delta_u$ y existe al menos un punto $\mathbf{u}_j \in U_i$ tal que $\|\mathbf{u}_j - \mathbf{u}_i\| > 2\Delta_u$, entonces definir $\hat{\mathbf{u}}_i = (\mathbf{u}_j + \mathbf{u}_i)/2$.

De lo contrario definir $\hat{\mathbf{u}}_i = \mathbf{u}_i^*$.

(VI) Evalúe el proceso en el punto operativo $\hat{\mathbf{u}}_i$ y determine:

a) Movimiento Exitoso: Si $J_p(\hat{\mathbf{u}}_i) < J_p(\mathbf{u}_i)$, (y no hay violación en las restricciones) este es un mejor punto operativo. Luego actualice $\mathbf{u}_{i+1} = \hat{\mathbf{u}}_i$. Incremente el índice i por uno y vuelva al paso 3.

b) Movimiento fallido. Si $J_p(\hat{\mathbf{u}}_i) \geq J_p(\mathbf{u}_i)$, este no es un mejor punto de operación, pero puede considerarse como un paso exploratorio. No actualizar \mathbf{u}_i . Reducir el tamaño de la región de confianza γ por el factor r seleccionado, y volver al paso 3 (esto mantendrá centrada la interpolación en \mathbf{u}_i pero tendremos un nuevo punto a considerar al buscar formar U_i).

Tenga en cuenta que aquí el subíndice i hace referencia a un paso de mejora. Este índice cambiará solo cuando se realice un movimiento exitoso, sin embargo, tanto los movimientos exitosos como los fallidos son operaciones de planta que deben tenerse en cuenta para evaluar el desempeño del esquema propuesto.

5.3.4 Aspectos de la implementación

El algoritmo aquí sugerido, para utilizar los modificadores propuestos en un esquema de RTO, se basa en los algoritmos existentes en la literatura de DFO [79] como así también en el algoritmo utilizado por Gao et al. [58]. Las sucesivas simulaciones del caso de estudio, que se presentan en la próxima sección, muestran en general la idoneidad para converger a la cercanía del óptimo real de la planta para escenarios realistas que incluyen restricciones y ruido en las mediciones. La calidad de la aproximación cuadrática, y en consecuencia la calidad de los modificadores, dependerá de la adecuada distribución de los puntos U_i en el espacio. El paso crítico, común en los métodos DFO [79] y utilizado también en [58] y en el presente trabajo, busca mejorar el acondicionamiento del conjunto de puntos al reducir la distancia entre los puntos del conjunto interpolador. En la imagen de la izquierda de la Figura 5.1 se ilustra el efecto de aplicar el paso crítico. Como puede observarse el punto sugerido $\hat{\mathbf{u}}_i = (\mathbf{u}_j + \mathbf{u}_i)/2$ mejorará el acondicionamiento al proporcionar, para la próxima selección del conjunto U_i , el punto $\hat{\mathbf{u}}_i$. Sin embargo, que todos los puntos pertenecientes a U_i estén dentro de la esfera determinada por $2\Delta_u$ no nos asegura que exista una buena distribución espacial de los mismos.

En los métodos DFO [79] la buena disposición de puntos se logra realizando movimientos que buscan mejorar el acondicionamiento del conjunto de puntos, ver por ejemplo los algoritmos presentados en el Capítulo 6, en Conn et al. [79]. Siguiendo el ejemplo anterior, en la imagen derecha de la Figura 5.1 se muestra un posible punto $\hat{\mathbf{u}}_i$ que proporcionará una mejora en el conjunto U_i para la próxima interpolación cuadrática. Sin embargo, aplicar los algoritmos de mejora de conjunto de puntos en un contexto de procesos industriales implica realizar modificaciones adicionales en el punto de operación de la planta. A su vez, es necesario ser cautos al aplicar directamente dichos algoritmos en presencia de restricciones, ya que estos movimientos podrían llevarnos a violar momentáneamente las restricciones existentes.

En Gao et al. [81] los autores sostienen que la mejora del conjunto de puntos en su algoritmo es consecuencia de los movimientos fallidos o exploratorios. Estos movimientos exploratorios, al estar basados en una aproximación de baja calidad, pueden también llevarnos a violar momentáneamente las restricciones. La ventaja que alegan Gao et al. [81] es que no se requiere realizar en cada iteración del RTO movimientos adicionales (puntos operativos en la aplicación industrial). El algoritmo aquí planteado sigue esta idea de Gao et al. [81]. Sin embargo, consideramos que es necesario ahondar en

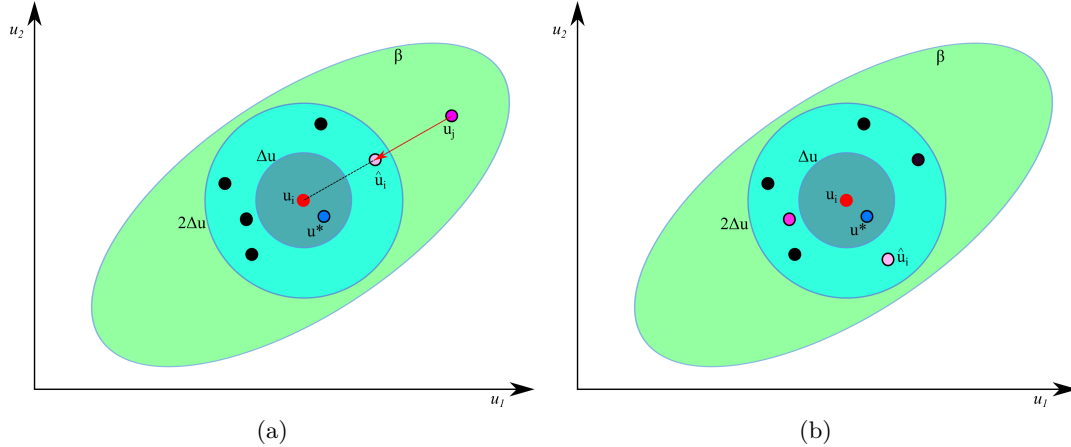


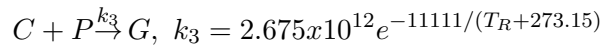
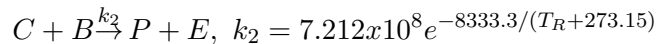
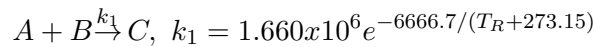
Figura 5.1: En la imagen de la izquierda se ilustra el paso crítico y en la derecha el paso de mejora del conjunto de puntos. El área verde es la región de confianza y las áreas celeste y azul las correspondientes a $2\Delta_u$ y Δ_u respectivamente. El punto rojo \mathbf{u}_i es el último punto de mejora, los puntos negros sólidos y el magenta pertenecen al actual conjunto U_i . El punto azul \mathbf{u}_i^* es el recomendado según la optimización por adaptación de modificares (paso IV del algoritmo presentado) y se encuentra dentro de la esfera Δ_u . El punto rosa $\hat{\mathbf{u}}_i$ es el nuevo punto operativo recomendado según se aplica el paso crítico (imagen de la izquierda) o el paso de mejora del conjunto de puntos (imagen de la derecha).

casos de estudios y, de ser necesario, incorporar explícitamente un paso de mejora del conjunto de puntos como los utilizados en la bibliografía de DFO [79].

5.4 Resultados de Simulación

Usaremos el problema de referencia del reactor de Williams-Otto [67] para evaluar el desempeño del esquema de RTO propuesto. Evaluaremos un primer caso donde se formula como un problema sin restricciones y una segunda formulación donde se convierte en un problema restringido (e inadecuado para algunos esquemas de RTO existentes).

En el reactor verdadero (realidad simulada) ocurren las siguientes reacciones:



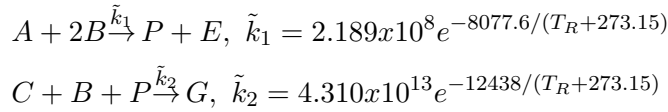
donde k_1, k_2, k_3 son las constantes de velocidad de reacción. Los productos deseados son P y E , C es un producto intermedio, y G es un subproducto indeseado. Los reactivos A y B se alimentan con caudales máxicos F_A y F_B , la masa retenida en el reactor es 2105 Kg y la temperatura de operación isotérmica es T_R .

Para este problema, la función objetivo se expresa como una ganancia económica a maximizar:

$$\phi(u, y) = 1143.38X_P F + 25.92X_E F - 76.23F_A - 114.34F_B$$

donde X_E y X_P son las concentraciones de los productos E y P . El caudal F está dado por $F = F_A + F_B$.

El error planta-modelo se introduce asumiendo que las reacciones anteriores son desconocidas, y solo se ha modelado una aproximación mediante dos reacciones:



En [82] y en el apéndice se encuentran las ecuaciones de la planta y el modelo utilizados.

5.4.1 Problema Sin Restricciones para el Reactor de Williams-Otto

Para las simulaciones sin restricciones, consideramos la variable de entrada $F_A = 1.8275$ kg/s fija. Las variables de decisión son entonces F_B y T_R , es decir, $\mathbf{u} = [F_B \quad T_R]^T$. Se introduce el mismo nivel de ruido usado en [56] y [58] para la función objetivo medida (distribución normal gaussiana con desviación estándar de 0,5).

El problema de optimización es el siguiente:

$$\begin{aligned} \underset{F_B, T_R}{\text{máx}} \quad & 1143.38X_P F + 25.92X_E F - 76.23F_A - 114.34F_B \\ \text{s.t.} \quad & \text{ecuaciones del modelo en estado estacionario,} \\ & F_B \in [3, 6], \quad T_R \in [70, 100]. \end{aligned} \tag{5.21}$$

Para la inicialización en el paso (II) del algoritmo presentado en la Subsección 5.3.3 se utiliza el método de adaptación por modificadores de primer orden mediante interpolación lineal. Se utilizará también como restricción adicional la región de confianza dada por la ecuación (5.20). Los parámetros para el método de selección de datos y para la definición de la región de confianza se ajustan con $\gamma^2 = 9$, un factor de reducción $r = 5/6$ y un mínimo $\gamma_{\min} = 0.05$. El valor de γ está relacionado con el tamaño del espacio de búsqueda definido por la ecuación (5.20). El radio del círculo se fija en $\Delta_u = 0.06$ y establece el límite entre los puntos cercanos y los puntos lejanos para la construcción de la aproximación cuadrática (es también un valor definido por el usuario [58]).

