

Juegos poblacionales y modelos dinámicos de pago: Un nuevo paradigma para control y optimización

Germán Obando^{a,*}, Juan Martínez-Piazuelo^b, Nicanor Quijano^c, Carlos Ocampo-Martínez^b

^aUniversidad de Nariño, Pasto (Colombia)

^bUniversitat Politècnica de Catalunya - BarcelonaTECH, Barcelona (España)

^cUniversidad de los Andes, Bogotá (Colombia)

To cite this article: Obando, G., Martínez-Piazuelo, J., Quijano, N., Ocampo-Martínez, C. 2024. Population games and dynamic payment models: A new paradigm for control and optimization. Revista Iberoamericana de Automática e Informática Industrial 21, 287-305. <https://doi.org/10.4995/riai.2024.21215>

Resumen

En la última década, se han venido desarrollando técnicas inspiradas por la naturaleza y la economía con el fin de resolver problemas de control y toma de decisiones. En este artículo, se presenta este nuevo paradigma que combina los juegos poblacionales y los modelos dinámicos de pago. Se introducen conceptos fundamentales en torno a estas áreas, incluyendo un desarrollo matemático formal (basado en teoría de pasividad para sistemas dinámicos, estabilidad de Lyapunov e invarianza de conjuntos) que valida su uso tanto para abordar problemas de optimización como para diseñar sistemas de control en lazo cerrado con restricciones (físicas y operacionales). Específicamente, nos enfocamos en problemas cuyos objetivos se alinean con la distribución dinámica de recursos y el alcance de equilibrios generalizados de Nash. La pertinencia del paradigma formulado se ilustra a través de diferentes problemas de ingeniería con aplicaciones en múltiples campos.

Palabras clave: Juegos evolutivos, Asignación dinámica de recursos, Equilibrios de Nash, Modelos dinámicos de pago.

Population games and dynamic payment models: A new paradigm for control and optimization

Abstract

Over the course of the past ten years, methods that are derived from both economics and nature have been created in order to solve optimization and control problems. Delving in this framework, we present a new paradigm that combines population games and dynamic payment models. Fundamental concepts in these areas are introduced, including a formal mathematical development (which is based on passivity theory, Lyapunov stability, and sets invariance) that supports the design of optimization algorithms and closed-loop control systems with physical and operational constraints. Specifically, we focus on problems whose goals align with the dynamic distribution of resources and the attainment of generalized Nash equilibria. The relevance of the formulated paradigm is illustrated through different engineering problems with applications in multiple fields.

Keywords: Evolutionary game theory, Dynamic resource allocation, Nash equilibria, Payoff dynamic models.

1. Introducción

La teoría de juegos es el nombre que se le da a la metodología que utiliza herramientas matemáticas para modelar y analizar situaciones en las que interactivamente se toman decisiones. Estas nociones se habían comenzado a discutir a finales del siglo XIX, principios del siglo XX, pero no es hasta el desa-

rollo del teorema del minimax por parte de John von Neumann (von Neumann, 1928) y el posterior libro que publica con Oskar Morgenstern (von Neumann and Morgenstern, 1947) que no se aterrizan las ideas iniciales. La diseminación del concepto toma varios años, y es gracias a los aportes de la corporación RAND que estas nociones logran permear diferentes áreas del conocimiento (Bhattacharya, 2021).

*Autor para correspondencia: gdo bando@udenar.edu.co
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

Por otro lado, la toma óptima de decisiones por parte de múltiples agentes que interactúan entre sí por medio de una red (sea física o de comunicaciones), ha sido uno de los temas que más auge ha tenido en la comunidad de sistemas dinámicos y control en las últimas décadas. ¿La razón? Estos problemas emergen en ingeniería, ciencias sociales y económicas, sistemas urbanos o inteligencia artificial, donde se encuentran aplicaciones como el análisis de redes sociales (Jackson et al., 2008), manejo y/o control de redes inteligentes (Mojica-Nava et al., 2013; Ananduta et al., 2018), redes inalámbricas (Han et al., 2019), ciberseguridad (Pawlick and Zhu, 2021), infraestructura crítica (Rass et al., 2020), o sistemas ciber-físicos (Groot et al., 2014).