Se considerarán simulaciones para el caso de interpolación en el que los modificadores se calculan mediante la ecuación (5.15) y el caso de regresión en el que los modificadores se determinan resolviendo el Problema (5.16) y se compararán los resultados obtenidos por ambos métodos. Las simulaciones se llevan a cabo para 60 evaluaciones de planta con puntos de partida aleatorios factibles.

El gráfico de la izquierda de la Figura 5.2 muestra los valores de la ganancia económica obtenida en función de los puntos operativos implementados en el proceso (movimientos exitosos y movimientos fallidos del esquema de RTO propuesto) para el caso donde los modificadores se obtienen mediante el uso de interpolación. Se puede observar cómo la operación tiende a converger a las cercanías del óptimo de la planta. Se observan también movimientos erráticos (a causa de una mala aproximación o a causa de los pasos críticos) y posteriormente correcciones. La línea discontinua, en la gráfica de la izquierda, marca el beneficio máximo. La gráfica de la derecha de la Figura 5.2 muestra los puntos iniciales y finales obtenidos después de 60 evaluaciones operativas de planta para los 50 puntos iniciales aleatorios. La simulación se repite utilizando los modificadores calculados mediante regresión y los resultados se muestran en la Figura 5.3.

En ambos enfoques, el algoritmo propuesto es eficaz para alcanzar las cercanías del óptimo de la planta.

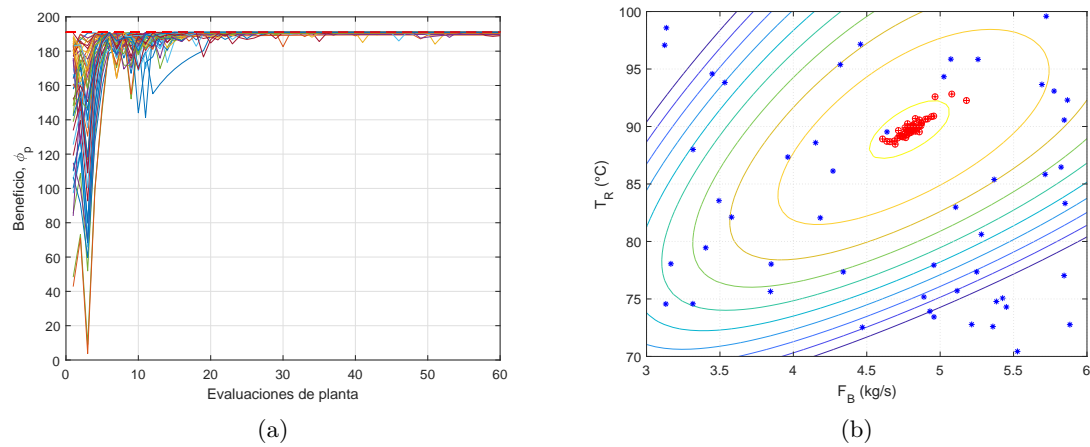


Figura 5.2: Escenario de interpolación, simulado para 50 puntos iniciales aleatorios. Gráfica izquierda: evolución del beneficio en función de las 60 evaluaciones de planta para los puntos iniciales aleatorios. Línea discontinua roja: valor óptimo de la planta. Gráfica derecha: Puntos de inicio (azul *) y puntos de finalización (rojo \oplus) después de 60 puntos operativos aplicados al proceso. Líneas de contorno: beneficio de la planta.

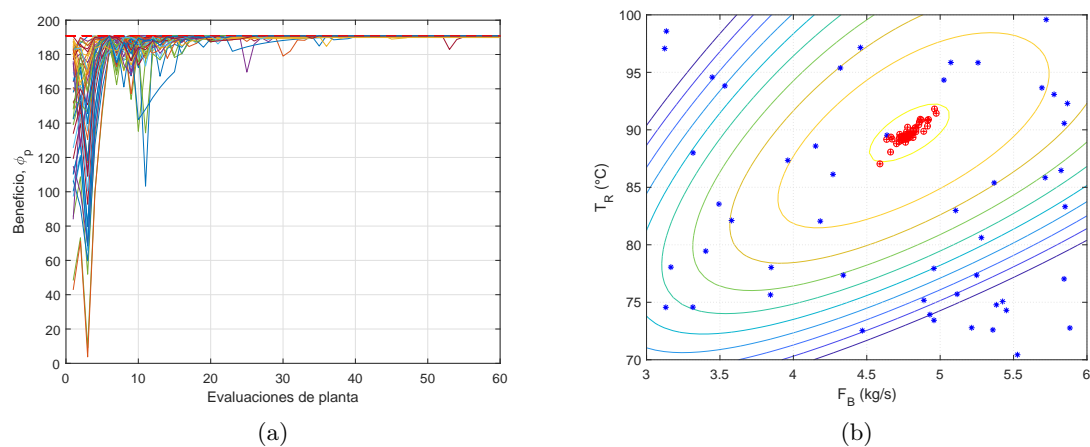


Figura 5.3: La simulación se repite para el esquema propuesto utilizando modificadores obtenidos mediante regresión cuadrática.

5.4.2 Problema Con Restricciones para el Reactor de Williams-Otto

Para esta simulación, agregaremos una restricción al proceso planteado anteriormente, a saber, $g = X_G - 0.08 \leq 0$. La función objetivo es la misma presentada para el caso general. El problema de optimización se puede escribir como:

$$\begin{aligned}
 & \underset{F_B, T_R}{\text{máx}} && 1143.38X_P F + 25.92X_E F - 76.23F_A - 114.34F_B \\
 & \text{s.t.} && \text{ecuaciones del modelo en estado estacionario,} \\
 & && X_G - 0.08 \leq 0, \\
 & && F_B \in [3, 6], \quad T_R \in [70, 100].
 \end{aligned} \tag{5.22}$$

El punto de operación óptimo de la planta es $\hat{\mathbf{u}} = [4, 97 \ 84, 3]^T$. En dicho punto, la restricción sobre X_G está activa, es decir, $X_G = 0.08$, y la ganancia económica alcanza un valor de $\phi_p = 178.63$. De nuevo, consideraremos la presencia de ruido gaussiano aleatorio en la función objetivo con: $\sigma_\phi = 0.5$. Para la determinación de la región de confianza, establecemos $\gamma^2 = 9$, un factor de reducción $r = 10/11$ y un mínimo $\gamma_{\text{mín}} = 0,05$. Para la selección de puntos, se fija el parámetro $\Delta_u = 0.08$. A partir de puntos de operación factibles iniciales tomados en forma aleatoria, realizamos 60 iteraciones RTO.

El gráfico de la izquierda de la Figura 5.4 ilustra la evolución del valor objetivo y de la variable restringida para los puntos sucesivos del esquema de RTO mediante adaptación por modificadores de segundo orden obtenidos por interpolación cuadrática. Se observa en el gráfico izquierdo superior de la Figura 5.4 que el algoritmo se aproxima al valor óptimo rápidamente. A medida que se logra una mejor representación de la aproximación cuadrática de la diferencia planta-modelo, se va convergiendo a un punto cercano al óptimo real. Se observan también puntos que están violando la restricción $X_G = 0.08$, es decir, no hay garantía de que se seleccionen puntos factibles en algunas evaluaciones de planta para el esquema de RTO propuesto. Esto es consecuencia de un movimiento realizado con una mala aproximación cuadrática. Los puntos de inicio y finalización se observan en el gráfico de la derecha en la Figura 5.4. En la Figura 5.5 se presentan los resultados obtenidos mediante modificadores de segundo orden por regresión cuadrática. Se observa que ambas técnicas de cálculo de modificadores nos permiten acercarnos y mantenernos cerca del óptimo real de la planta.

5.4.3 Problema Alternativo con Restricciones para el Reactor Williams-Otto

En este caso utilizaremos una formulación modificada del problema anterior con los cambios presentados en [78]: el caudal F_A es ahora una variable de decisión, es decir, $\mathbf{u} = [F_A \ F_B \ T_R]^T$. Los límites inferiores y superiores de las variables de decisión (o de entrada) son: $\mathbf{u}^L = [3 \ 6 \ 80]^T$, $\mathbf{u}^U = [4.5 \ 11 \ 105]^T$. La restricción es $g = X_G - 0.08 \leq 0$ y la función objetivo es la misma presentada para el caso general. El problema de optimización puede entonces escribirse de la siguiente manera:

$$\begin{aligned}
 & \underset{F_A, F_B, T_R}{\text{máx}} && 1143.38X_P F + 25.92X_E F - 76.23F_A - 114.34F_B \\
 & \text{s.t.} && \text{ecuaciones del modelo en estado estacionario,} \\
 & && X_G - 0.08 \leq 0, \\
 & && F_A \in [3, 4.5], \quad F_B \in [6, 1], \quad T_R \in [80, 105].
 \end{aligned} \tag{5.23}$$

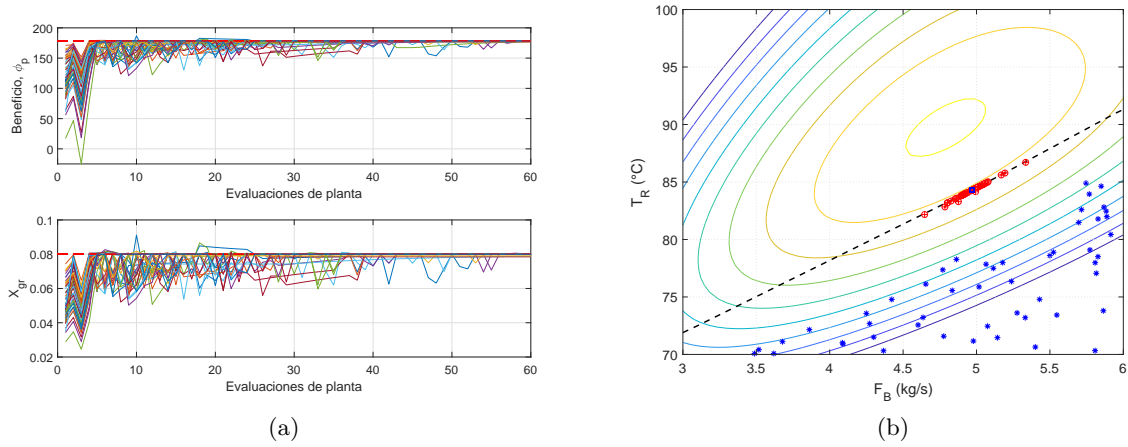


Figura 5.4: Gráfico de la izquierda: Superior: Beneficio económico obtenido en los primeros 60 puntos operativos del método RTO con modelo adaptado por modificadores de segundo orden mediante interpolación. Inferior: Evolución de la variable restringida para las evaluaciones de planta. Gráfico de la derecha puntos iniciales y finales. Esquema de interpolación. Simulado para 50 puntos de partida aleatorios.