Una de las maneras de modelar estos sistemas complejos de gran escala con múltiples decisiones y acciones, en la que los diferentes entes/controladores interactúan entre sí, es mediante el uso de la teoría de juegos (Muros, 2021; Martínez et al., 2023). Dentro de los trabajos desarrollados, se puede mencionar el de Bacci et al. (2016), el cual muestra la relación entre los juegos, la optimización y el aprendizaje para el procesamiento de señales en red. Otros ejemplos similares incluyen la carga de vehículos eléctricos (Grammatico et al., 2016), la coordinación de redes de robots (Jaleel and Shamma, 2020; Park and Barreiro-Gomez, 2023), problemas de congestión vehicular (Kara et al., 2022), problemas de control de pandemias y epidemias (Martins et al., 2023), técnicas de aprendizaje por refuerzo (Gao and Pavel, 2021), respuesta a la demanda (Genis-Mendoza et al., 2022) y manejo de recursos y regulación de sistemas de agua (Lu et al., 2022). En estos trabajos, los autores abordan los problemas desde diferentes ángulos de la teoría de juegos. Algunos parten de los juegos matriciales, lo cuales son conocidos por su forma normal, en la que la interacción simultánea entre jugadores se da de manera estática y cada jugador se entiende como un ente individual. Por otra parte, en juegos continuos, los jugadores pueden elegir entre una amplia gama de estrategias que cambian con el tiempo.

Hay otros juegos que se conocen como juegos dinámicos, que son los que utilizan un mecanismo de aprendizaje que permite ajustar las acciones basadas en eventos previos. Estos juegos se distinguen por tres problemas principales: i) modelar el ambiente en el que interactúan los jugadores; ii) modelar los objetivos que persiguen los jugadores; y iii) describir el orden en el que los jugadores toman decisiones y la cantidad de información que tienen. En este caso, se asume que la interacción ocurre entre un gran número (desconocido) de jugadores, y lo que nos interesa estudiar es la proporción de individuos que finalmente utilizan una estrategia u otra. Los juegos evolutivos, que fueron creados con base en el comportamiento ecológico, se clasifican como juegos dinámicos. Las nociones de las estrategias evolutivamente estables fueron desarrolladas por Maynard-Smith y Price, quienes fueron los pioneros en este concepto (Maynard-Smith and Price, 1973). Después, Taylor y Jonker en 1978 crean el modelo de replicadores (*replicator dynamics*), que se utiliza ampliamente en aplicaciones de ingeniería, para ver el comportamiento dinámico y su relación con la parte genética (Taylor and Jonker, 1978). Complementariamente, mediante la intro-

ducción de protocolos de revisión y *mean dynamics*, se han desarrollado aproximaciones de juegos evolutivos desde el punto de vista de sistemas económicos (Sandholm, 2010).

Este artículo muestra algunos ejemplos de asignación dinámica de recursos a través de juegos poblacionales y modelos dinámicos de pago (Park et al., 2019). El objetivo principal es destacar la utilidad y la idoneidad de estas técnicas para modelar dinámicas de sistemas complejos de ingeniería, así como para diseñar estrategias de gestión y control siguiendo políticas particulares y contemplando restricciones físicas y operativas, tanto locales como globales. Las estrategias desarrolladas por medio de este paradigma son de fácil implementación física, como se puede ver en trabajos como Martínez-Piazuelo et al. (2022a); Barreiro-Gomez et al. (2021), en los que se muestran criterios para seleccionar parámetros y su implementación. Por otra parte, este nuevo paradigma, también es capaz de abarcar problemáticas como los retrasos (Obando et al., 2016; Park and Leonard, 2021), dando pie a una nueva alternativa respecto a otras técnicas desarrolladas, cuya implementación tiene sus dificultades.

Concretamente, nuestro artículo se centra en presentar conceptos esenciales sobre los modelos dinámicos de pago y los juegos poblacionales, incluyendo una fundamentación matemática rigurosa. Para ello, exponemos un compendio de herramientas que formalizan el uso de juegos poblacionales y modelos dinámicos de pago tanto para diseñar estrategias de control como para resolver problemas de optimización. A diferencia de otros artículos que utilizan estas herramientas para resolver problemas específicos (desde enrutamiento de tráfico (Tan et al., 2022), pasando por el diseño de filtros para mejorar imágenes (Horstmann, 2005), hasta explicar mecanismos que usan las células cancerosas para desarrollar resistencia a las terapias (Brown et al., 2016)), nuestro propósito es establecer un marco de referencia general. Para tal fin, proponemos una serie de resultados que, por un lado, evidencian propiedades que pueden explotarse en el diseño de controladores y algoritmos de optimización, y por otro, imponen condiciones que tipifican los problemas que pueden abordarse utilizando el paradigma propuesto. En particular, mostramos cómo un conjunto de ecuaciones que modela la selección natural puede estabilizar un sistema y llevarlo a un estado preestablecido. Además, probamos formalmente que cierto tipo de dinámicas, típicamente utilizadas para modelar el comportamiento de grupos de vehículos, pueden emplearse para solucionar problemas de optimización.

Para que el lector pueda llevarse una idea de este nuevo paradigma, varios ejemplos se ilustran a lo largo del documento¹. Con ellos, se busca mostrar la versatilidad que tienen las técnicas desarrolladas, no sólo desde el punto de vista teórico sino de implementación. Los ejemplos de simulación se pueden realizar utilizando Matlab/Python como se ha hecho en este trabajo, o por medio de herramientas como EGTtools24 STAR Methods. Esta es una biblioteca híbrida C++/Python que ofrece implementaciones rápidas y paralelas de métodos analíticos, numéricos y Monte-Carlo. El objetivo es proporcionar una forma estandarizada para que los investigadores accedan y compartan modelos y métodos relacionados con la teoría de juegos evolu-

¹ Los archivos de simulación de los ejemplos pueden encontrarse en: <https://github.com/elessardari/poblacionalesRIAI24.git>

tivos. EGTtools24 permite simulaciones de múltiples agentes a gran escala, procesos de Markov a pequeña escala y análisis determinista clásico (Domingos et al., 2023). Existen otras aproximaciones de múltiples agentes como las de Izquierdo et al. (2019).

Algunas ideas que presentamos en este artículo se derivan de la evolución de trabajos anteriores. Se puede encontrar un resumen de los trabajos que se presentaron hasta el año 2017 en Quijano et al. (2017). Se han hecho contribuciones desde entonces en términos de dinámicas distribuidas en tiempo continuo (Barreiro-Gomez et al., 2017a) y en tiempo discreto (Martinez-Piazuelo et al., 2022a), así como en la combinación de técnicas en sistemas híbridos (Ochoa et al., 2021). Últimamente, se han desarrollado temas como aplicaciones para vehículos autónomos no tripulados (Barreiro-Gomez et al., 2021) o la combinación de técnicas de control para aplicaciones en redes de agua (Barreiro-Gomez et al., 2017b; Obando et al., 2022). Además, se han reportado recientemente resultados significativos sobre la relación entre estas poblaciones dinámicas y los equilibrios generalizados de Nash (Martinez-Piazuelo et al., 2022c,d,b; Sánchez-Amores et al., 2023; Martinez-Piazuelo et al., 2023).

El resto del artículo se organiza de la siguiente manera: la Sección 2 introduce los principales conceptos en torno a juegos poblacionales y dinámicas de pago. La Sección 3 explora la relación que existe entre juegos poblacionales y clásicos. Las Secciones 4 y 5 describen aplicaciones en control y optimización, respectivamente. Finalmente, en las Secciones 6 y 7 se presenta una discusión alrededor del paradigma desarrollado y se plantean las conclusiones.