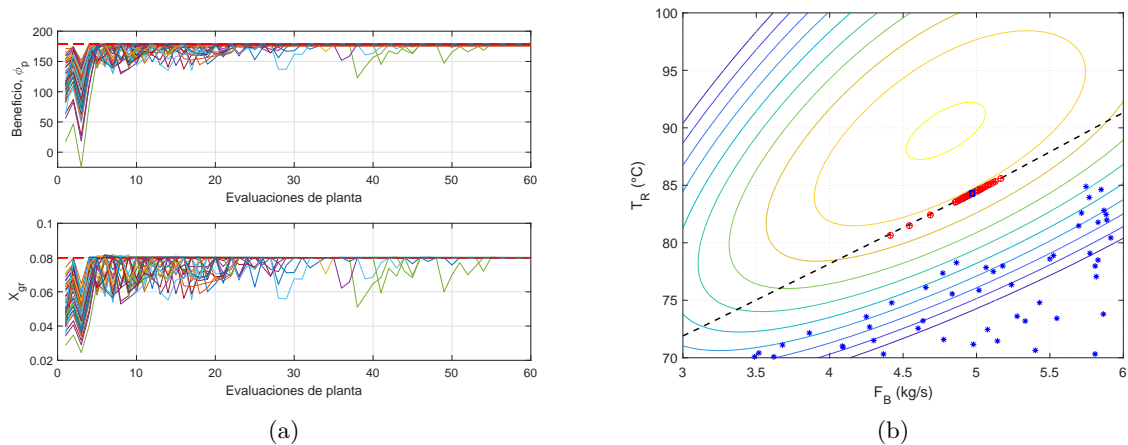


Figura 5.5: La simulación se repite para método RTO con modelo adaptado por modificadores de segundo orden mediante regresión. Simulado para 50 puntos de partida aleatorios.

En este caso, el punto de operación óptimo de la planta es $\mathbf{u} = [3.887 \ 9.369 \ 91.2]^T$. Para este punto la restricción sobre X_G está activa y la ganancia alcanza $\phi_p = 210,33$.

Para este problema de optimización, el método estándar de adaptación por modificadores de primer orden no converge. Se muestra en [78] que el modelo que se posee no es adecuado para el algoritmo mencionado. Esto se debe a que el Hessiano reducido del Lagrangiano del problema de optimización en el punto \mathbf{u}_p^* no es positivo semi-definido (ver Proposición 1). Para lidiar con esto, en [78] los autores propusieron un esquema de optimización que considera una situación ideal donde el gradiente de costo y el gradiente de la restricción de la planta y del modelo están disponibles en cada iteración de RTO. El efecto del ruido no se consideró allí y no se considerará aquí para una primera comparación de los métodos.

Mediante este caso de estudio se mostrará cómo el método propuesto de adaptación de modificadores de segundo orden mediante aproximación cuadrática podría responder a este tipo de inconvenientes. Para este caso de estudio, necesitamos al menos $p1 = 10$ puntos para una interpolación única. Inicializamos el algoritmo como se presentó anteriormente, es decir, mediante una interpolación lineal dentro de una región de confianza.

Partimos de un punto factible dado $\mathbf{u}_0 = [3.6 \ 10 \ 85]^T$ que será utilizado para las simulaciones del método de adaptación por modificadores estándar y para el método de adaptación por modificadores de segundo orden por aproximación cuadrática propuesto en esta tesis. La evolución de las iteraciones, tanto del funcional objetivo como del valor de la restricción del proceso, se muestran en el gráfico de la izquierda de la Figura 5.6 para el método de adaptación por modificadores estándar. En el gráfico de la derecha de la Figura 5.6 también se representa la superficie que corresponde a la restricción activa $X_G = 0.08$ y en ella se puede observar las líneas de contorno de la ganancia económica. Notar que la restricción sobre X_G se viola por encima de la superficie mostrada.

En el gráfico de la izquierda en la Figura 5.6 se observa un comportamiento inestable que no converge hacia el óptimo y queda violando la restricción.

A continuación, se simula el método propuesto. Desde el punto inicial factible dado, procederemos usando el esquema presentado en esta tesis en su versión de primer orden (observación 8) hasta alcanzar la cantidad de puntos necesarios que nos permitan la construcción de la aproximación cuadrática.

Para definir la región de confianza usamos: $\gamma^2 = 9$, un factor de reducción de $r = 10/11$, y un valor mínimo de $\gamma_{\min} = 0.05$. El radio que establece el límite de puntos cercanos es fijado en $\Delta_u = 0.08$.

Como se puede observar en la imagen de la izquierda de la Figura 5.7, el nuevo esquema RTO con adaptación por modificadores se aproxima rápidamente al valor óptimo y luego de movimientos exploratorios puede mantenerse en él. Sin embargo, como ya se notó en el ejemplo anterior, no hay garantía de factibilidad en algunas de las iteraciones del optimizador.

La razón de dicho inconveniente es la construcción de la diferencia planta-modelo por interpolación tanto de la función costo como también de la restricción. Nuestro conjunto de puntos para realizar la interpolación/regresión cuadrática es el mejor disponible según los datos históricos, pero esto no significa que sea lo suficientemente bueno para representar con precisión la diferencia planta-modelo, y por ende los modificadores. Por lo tanto, el siguiente paso puede llevarnos a violar las restricciones. Como se presentó en la publicación [81] estos movimientos exploratorios (que en este caso pueden violar las restricciones) mejorarán el conjunto de interpolación y nos conducirán a un conjunto más equilibrado.

A pesar de esto, existen puntos positivos a destacar. El método estándar no converge y queda violando la restricción, mientras que el método presentado por Marchetti et al. [78] para abordar este problema, necesita 150 iteraciones para converger. En cambio, utilizando el método presentado en esta

tesis estamos en aproximadamente 30 evaluaciones de planta convergiendo al óptimo en un escenario más realista donde no se necesita disponer de los gradientes ni hessianos analíticos de la planta (la comparación se hace para el mismo punto de partida ($u_0 = [3.6 \ 10 \ 85]^T$)).

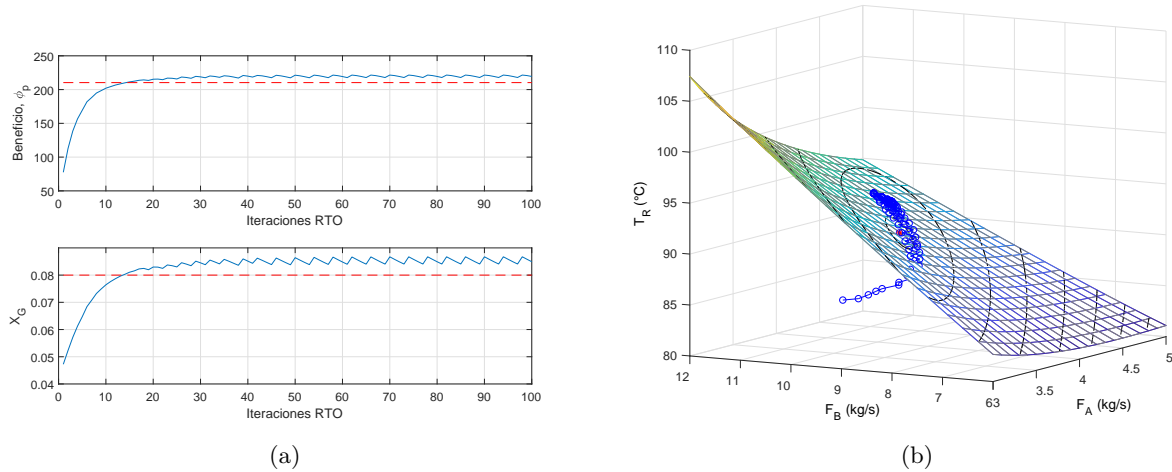


Figura 5.6: Método de adaptación por modificadores estándar. A la izquierda, imagen superior, se representa la evolución de la ganancia económica en función de las evaluaciones de planta. En líneas de puntos rojos se indica el valor óptimo. En la imagen inferior se representa el valor de la variable de salida restringida X_G y en líneas de puntos rojos su valor límite superior. La imagen de la derecha representa la trayectoria de los puntos operativos de la planta como así también la superficie que corresponde a la restricción activa $X_G = 0.08$.

Finalmente simularemos el ejemplo incorporando ruido para la función objetivo medida (distribución normal gaussiana con desviación estándar de 0,5). En la Figura 5.8 se muestran los resultados utilizando el esquema de interpolación para 50 puntos de inicio factibles aleatorios y para 120 evaluaciones de planta. La imagen de la izquierda de la Figura 5.8 ilustra los resultados obtenidos de la función objetivo (gráfico superior) como así también los valores alcanzados por la variable restringida (gráfico inferior). En la Figura 5.9 se repiten las simulaciones para el esquema propuesto utilizando modificadores de segundo orden obtenidos mediante regresión cuadrática.

Los resultados muestran la capacidad del método para converger hacia la cercanía del óptimo real de la planta.

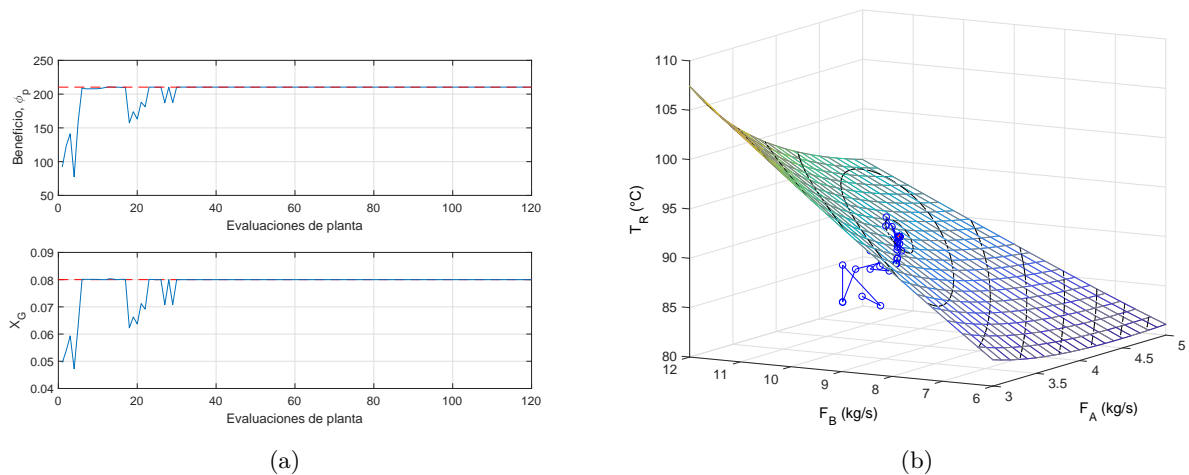


Figura 5.7: Método de adaptación por modificadores vía aproximación cuadrática. En la imagen de la izquierda se observa la evolución de la ganancia económica (imagen superior) y de la variable de salida restringida X_G (imagen inferior) en función de las evaluaciones de planta. La imagen de la derecha muestra la trayectoria operativa según el método propuesto.