2. Juegos Poblacionales y Modelos Dinámicos de Pago

Como su nombre lo indica, las dinámicas poblacionales son modelos que describen la evolución de grupos de individuos (personas o animales) y las razones de dicha evolución. Cabe resaltar que, cuando hablamos de evolución, nos referimos al cambio que experimentan las poblaciones en, por ejemplo, composición o tamaño. Por la amplitud del término, las dinámicas poblacionales admiten múltiples interpretaciones (Sandholm, 2010; Quijano et al., 2017). Una de ellas, considera las dinámicas poblacionales como modelos que describen el proceso de selección natural, en el que las estrategias más rentables son las que consiguen más adeptos. Notemos que al hablar de *estrategias* y *rentabilidad*, estamos usando terminología típica de la teoría de juegos. Efectivamente, la mejor forma de explicar los fundamentos de las dinámicas poblacionales es a través de un juego en el que intervienen una gran cantidad de agentes.

Consideremos un conjunto de $N \in \mathbb{Z}_{>0}$ poblaciones, cada una compuesta por un número grande y constante de agentes que toman decisiones estratégicamente. A lo largo de este artículo, el conjunto de poblaciones se indexa por $\mathcal{P} = \{1, 2, \dots, N\}$, el conjunto de estrategias disponibles para los agentes de la población $k \in \mathcal{P}$ se indexa por $\mathcal{S}^k = \{1, 2, \dots, n^k\}$, donde $n^k \in \mathbb{Z}_{\geq 2}$, y el total de agentes de cada población $k \in \mathcal{P}$ se modela como un continuo de masa $m^k \in \mathbb{R}_{>0}$. Asimismo, en un instante de tiempo t , la masa de agentes que selecciona la estrategia $i \in \mathcal{S}^k$ en la población $k \in \mathcal{P}$ está dada por $x_i^k(t) \in \mathbb{R}_{\geq 0}$. Por lo tanto, los vectores $\mathbf{x}^k(t) = \text{col}(x_1^k(t), x_2^k(t), \dots, x_{n^k}^k(t)) \in$

Δ^k y $\mathbf{x}(t) = \text{col}(\mathbf{x}^1(t), \mathbf{x}^2(t), \dots, \mathbf{x}^N(t)) \in \Delta$ denotan las distribuciones estratégicas de la población $k \in \mathcal{P}$ y de toda la sociedad, respectivamente. Aquí, $\text{col}(\cdot)$ es la operación de concatenación en vector columna; $n = \sum_{k \in \mathcal{P}} n^k$; $\Delta^k = \{\mathbf{x}^k(t) \in \mathbb{R}_{\geq 0}^{n^k} : \sum_{i \in \mathcal{S}^k} x_i^k(t) = m^k\}$; y $\Delta = \{\mathbf{x}(t) \in \mathbb{R}_{\geq 0}^n : \mathbf{x}^k(t) \in \Delta^k, \forall k \in \mathcal{P}\}$. En particular, Δ^k es el conjunto de todas las posibles distribuciones estratégicas de la población $k \in \mathcal{P}$, mientras que Δ es el conjunto de todas las posibles distribuciones estratégicas de toda la sociedad.

En resumen, los ingredientes de un juego poblacional son dos: un conjunto de agentes (jugadores) divididos en poblaciones, y un conjunto de estrategias que dichos agentes pueden elegir para jugar. Como describiremos a continuación, el proceso que siguen los agentes para decidir entre las múltiples estrategias que tienen a disposición está bien establecido.

2.1. Protocolos de Revisión

Por facilidad, asumiremos que en nuestro juego poblacional interviene una única población (entonces, $N = 1$). Sin embargo, los conceptos explicados en esta subsección pueden ser fácilmente extrapolados a juegos poblacionales con múltiples poblaciones. Consideremos que el conjunto de estrategias que los jugadores tienen a disposición (esto es, $\mathcal{S}^1 = \{1, 2, \dots, n^1\}$) tiene un pago asociado que varía en el tiempo. Vamos a denotar el pago asociado a la i -ésima estrategia como $p_i^1(t)$. Demos claridad a este marco de referencia a través de un ejemplo: supongamos que un grupo de animales tiene que elegir entre dos hábitats para vivir. La elección de cada animal dependerá de la cantidad de comida que ofrezca cada uno de esos hábitats. Bajo estas suposiciones, los jugadores son los animales; las estrategias son los hábitats disponibles, es decir, $\mathcal{S}^1 = \{1, 2\}$; y las funciones de pago, $p_1^1(t)$ y $p_2^1(t)$, la cantidad de comida disponible en el hábitat 1 y 2, respectivamente. Observemos algo interesante: si inicialmente hay más comida en el hábitat 1 (es decir, $p_1^1(0) > p_2^1(0)$), es probable que muchos animales elijan inicialmente vivir allí. Esta elección hará que la comida del hábitat 1 se vaya consumiendo y eventualmente escasee, provocando que en un tiempo determinado, \bar{t} , el vivir en el hábitat 2 sea más rentable para los animales (esto es, $p_2^1(\bar{t}) > p_1^1(\bar{t})$). En conclusión, la situación planteada da lugar a un juego dinámico en el que los jugadores pueden cambiar de estrategia a medida que el tiempo transcurre para mejorar el pago que reciben. Las dinámicas poblacionales justamente modelan este proceso.

Concretamente, el proceso descrito por las dinámicas poblacionales, que se ilustra en la Figura 1, puede resumirse en dos etapas que se ejecutan periódicamente. En la *etapa 1*, se empareja a los jugadores de forma aleatoria. Luego, en la *etapa 2*, los jugadores emparejados tienen la posibilidad o bien de continuar con la estrategia que están jugando o de cambiarse a la estrategia de su par. Esta elección depende del pago que el jugador y su par estén recibiendo al momento del emparejamiento.

La decisión señalada en la *Etapa 2* se modela a través de una función conocida como “protocolo de revisión”. Esta función describe la probabilidad de que un jugador cambie de estrategia. Por ejemplo, supongamos que se empareja un jugador que está jugando la estrategia $i \in \mathcal{S}^1$ con otro que está jugando la estrategia $j \in \mathcal{S}^1$. El protocolo de revisión ρ_{ij} es una función que tiene como entrada los pagos de las estrategias i y j (esto es, $p_i^1(t)$ y $p_j^1(t)$), y la proporción de jugadores que ha elegido estas

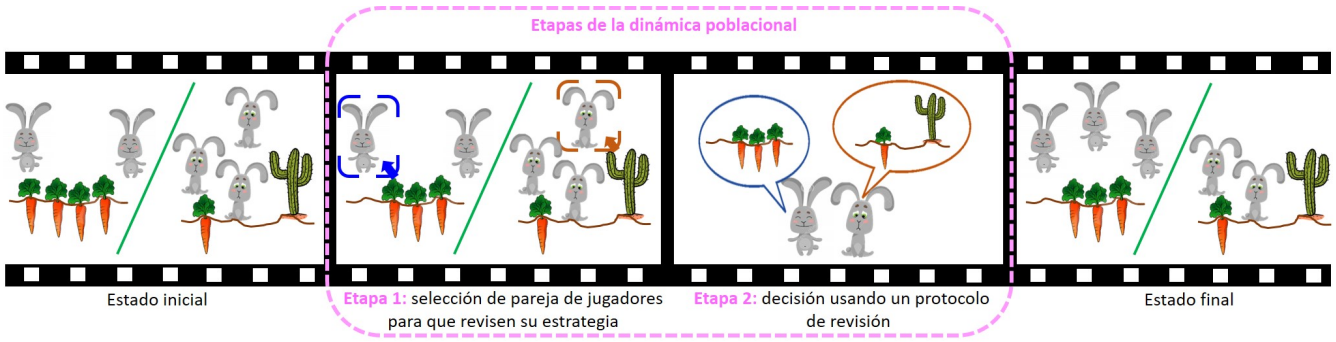


Figura 1: Proceso de toma de decisión de los individuos inmersos en un juego poblacional.

estrategias (es decir, $x_i^1(t)$ y $x_j^1(t)$). El resultado del protocolo de revisión es un número no negativo. Entre más grande sea este número, hay mayor probabilidad de que el jugador cambie la i -ésima por la j -ésima estrategia.