5.5 Discusiones y Conclusiones del Capítulo

En este capítulo se ha propuesto un procedimiento para estimar modificadores de segundo orden que pueden utilizarse en el método de optimización en tiempo real de adaptación por modificadores.

Se propone incorporar el procedimiento de estimación de modificadores de segundo orden a un algoritmo que, en conjunto con técnicas existentes de selección de datos operativos pasados y de región de confianza, nos permite converger a un punto operativo cercano al óptimo real de la planta.

En cuanto al cálculo de los modificadores es de destacar que es posible estimarlos para un escenario realista, es decir, sin el conocimiento de los gradientes y la matriz hessiana de la planta real.

Los resultados de la simulación del reactor Williams-Otto muestran que el algoritmo propuesto se acerca al óptimo, procede con iteraciones exploratorias que mejoran el conjunto de puntos y finalmente converge a un entorno cercano al óptimo real de la planta. La convergencia es posible aún en caso de que no se disponga de un modelo adecuado según la Proposición 1.

Se sugiere avanzar en casos de estudios con procesos que presenten más grados de libertad y restricciones para encontrar las potenciales limitaciones y buscar mejoras de la estrategia presentada. A su vez es de interés mejorar el algoritmo a fin de evitar o minimizar las posibles violaciones de las restricciones.

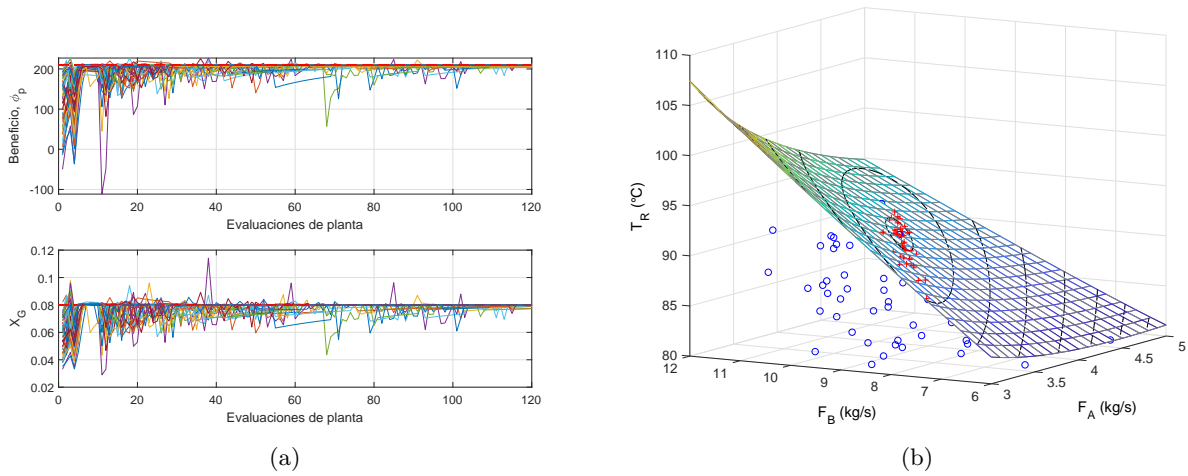


Figura 5.8: Método de adaptación por modificadores vía interpolación cuadrática. En la imagen de la izquierda se observa la evolución de la ganancia económica (imagen superior) y de la variable de salida restringida X_G (imagen inferior) en función de las evaluaciones de planta. La imagen de la derecha muestra los puntos iniciales (azul *) y puntos de finalización (rojo \oplus) después de 120 puntos operativos aplicados al proceso

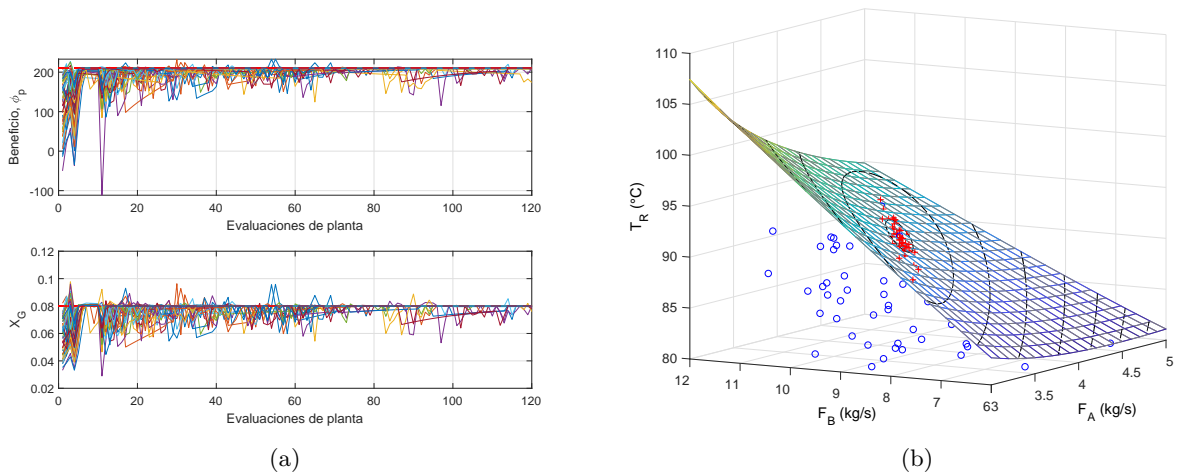


Figura 5.9: Método de adaptación por modificadores vía regresión cuadrática. La imagen de la izquierda ilustra la evolución de la ganancia económica (imagen superior) y de los valores alcanzados por la variable de salida restringida X_G (imagen inferior) en función de las evaluaciones de planta. La imagen de la derecha muestra los puntos iniciales (azul *) y puntos de finalización (rojo \oplus) después de 120 puntos operativos aplicados al proceso.

Capítulo 6

Conclusiones Generales y Trabajos Futuros

6.1 Conclusiones Generales

A lo largo de la tesis se propusieron diversos métodos para la operación óptima de plantas industriales en base a criterios económicos de estado estacionario. El eje central de la tesis es la selección de la estructura de control que garantice una operación económicamente óptima a la vez que se satisfacen las restricciones operativas ante el efecto de las perturbaciones externas.

En el Capítulo 3 se estudiaron y se vincularon los métodos de los márgenes de seguridad en estado estacionario con los métodos de control auto optimizantes para la selección de la estructura de control. Al formular el problema de los márgenes de seguridad minimizando el costo promedio y dado que a fines resolutivos minimizar el costo promedio y la pérdida promedio no presentan diferencias, se estableció el vínculo entre ambos los métodos. La búsqueda de la estructura de control óptima según el método de los márgenes de seguridad y costo promedio se formuló mediante dos problemas de optimización matemática. En una primera formulación se consideran estructuras de control clásicas y la búsqueda se resuelve mediante un Problema MINLP (3.4). En una segunda formulación se consideran posibles combinaciones lineales de variables a controlar y la búsqueda de la estructura de control óptima se resuelve mediante un Problema NLP (3.10). Esta formulación NLP representa un problema de control auto optimizante con combinaciones lineales del tipo global al buscar una solución óptima que contempla todos los valores de las perturbaciones. Debido a la posible dificultad de resolver problemas MINLP, según el tamaño y complejidad de la planta estudiada, en la Sección 3.4.2 se formularon aproximaciones MIQP que mostraron idoneidad para proporcionar resultados en un tiempo de cómputo reducido.

Debido al efecto de las perturbaciones es probable que las restricciones activas y el punto operativo óptimo cambien ante distintos valores del vector de perturbaciones. En consecuencia, es posible que la estructura de control fija con valores de consigna fijos nos proporcionen una operación subóptima de la planta. Esta operación de planta con puntos operativos óptimos cambiantes se suelen abordar mediante un sistema de control que involucre un control supervisor. Dicho control supervisor asignará cambios a los valores de consigna, o directamente cambios en los lazos de control, según se detecten cambios en las condiciones operativas óptimas.

En el Capítulo 4 se estudiaron y propusieron diversas estrategias para incorporar un control supervisor. Un paso importante para la aplicación de los métodos propuestos es distinguir entre perturbaciones medibles o estimables y perturbaciones no medibles. Esta división nos permite por

ejemplo analizar las potenciales mejoras económicas que se pueden obtener al incluir, en el nivel de control supervisor, un optimizador en tiempo real que mediante mediciones de las condiciones actuales del proceso recalcula el punto operativo óptimo. Es necesario destacar que dicha búsqueda del optimizador en tiempo real se encontrará restringida a la estructura de control regulador instalada dado que el resultado de la optimización se transmite como valores de consigna a las variables controladas o a las variables manipuladas fijas. Es por esto que se hace necesario plantear el problema de selección de estructura de control en conjunto con la existencia del nivel superior de control en tiempo real. En el Capítulo 4 se formuló la selección de la estructura de control en base a criterios económicos de estado estacionario que contempla la posibilidad de un sistema de control supervisor operando en base a las perturbaciones medibles. Para la búsqueda de la estructura de control óptima correspondiente a dicho escenario se formuló un Problema MINLP (4.1).

Una alternativa a un optimizador en tiempo real es un sistema supervisor que trabaje identificando diversas regiones operativas y produciendo cambios en la estructura de control según el estado actual de la planta. Pensando en la operación óptima de la planta se sugiere que las regiones se caractericen según que restricciones del problema de optimización (2.5) se encuentran activas ante un particular escenario de perturbaciones. Una vez identificadas las regiones de restricciones activas se pueden entonces formular problemas de optimización que permitan cambios en las estructuras de control según se opere en una u otra región. Para la búsqueda de las estructuras de control óptimas correspondientes a esta estrategia se formularon los Problemas MINLP (4.4) y (4.6). Las diversas estrategias de selección de estructura de control se pueden generalizar con la formulación (4.8) e incluyendo restricciones adicionales.

Finalmente en el Capítulo 5 se abordó el posible escenario en el que el modelo de planta disponible es deficiente y difiere de la planta real. En tal caso, la optimización de la operación en base al modelo probablemente no coincida con el óptimo real de la planta. Para hacer frente a este inconveniente los métodos de optimización en tiempo real proceden adaptando el modelo según la información real de la planta obtenida con los instrumentos de medición. Posteriormente ejecutan el problema de optimización con el modelo adaptado y obtienen el nuevo punto operativo. Los métodos clásicos de optimización en tiempo real con adaptación por modificadores han mostrado su capacidad de converger hacia el óptimo real de la planta en una gran cantidad de escenarios. Sin embargo, para que esto suceda es necesario que el modelo presente cierta adecuación para su uso (ver Proposición 1). Si no es posible asegurar la adecuación del modelo para una adaptación por modificadores de primer orden, es sugerido entonces proceder con una adaptación por modificadores de segundo orden. En esta tesis se ha trabajado y propuesto un método para el cálculo de los modificadores de segundo orden mediante interpolación cuadrática en base a la utilización de la información de planta de puntos operativos pasados. La utilización de los modificadores se sintetizó mediante el algoritmo para la adaptación del modelo por modificadores de segundo orden presentado en la Sección 5.3.3.

La distintas simulaciones de casos de estudio han mostrado a lo largo de la tesis la capacidad de los métodos propuestos para conducir a los ingenieros hacia una operación de planta económicamente óptima en base a la correcta selección de la estructura de control.

De acuerdo a lo presentado en esta tesis, se proporcionan las siguientes recomendaciones para la selección de estructuras de control en base a criterios económicos de estado estacionario:

- Se recomienda minimizar el costo promedio en lugar del costo nominal en los métodos basados en los márgenes de seguridad. Se ha demostrado que la minimización del costo nominal puede dar lugar a estructuras de control con bajo rendimiento económico para valores de perturbación diferentes a los nominales.
- Se recomienda minimizar el costo promedio en lugar de la pérdida económica promedio en los métodos de los márgenes de seguridad y en los métodos de control auto optimizante para la

selección de estructuras de control. Evaluar la pérdida con respecto al funcionamiento óptimo es innecesario en vista de la Observación 1. Una vez que se determina la estructura de control óptima, se puede evaluar la pérdida económica asociada.

- Para garantizar la viabilidad para todos los escenarios de perturbación, se recomienda calcular valores de consigna óptimos para todas las variables controladas y para todas las variables de entrada que no se seleccionan como variables manipuladas. Fijar las variables de entrada en sus valores nominales óptimos puede descartar importantes estructuras de control al verlas como inviábiles o económicamente poco atractivas.
- Se recomienda que el estudio de un sistema de control por niveles que busque la operación óptima en base a criterios económicos se realice de manera integrada. Seleccionar la estructura de control regulador y posteriormente diseñar el control supervisor nos puede llevar a resultados sub-óptimos.
- Se recomienda analizar diversas estrategias de control supervisor a fin de poder identificar los potenciales beneficios económicos de cada una de ellas.
- En caso de disponer un modelo deficiente de la planta se recomienda el uso de un sistema de control supervisor del tipo Optimizador en Tiempo Real que incorpore una estrategia de adaptación de modelo.
- Si se desconoce la causa de la diferencia entre el modelo disponible y la planta real, y/o no se tiene certeza de la adecuación del modelo según Proposición 1, se recomienda utilizar métodos de optimización en tiempo real que incorporen modificadores de segundo orden.

6.2 Trabajos futuros

A partir de los trabajos realizados a lo largo del doctorado, han surgido diversas líneas de investigación a profundizar en el corto y mediano plazo:

- **Detección en línea de regiones activas:** en la tesis se propuso un problema que busca la o las estructuras de control óptimas para un sistema que incluye un nivel superior de detección de regiones activas. Sin embargo, el funcionamiento de dicho bloque para operar en línea es un tema de investigación en sí mismo y aún no hay un enfoque sistemático para diseñarlo.
- **Estudio del problema considerando la dinámica de la planta:** Los métodos propuestos a lo largo de la tesis buscan la operación óptima de estado estacionario de la planta. Posterior al diseño basado en estado estacionario es necesario simular la planta y verificar que durante los transitorios no se violen las restricciones operativas. Este enfoque secuencial puede revelar la necesidad de modificar los valores de consigna de las variables controladas a fin de que las restricciones no se violen durante los efectos transitorios. Al realizar cambios en los valores de consigna será necesario entonces re calcular su desempeño económico y consecuentemente puede ser necesario volver a iterar en la búsqueda de la estructura de control económicamente óptima. Incluir la dinámica de la planta en el problema de optimización nos proporcionaría un enfoque integrado para la selección de la estructura de control.
- **Algoritmo de búsqueda de puntos para aproximación cuadrática:** En el capítulo de optimización en tiempo real con adaptación por modificadores se propuso calcular los modificadores de hasta segundo orden mediante aproximación cuadrática. La calidad de la aproximación cuadrática de la diferencia planta-modelo, y por ende de los modificadores propuestos, depende de la adecuada distribución espacial de los puntos pertenecientes al conjunto interpolador. El algoritmo actual no asegura en cada iteración del optimizador en tiempo real tener una aproximación cuadrática de calidad. Consecuentemente, en problemas con restricciones, es posible que el resultado de optimizar el modelo adaptado nos lleve a puntos operativos que violan las restricciones. Es necesario entonces profundizar el estudio a fin de encontrar mejoras en el algoritmo que eviten dicho inconveniente.

Apéndice A

Modelos de Casos de Estudio

A.1 Modelo de Reactor Tanque Continuo Agitado (CSTR)

El reactor Williams-Otto [67], que consiste en un CSTR ideal se ilustra en la imagen A.1 y tienen lugar las siguientes reacciones:

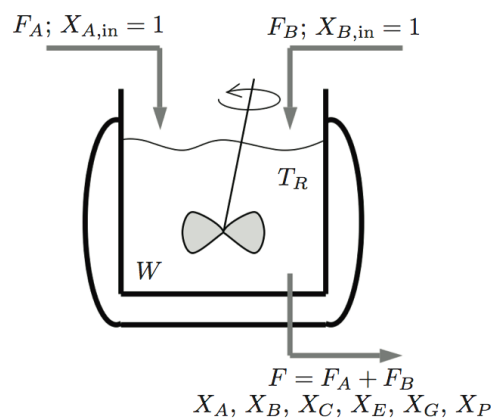
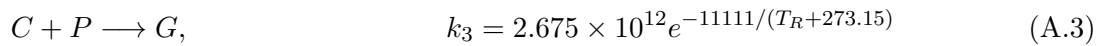
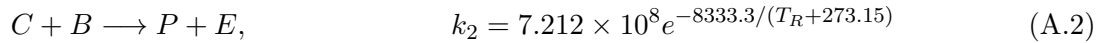


Figura A.1: Reactor de agitación continua

Los reactivos A y B se alimentan con caudales másicos F_A y F_B , respectivamente. Los productos deseados son P y E , C es un producto intermedio y G es un producto no deseado. El reactor funciona isotérmicamente a temperatura T_R , con una retención de masa constante W de 2105 kg. Las ecuaciones

del modelo en estado estacionario en base al balance de materia son:

$$0 = F_A - (F_A + F_B)X_A - Wr_1 \quad (\text{A.4})$$

$$0 = F_B - (F_A + F_B)X_B - \frac{M_B}{M_A}Wr_1 - Wr_2 \quad (\text{A.5})$$

$$0 = -(F_A + F_B)X_C - \frac{M_C}{M_A}Wr_1 - \frac{M_C}{M_B}Wr_2 - Wr_3 \quad (\text{A.6})$$

$$0 = -(F_A + F_B)X_P - \frac{M_P}{M_B}Wr_2 - \frac{M_P}{M_C}Wr_3 \quad (\text{A.7})$$

$$0 = -(F_A + F_B)X_G - \frac{M_G}{M_C}Wr_3 \quad (\text{A.8})$$

$$X_E = \frac{M_E}{M_P}X_P + \frac{M_E}{M_G}X_G \quad (\text{A.9})$$

donde,

$$r_1 = k_1X_AX_B, \quad r_2 = k_2X_BX_C, \quad r_3 = k_3X_CX_P \quad (\text{A.10})$$

Las variables X_i son la fracciones de masa de las especies i , M_i es el peso molecular de la especie i , W la retención de masa en el reactor y k_j es el coeficiente cinético de reacción j dado en [A.1-A.3](#).

A.2 Modelo de Evaporador

Consideramos el evaporador de circulación forzada descrito por Newell y Lee [68]. El diseño del evaporador se muestra en la Figura A.2 y sus principales variables con sus valores nominales se enumeran en la Tabla A.1. El evaporador se alimenta con la alimentación líquida F_1 y el licor de recirculación F_3 . El suministro de vapor P_{100} se utiliza para calentar el evaporador y la mezcla de vapor-líquido generada F_4 se fracciona en el separador. El líquido separado se recircula principalmente, mientras que F_2 se extrae como producto.

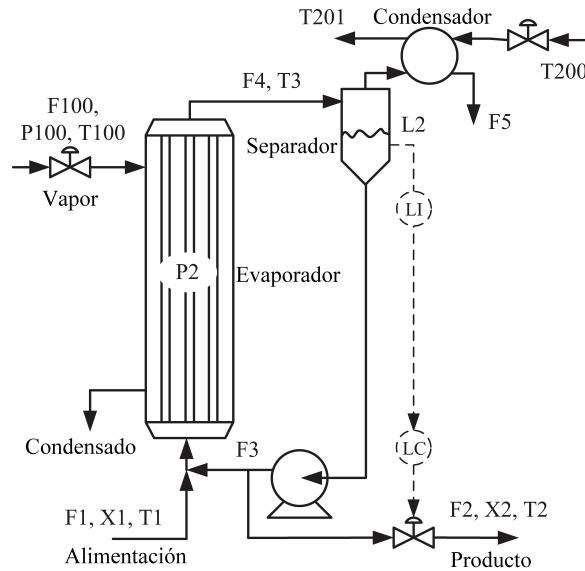


Figura A.2: Sistema de Evaporador.

Tabla A.1: Evaporador: Variables y Valores en el Punto Óptimo Nominal.

Variable	Descripción	Valor	Unidad
F_{200}	Caudal del agua refrigerante	217.74	[kg/min]
P_{100}	Presión del vapor de entrada	400.00	[kPa]
F_3	Caudal de circulación	24.72	[kg/min]
F_1	Caudal de alimentación	9.469	[kg/min]
X_2	Composición del producto	35.50	[%]
F_4	Caudal de la mezcla de salida	8.135	[kg/min]
F_5	Caudal del condensado	8.135	[kg/min]
T_2	Temperatura del producto	88.4	[°C]
T_3	Temperatura de la mezcla de salida	81.1	[°C]
P_2	Presión operativa	51.41	[kPa]
F_{100}	Caudal del vapor de entrada	9.434	[kg/min]
T_{100}	Temperatura del vapor de entrada	151.5	[°C]
T_{201}	Temperatura de salida del agua refrigeración	45.55	[°C]
X_1	Composición de la alimentación	5.00	[%]
T_1	Temperatura de la alimentación	40.0	[°C]
T_{200}	Temperatura de la entrada del agua refrigeración	25.0	[°C]
J	Ganancia económica	582.23	[\$/h]

Tenga en cuenta que el nivel L_2 del separador se controla mediante un controlador PI que manipula el caudal de producto F_2 . Las ecuaciones del modelo para el evaporador son [31]:

$$\frac{dL_2}{dt} = \frac{F_1 - F_4 - F_2}{20} \quad (\text{A.11})$$

$$\frac{dX_2}{dt} = \frac{F_1 X_1 - F_2 X_2}{20} \quad (\text{A.12})$$

$$\frac{dP_2}{dt} = \frac{F_4 - F_5}{4} \quad (\text{A.13})$$

$$T_2 = 0.5616P_2 + 0.3126X_2 + 48.43 \quad (\text{A.14})$$

$$T_3 = 0.507P_2 + 55.0 \quad (\text{A.15})$$

$$F_4 = \frac{Q_{100} - 0.07F_1(T_2 - T_1)}{38.5} \quad (\text{A.16})$$

$$T_{100} = 0.1538P_{100} + 90.0 \quad (\text{A.17})$$

$$Q_{100} = 0.16(F_1 + F_3)(T_{100} - T_2) \quad (\text{A.18})$$

$$F_{100} = \frac{Q_{100}}{36.6} \quad (\text{A.19})$$

$$Q_{200} = \frac{0.9576F_{200}(T_3 - T_{200})}{0.14F_{200} + 6.84} \quad (\text{A.20})$$

$$T_{201} = T_{200} + \frac{13.68(T_3 - T_{200})}{0.14F_{200} + 6.84} \quad (\text{A.21})$$

$$F_5 = \frac{Q_{200}}{38.5} \quad (\text{A.22})$$

donde además de las variables principales enumeradas en la Tabla [A.1](#), tenemos el nivel del separador L_2 [m], la tasa de flujo del producto F_2 [kg / min], la carga térmica Q_{100} [kW] y la carga del condensador Q_{200} [kW].

A.3 Modelo Sistema de Reacción-Separación con Reciclo

Consideramos el proceso reactor-separador-reciclo descrito en Jacobsen y Skogestad [44]. El diagrama de flujo del proceso se muestra en la Figura A.3 y las principales variables se enumeran en la Tabla A.2, junto con sus valores de estado estacionario en el óptimo nominal.

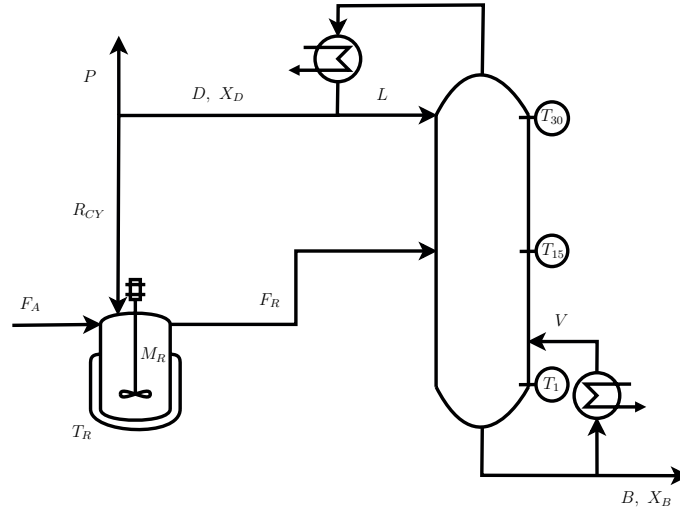
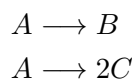


Figura A.3: Proceso Reactor-Separador-Reciclo.

Tabla A.2: Proceso reactor-separador-reciclo: Principales variables y sus valores en un óptimo nominal.

Variable	Descripción	Valor	Unidad
L	Caudal de reflujo en la columna	1.929	[mol/s]
V	Caudal de Vapor en la columna	2.121	[mol/s]
R_{CY}	Caudal de reciclaje	0.000	[mol/s]
T_R	Temperatura en el reactor	390.000	[K]
F_R	Caudal de salida del reactor	1.121	[mol/s]
M_R	Volumen de Reactor	11000.000	[mol]
P	Caudal de purga	0.192	[mol/s]
B	Caudal de fondo de Columna	0.927	[mol/s]
$X_{B,A}$	Fracción molar de A en el fondo de columna	0.087	
$X_{B,B}$	Fracción molar de B en el fondo de columna	0.900	
$X_{B,C}$	Fracción molar de C en el fondo de columna	0.013	
D	Caudal de destilado en la columna	0.192	[mol/s]
$X_{D,A}$	Fracción molar de A en el destilado de columna	0.068	
$X_{D,B}$	Fracción molar de B en el destilado de columna	0.071	
$X_{D,C}$	Fracción molar de C en el destilado de columna	0.861	
F_A	Caudal de alimentación de A en el reactor	1.030	[mol/s]
p_V	Precio de la energía	0.060	[\$/mol]
J	Ganancia económica	0.789	[\$/s]

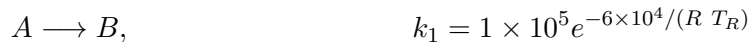
El reactor de tanque agitado continuamente (CSTR) se alimenta con el ingreso F_A de producto crudo A y con el flujo de reciclaje R_{CY} . Las reacciones que tienen lugar en el CSTR son:



donde B es el producto deseado y C un subproducto. El flujo de salida del reactor se envía a la columna de destilación. El producto deseado B se separa luego como un producto de fondo, y el subproducto

C y el A sin reaccionar se destilan. El flujo de destilado F_D luego se separa en el flujo de purga y reciclado al CSTR.

Ecuaciones del Modelo de Reactor Las reacciones se modelan con cinéticas de primer orden como sigue:



donde R es la constante del gas y T_R la temperatura del reactor. Los ratios de reacción para el modelo cinético de primer orden son:

$$r_1 = k_1 x_{R,A} \quad (\text{A.23})$$

$$r_2 = k_2 x_{R,A} \quad (\text{A.24})$$

Las ecuaciones de balance de masa en estado estacionario considerando un flujo de entrada y un flujo de salida:

$$\frac{dn_i}{dt} = F_F x_{F,i} + M_R v r - F_R x_{R,i}, \quad i = 1, 2, 3 \quad (\text{A.25})$$

donde $i = 1, 2, 3$ representa los componentes A , B y C , respectivamente; n_i , expresado en unidades de mol, es la retención del reactor del componente i ; r es un vector de ratios de reacción; v es una matriz con los coeficientes estequiométricos $([-1, -1; 1, 0; 0, 2])$; y x_R es la composición interior del reactor (que se considera igual a la composición en la corriente de salida). F_F representa la corriente en [mol/s] que ingresa al reactor,

$$F_F = R_{CY} + F_A \quad (\text{A.26})$$

$$F_F x_{F,i} = F_A x_{A,i} + R_{CY} x_{D,i} \quad (\text{A.27})$$

Además tenemos las siguientes ecuaciones:

$$R_{CY} = D(1 - P) \quad (\text{A.28})$$

$$\sum n_i = M_R \quad (\text{A.29})$$

$$x_{R,i} = n_i / M_R \quad (\text{A.30})$$

Ecuaciones del Modelo de Columna de Destilación La columna de destilación se basa en el modelo de columna 'A' descrito en Skogestad y Morari [77], que considera los siguientes supuestos: volatilidades relativas constantes, desbordamiento molar constante, presión constante en toda la columna, equilibrio en cada etapa y retención de vapor insignificante. Los parámetros de la columna de destilación se enumeran en la Tabla A.3.

Tabla A.3: Columna de destilacion: parametros principales

Parámetros	Valor
α_{AC}	0.70
α_{BC}	0.60
NT: número de platos	30
NF: ubicación del plato de alimentación	15
fracción líquida en alimentación	1
V_{max}	30

Equilibrio Vapor-Líquido: Para $j = 1, \dots, 30$:

$$y_{j,A} = \frac{\alpha_{AC}x_{j,A}}{1 + (\alpha_{AC} - 1)x_{j,A} + (\alpha_{BC} - 1)x_{j,B}} \quad (\text{A.31})$$

$$y_{j,B} = \frac{\alpha_{BC}x_{j,B}}{1 + (\alpha_{AC} - 1)x_{j,A} + (\alpha_{BC} - 1)x_{j,B}} \quad (\text{A.32})$$

$$y_{j,A} + y_{j,B} + y_{j,C} = 1 \quad (\text{A.33})$$

donde $x_{j,i}$ son las fracciones molares en la fase líquida e $y_{j,i}$ en la fase de vapor, α_{AC}, α_{BC} son las volatilidades relativas. Consulte la Tabla A.3 para conocer los valores utilizados en este ejemplo.

Flujos en Columna:

- Flujos de vapor (asumiendo flujos molares constantes):

$$V_j = V; \quad j \leq NF \quad (\text{A.34})$$

$$V_j = V + (1 - q_{FR})F_R; \quad j > NF \quad (\text{A.35})$$

donde q_{FR} es la fracción de líquido en la alimentación de la columna (para este ejemplo = 1)

- Flujos de líquido (fórmula de vertedero de Francis):

$$L_j = K_{bf}(M_j - M_{uw})^{1.5}; \quad 2 \leq j \leq NF \quad (\text{A.36})$$

$$L_j = K_{uf}(M_j - M_{uw})^{1.5}; \quad NF < j < NT \quad (\text{A.37})$$

$$L_j = L; \quad j = NT \quad (\text{A.38})$$

donde K_{bf}, K_{uf} son parámetros constantes para la parte inferior y superior de la bandeja de alimentación, M_{uw} es la retención constante de líquido debajo del vertedero.

Balances de material (utilizamos la versión de estado estacionario de las ecuaciones diferenciales):

Caso General ($j \neq 1, NF, NT$)

$$\frac{dM_j}{dt} = L_{j+1} - L_j + V_{j-1} - V_j; \quad j \neq NF, \quad 2 \leq j \leq NT - 1 \quad (\text{A.39})$$

$$\frac{dMx_{j,i}}{dt} = L_{j+1}x_{j+1,i} - L_jx_{j,i} + V_{j-1}y_{j-1,i} - V_jy_{j,i}; \quad j \neq NF, \quad 2 \leq j \leq NT - 1 \quad (\text{A.40})$$

En la etapa de alimentación ($j = NF$):

$$\frac{dM_j}{dt} = L_{j+1} - L_j + V_{j-1} - V_j + F_R; \quad j = NF \quad (\text{A.41})$$

$$\frac{dMx_{j,i}}{dt} = L_{j+1}x_{j+1,i} - L_jx_{j,i} + V_{j-1}y_{j-1,i} - V_jy_{j,i} + F_Rx_{R,i}; \quad j = NF \quad (\text{A.42})$$

En el hervidor ($j = 1$):

$$\frac{dM_1}{dt} = L_2 - B - V_1; \quad (\text{A.43})$$

$$\frac{dMx_{1,i}}{dt} = L_2x_{2,i} - Bx_{B,i} - V_1y_{1,i}; \quad (\text{A.44})$$

En el Condensador ($j = NT$):

$$\frac{dM_{NT}}{dt} = V_{NT-1} - L - D; \quad (\text{A.45})$$

$$\frac{dMx_{NT,i}}{dt} = V_{NT-1}y_{NT-1,i} - Lx_{D,i} - Dx_{D,i}; \quad (\text{A.46})$$

Para calcular la derivada de la fracción molar ($\frac{dx}{dt}$):

$$\frac{dMx}{dt} = x \frac{dM}{dt} + M \frac{dx}{dt}; \quad (\text{A.47})$$

Ecuaciones adicionales para el nivel de la columna:

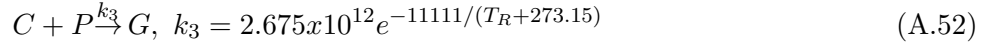
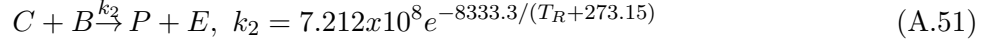
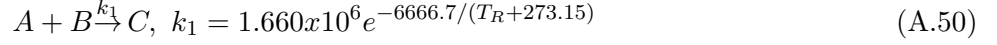
$$D = D_s + (M_{NT} - M_{D_s}) * K_{cd} \quad (\text{A.48})$$

$$B = B_s + (M_1 - M_{B_s}) * K_{cb} \quad (\text{A.49})$$

donde los parámetros D_s, B_s son flujos de referencia, M_{D_s}, M_{B_s} son las retenciones de referencia y K_{cb}, K_{cd} son las ganancias del controlador .

A.4 Modelo de Reactor Tanque Continuo Agitado (CSTR) usado en Adaptación Por Modificadores

La planta verdadera (realidad simulada) se describe mediante las siguientes reacciones:



donde k_1 , k_2 and k_3 son las constantes de reacción. Los productos deseados son P and E , C es un producto intermedio, y G es un producto indeseado. Los reactivos A y B se alimentan con caudales másicos F_A y F_B , la masa retenida en el reactor es 2105 Kg y la temperatura de operación isotérmica es T_R .

Las ecuaciones de la planta real en estado estacionario en base al balance de materia son:

$$0 = F_A - (F_A + F_B)X_A - W r_1 \quad (\text{A.53})$$

$$0 = F_B - (F_A + F_B)X_B - \frac{M_B}{M_A} W r_1 - W r_2 \quad (\text{A.54})$$

$$0 = -(F_A + F_B)X_C - \frac{M_C}{M_A} W r_1 - \frac{M_C}{M_B} W r_2 - W r_3 \quad (\text{A.55})$$

$$0 = -(F_A + F_B)X_P - \frac{M_P}{M_B} W r_2 - \frac{M_P}{M_C} W r_3 \quad (\text{A.56})$$

$$0 = -(F_A + F_B)X_G - \frac{M_G}{M_C} W r_3 \quad (\text{A.57})$$

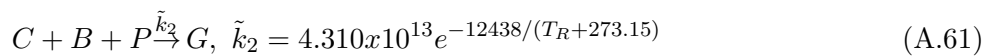
$$X_E = \frac{M_E}{M_P} X_P + \frac{M_E}{M_G} X_G \quad (\text{A.58})$$

donde,

$$r_1 = k_1 X_A X_B, \quad r_2 = k_2 X_B X_C, \quad r_3 = k_3 X_C X_P \quad (\text{A.59})$$

Las variables X_i son la fracciones de masa de las especies i , M_i es el peso molecular de la especie i , W la retención de masa en el reactor y k_j es el coeficiente cinético de reacción j dado en [A.50-A.52](#).

El modelo disponible que aproxima la realidad se basa en dos reacciones:



Las ecuaciones del modelo en estado estacionario en base al balance de materia son:

$$0 = F_A - (F_A + F_B)X_A + Wr_1 - Wr_2 \quad (\text{A.62})$$

$$0 = F_B - (F_A + F_B)X_B - \frac{M_B}{M_A}2Wr_1 - \frac{M_B}{M_A}Wr_2 \quad (\text{A.63})$$

$$0 = -(F_A + F_B)X_P - \frac{M_P}{M_A}Wr_1 - \frac{M_P}{M_A}Wr_2 \quad (\text{A.64})$$

$$0 = -(F_A + F_B)X_E - \frac{M_E}{M_A}Wr_1 \quad (\text{A.65})$$

$$X_G = \frac{M_G}{M_E}X_E + \frac{M_G}{M_P}X_P \quad (\text{A.66})$$

donde,

$$r_1 = k_1X_AX_B^2, \quad r_2 = k_2X_AX_BX_P \quad (\text{A.67})$$

donde los coeficientes cinéticos k_j están dados en en [A.60-A.61](#).

Bibliografía

- [1] V. Sakizlis, J. D. Perkins, y E. N. Pistikopoulos. Recent advances in optimisation-based simultaneous process and control design. *Comp. Chem. Eng.*, 28:2069–2086, 2004.
- [2] M. J. Mohideen, J. D Perkins, y E. N. Pistikopoulos. Optimal design of dynamic systems under uncertainty. *AIChE J.*, 42(8):2251–2272, 1996.
- [3] L. A. Ricardez-Sandoval, H. M. Budman, y P. L. Douglas. Integration of design and control for chemical processes: A review of the literature and some new results. *Annual Reviews in Control*, 33:158–171, 2009.
- [4] J. A. Heath, I. K. Kookos, y J. D. Perkins. Process control structure selection based on economics. *AIChE J.*, 46(10):1998–2016, 2000.
- [5] S. Skogestad. Plantwide control: the search for the self-optimizing control structure. *J. Process Contr.*, 10:487–507, 2000.
- [6] A. Marchetti, B. Chachuat, y D. Bonvin. Modifier-adaptation methodology for real-time optimization. *Ind. Eng. Chem. Res.*, 48(13):6022–6033, 2009.
- [7] A. G. Marchetti, A. Ferramosca, y A. H. González. Steady-state target optimization designs for integrating real-time optimization and model predictive control. *J. Process Contr.*, 24:129–145, 2014.
- [8] I. K. Kookos. Real-time regulatory control structure selection based on economics. *Ind. Eng. Chem. Res.*, 44:3993–4000, 2005.
- [9] H. Manum y S. Skogestad. Self-optimizing control with active set changes. *J. Process Contr.*, 22:873–883, 2012.
- [10] M. S. Bazaraa, H. D. Sherali, y C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, New Jersey, 3rd edition, 2006.
- [11] P. A. Bahri, J. A. Bandoni, y J. A. Romagnoli. Effect of disturbances in optimizing control: Steady-state open-loop backoff problem. *AIChE J.*, 42(4):983–994, 1996.
- [12] P. A. Bahri, J. A. Bandoni, G. W. Barton, y J. A. Romagnoli. Back-off calculations in optimising control: A dynamic approach. *Comp. Chem. Eng.*, 19, Suppl.:S699–S708, 1995.
- [13] J. L. Figueroa, P. A. Bahri, J. A. Bandoni, y J. A. Romagnoli. Economic impact of disturbances and uncertain parameters in chemical processes – A dynamic back-off analysis. *Comp. Chem. Eng.*, 20(4):453–461, 1996.
- [14] L. T. Narraway, J. D. Perkins, y G. W. Barton. Interaction between process design and process control: economic analysis of process dynamics. *J. Process Contr.*, 1:243–250, 1991.
- [15] C. Loeblein y J. D. Perkins. Structural design for on-line process optimization: I. Dynamic economics of MPC. *AIChE J.*, 45(5):1018–1029, 1999.
- [16] M. Soliman, C. L. E. Swartz, y R. Baker. A mixed-integer formulation for back-off under constrained predictive control. *Comp. Chem. Eng.*, 32:2409–2419, 2008.
- [17] L. T. Narraway y J. D. Perkins. Selection of process control structure based on linear dynamic economics. *Ind. Eng. Chem. Res.*, 32:2681–2692, 1993.
- [18] J. A. Heath, I. K. Kookos, y J. D. Perkins. Process control structure selection based on economics. *AIChE J.*, 46(10):1998–2016, 2000.
- [19] A. Psaltis, I. K. Kookos, y C. Kravaris. Plant-wide control structure selection methodology based on economics. *Comp. Chem. Eng.*, 52:240–248, 2013.

- [20] D. A. Zumoffen, L. Braccia, y A. G. Marchetti. Economic plant-wide control design with backoff estimations using internal model control. *J. Process Contr.*, 40:93–105, 2016.
- [21] I. K. Kookos y J. D. Perkins. Control structure selection based on economics: Generalization of the back-off methodology. *AIChE J.*, 62(9):3056–3064, 2016.
- [22] I. K. Kookos y J. D. Perkins. An algorithm for simultaneous process design and control. *Ind. Eng. Chem. Res.*, 40:4079–4088, 2001.
- [23] I. K. Kookos y J. D. Perkins. An algorithmic method for the selection of multivariable process control structures. *J. Process Contr.*, 12:85–99, 2002.
- [24] M. Sharifzadeh y N. F. Thornhill. Optimal selection of control structure using a steady-state inversely controlled process model. *Comp. Chem. Eng.*, 38:126–138, 2012.
- [25] M. Sharifzadeh y N. F. Thornhill. Integrated design and control using a dynamic inversely controlled process model. *Comp. Chem. Eng.*, 48:121–134, 2013.
- [26] V. Bansal, J. D. Perkins, y E. N. Pistikopoulos. A case study in simultaneous desing and control using rigorous mixed-integer dynamic optimization models. *Ind. Eng. Chem. Res.*, 41:760–778, 2002.
- [27] Z. Yuan, B. Chen, G. Sin, y R. Gani. State-of-the-art and progress in the optimization-based simultaneous design and control for chemical processes. *AIChE J.*, 58(6):1640–1659, 2012.
- [28] I. J. Halvorsen, S. Skogestad, J. C. Morud, y V. Alstad. Optimal selection of controlled variables. *Ind. Eng. Chem. Res.*, 42:3273–3284, 2003.
- [29] J. Jäschke, Y. Cao, y V. Kariwala. Self-optimizing control - A survey. *Annual Reviews in Control*, 43:199–223, 2017.
- [30] V. Alstad y S. Skogestad. Null space method for selecting optimal measurement combinations as controlled variables. *Ind. Eng. Chem. Res.*, 46:846–853, 2007.
- [31] V. Kariwala, Y. Cao, y S. Janardhanan. Local self-optimizing control with average loss minimization. *Ind. Eng. Chem. Res.*, 47:1150–1158, 2008.
- [32] V. Alstad, S. Skogestad, y E. S. Hori. Optimal measurement combinations as controlled variables. *J. Process Contr.*, 19:138–148, 2009.
- [33] J. Jäschke y S. Skogestad. Optimal operation by controlling the gradients to zero. In *Proceedings of the 18th IFAC World Congress*, pages 593–595, Milano, Italy, 2011.
- [34] W. Hu, L. M. Umar, G. Xiao, y V. Kariwala. Local self-optimizing control of constrained processes. *J. Process Contr.*, 22:488–493, 2012.
- [35] R. Yelchuru y S. Skogestad. Convex formulations for optimal selection of controlled variables and measurements using mixed integer quadratic programming. *Journal of Process Control*, 22(6):995–1007, 2012.
- [36] A. G. Marchetti y D. A. Zumoffen. Self-optimizing control structures with minimum number of process-dependent controlled variables. *Ind. Eng. Chem. Res.*, 53(24):10177–10193, 2014.
- [37] L. Ye, Y. Cao, y X. Yuan. Global approximation of self-optimizing controlled variables with average loss minimization. *Ind. Eng. Chem. Res.*, 54:12040–12053, 2015.
- [38] L. Ye, Y. Cao, y S. Skogestad. Global self-optimizing control for uncertain constrained process systems. *IFAC-PapersOnLine*, 50(1):4672–4677, 2017.
- [39] G. François, B. Srinivasan, y D. Bonvin. Equivalence between neighboring-extremal control and self-optimizing control for steady-state optimization of dynamical systems. *Ind. Eng. Chem. Res.*, 53(18):7470–7478, 2014.
- [40] A. Maarleveld y J. E. Rijnsdorp. Constraint control on distillation columns. *Automatica*, 6:51–58, 1970.
- [41] Y. Arkun y G. Stephanopoulos. Studies in the synthesis of control structures for chemical processes: Part IV. Design of steady-state optimizing control structures for chemical process units. *AIChE J.*, 26(6):975–991, 1980.
- [42] A. Reyes-Lúa y S. Skogestad. Systematic design of active constraint switching using classical advanced control structures. *Ind. Eng. Chem. Res.*, 59(6):2229–2241, 2020.
- [43] T. E. Marlin y A. N. Hrymak. Real-time operations optimization of continuous processes. In *AIChE Symposium Series - CPC-V*, volume 93, pages 156–164, 1997.

-
- [44] M. G. Jacobsen y S. Skogestad. Active constraint regions for optimal operation of chemical processes. *Ind. Eng. Chem. Res.*, 50(19):11226–11236, 2011.
- [45] B. Chachuat, B. Srinivasan, y D. Bonvin. Adaptation strategies for real-time optimization. *Comp. Chem. Eng.*, 33: 1557–1567, 2009.
- [46] C. Y. Chen y B. Joseph. On-line optimization using a two-phase approach: An application study. *Ind. Eng. Chem. Res.*, 26:1924–1930, 1987.
- [47] J. F. Forbes y T. E. Marlin. Design cost: A systematic approach to technology selection for model-based real-time optimization systems. *Comp. Chem. Eng.*, 20:717–734, 1996.
- [48] W. Gao, S. Wenzel, y S. Engell. Comparison of modifier adaptation schemes in real-time optimization. *IFAC-PapersOnLine*, 48(8):182–187, 2015.
- [49] B. Chachuat, A. Marchetti, y D. Bonvin. Process optimization via constraints adaptation. *J. Process Contr.*, 18: 244–257, 2008.
- [50] P. Tatjewski. Iterative optimizing set-point control - The basic principle redesigned. In *15th IFAC World Congress*, Barcelona, Spain, 2002.
- [51] W. Gao y S. Engell. Iterative set-point optimization of batch chromatography. *Comp. Chem. Eng.*, 29:1401–1409, 2005.
- [52] A. Marchetti, B. Chachuat, y D. Bonvin. A dual modifier-adaptation approach for real-time optimization. *J. Process Contr.*, 20:1027–1037, 2010.
- [53] T. Faulwasser y D. Bonvin. On the use of second-order modifiers for real-time optimization. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, volume 19, pages 7622–7628, 2014.
- [54] A. G. Marchetti, G. François, T. Faulwasser, y D. Bonvin. Modifier adaptation for real-time optimization—methods and applications. *Processes*, 4(4), 2016.
- [55] M. Brdyś y P. Tatjewski. *Iterative Algorithms for Multilayer Optimizing Control*. Imperial College Press, London UK, 2005.
- [56] A. G. Marchetti. A new dual modifier-adaptation approach for iterative process optimization with inaccurate models. *Comp. Chem. Eng.*, 59:89–100, 2013.
- [57] P. D. Roberts. Broyden derivative approximation in ISOPE optimising and optimal control algorithms. In *11th IFAC Workshop on Control Applications of Optimisation*, volume 1, pages 283–288, St Petersburg, Russia, 2000.
- [58] W. Gao, S. Wenzel, y S. Engell. A reliable modifier-adaptation strategy for real-time optimization. *Computers and Chemical Engineering*, 91:318–328, 2016.
- [59] W. L. Luyben, B. D. Tyréus, y M. L. Luyben. *Plantwide Process Control*. McGraw-Hill, New York, 1998.
- [60] M. van de Wal y A. de Jager. A review of methods for input/output selection. *Automatica*, 37:487–510, 2001.
- [61] S. Skogestad y I. Postlethwaite. *Multivariable Feedback Control. Analysis and Design*. John Wiley & Sons, New York, 2 edition, 2005.
- [62] M. Sharifzadeh. Integration of process desing and control: A review. *Chem. Eng. Res. Des.*, 91(12):2515–2549, 2013.
- [63] L. Braccia, P. A. Marchetti, P. Luppi, y D. Zumoffen. Multivariable control structure design based on mixed-integer quadratic programming. *Ind. Eng. Chem. Res.*, 56(39):11228–11244, 2017.
- [64] J. Nocedal y S. J. Wright. *Numerical Optimization*. Springer-Verlag, New York, 1999.
- [65] Yi Cao. Direct and indirect gradient control for static optimisation. *International Journal of Automation and Computing*, 2(1):60–66, 2005.
- [66] L. Ye, Y. Cao, Y. Li, y Z. Song. Approximating necessary conditions of optimality as controlled variables. *Ind. Eng. Chem. Res.*, 52(2):798–808, 2013.
- [67] T. J. Williams y R. E. Otto. A generalized chemical processing model for the investigation of computer control. *AIEE Trans.*, 79:458, 1960.
- [68] R. B. Newell y P. L. Lee. *Applied Process Control - A Case Study*. Prentice Hall, Englewood Cliffs, NJ, 1989.
-

- [69] A. G. Marchetti y D. Zumoffen. On the links between real-time optimization, neighboring-extremal control, and self-optimizing control. In *European Control Conference, ECC'13*, Zurich, Switzerland, 2013.
- [70] C. E. Garcia y M. Morari. Optimal operation of integrated processing systems. Part I: Open-loop on-line optimizing control. *AIChE J.*, 27(6):960–968, 1981.
- [71] M. L. Darby, M. Nikolaou, J. Jones, y D. Nicholson. RTO: An overview and assessment of current practice. *J. Process Contr.*, 21:874–884, 2011.
- [72] J. Jäschke y S. Skogestad. NCO tracking and self-optimizing control in the context of real-time optimization. *J. Process Contr.*, 21:1407–1416, 2011.
- [73] L. M. Umar, Y. Cao, y V. Kariwala. Incorporating feedforward action into self-optimising control policies. *Canadian J. Chem. Eng.*, 92(1):90–96, 2014.
- [74] M. Baotić, F. Borrelli, A. Bemporad, y M. Morari. Efficient on-line computation of constrained optimal control. *SIAM Journal on Control and Optimization*, 47(5):2470–2489, 2008.
- [75] R. E. Young. Petroleum refining process control and real-time optimization. *IEEE Contr. Syst. Mag.*, 26(6):73–83, 2006.
- [76] M. Baotic, F. Borrelli, A. Bemporad, y M. Morari. Efficient on-line computation of constrained optimal control. *SIAM J. Control. Optim.*, 47:2470–2489, 2008.
- [77] S. Skogestad y M. Morari. Understanding the dynamic behavior of distillation columns. *Ind. Eng. Chem. Res.*, 27(10):1848–1862, 1988.
- [78] A. G. Marchetti, T. Faulwasser, y D. Bonvin. A feasible-side globally convergent modifier-adaptation scheme. *Submitted to J. Process Control*, 2016.
- [79] A. R. Conn, K. Scheinberg, y L. N. Vicente. *Introduction to Derivative-Free Optimization. MPS/SIAM Book Series on Optimization*. SIAM, 2009.
- [80] C.T. Kelley. *Iterative Methods for Optimization*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 1999. ISBN 9780898714333.
- [81] W. Gao, R. Hernández, y S. Engell. A study of explorative moves during modifier adaptation with quadratic approximation. *Processes*, 4(4), 2016.
- [82] Y. Zhang y J. F. Forbes. Extended design cost: A performance criterion for real-time optimization systems. *Comp. Chem. Eng.*, 24:1829–1841, 2000.