Convenientemente, si aplicamos el proceso anterior una y otra vez, y al mismo tiempo asumimos que el número de jugadores es lo suficientemente grande, podemos modelar los cambios en la población de jugadores que eligen la i -ésima estrategia a través de la siguiente ecuación diferencial:

$$\dot{x}_i^1(t) = \underbrace{\sum_{j \in \mathcal{S}^1} x_j^1(t) \rho_{ji}}_{\text{proporción de jugadores que adoptan la estrategia } i} - \underbrace{\sum_{j \in \mathcal{S}^1} x_i^1(t) \rho_{ij}}_{\text{proporción de jugadores que cambian la estrategia } i \text{ por otra}} \quad (1)$$

En el campo de las dinámicas poblacionales, esta ecuación diferencial se conoce como *mean dynamics*. Observe que, bajo el proceso descrito por las *mean dynamics*, los jugadores únicamente pueden cambiar de estrategia, no pueden reproducirse ni morir. Lo anterior implica que la masa de la población, m^1 , se mantiene constante durante todo el tiempo. Esta propiedad se expresa formalmente en el Teorema 2.1².

Teorema 2.1. *Considere las mean dynamics en (1). Si $\mathbf{x}^1(0) \in \Delta^1$, entonces $\mathbf{x}^1(t) \in \Delta^1$, para todo $t \geq 0$.*

Demostración. Definamos la masa poblacional como $m^1(t) := \sum_{i \in \mathcal{S}^1} x_i^1(t)$. Note que, a lo largo de las trayectorias de (1),

$$\begin{aligned} \dot{m}^1(t) &= \sum_{i \in \mathcal{S}^1} \dot{x}_i^1(t) \\ &= \sum_{i \in \mathcal{S}^1} \sum_{j \in \mathcal{S}^1} x_j^1(t) \rho_{ji} - \sum_{i \in \mathcal{S}^1} \sum_{j \in \mathcal{S}^1} x_i^1(t) \rho_{ij} \\ &= \sum_{j \in \mathcal{S}^1} \sum_{i \in \mathcal{S}^1} x_i^1(t) \rho_{ij} - \sum_{i \in \mathcal{S}^1} \sum_{j \in \mathcal{S}^1} x_i^1(t) \rho_{ij}. \end{aligned} \quad (2)$$

Por lo tanto, intercambiando el orden de las sumas del segundo término, $\dot{m}^1(t) = \sum_{j \in \mathcal{S}^1} \sum_{i \in \mathcal{S}^1} x_i^1(t) \rho_{ij} - \sum_{j \in \mathcal{S}^1} \sum_{i \in \mathcal{S}^1} x_i^1(t) \rho_{ij} = 0$. Esto implica que $m^1(t) = m^1(0) = m^1$, por suposición. Por otro lado, en vista que $\rho_{ij} \geq 0$, note que si para algún tiempo \hat{t} , $x_i^1(\hat{t}) = 0$, entonces $\dot{x}_i^1(\hat{t}) \geq 0$ de acuerdo con (1). Es decir, $x_i^1(t)$ no puede decrecer más allá de cero. En conclusión, $\mathbf{x}^1(t) \in \mathbb{R}_{\geq 0}^n$, completando la prueba \square

Como era de esperarse, las *mean dynamics* dependen explícitamente del protocolo de revisión que estén empleando los jugadores. Así que la elección de dicho protocolo determina la forma cómo evoluciona la población (la forma cómo los jugadores emigran de una estrategia a otra). Cabe resaltar que, en la literatura, se han reportado diversos protocolos de revisión (ver, por ejemplo, la Tabla I en Barreiro-Gomez et al. (2017a)). Uno de los más famosos es el *protocolo de comparación modificado*:

$$\rho_{ij} = \frac{\max\{0, p_j^1(t) - p_i^1(t)\}}{x_i^1(t)}. \quad (3)$$

Note que, bajo este protocolo, mientras mayor sea la ganancia al cambiar la estrategia i por la j , más probable es que un jugador decida hacer este cambio. Por otro lado, dicha probabilidad es nula si el jugador que ha elegido la estrategia i está recibiendo un pago mayor que un jugador que ha elegido la estrategia j . Es posible demostrar que al aplicar este protocolo, las *mean dynamics* se reducen a la muy conocida ecuación de las *projection dynamics* (Sandholm, 2010),

$$\dot{x}_i^1(t) = p_i^1(t) - \bar{p}(t), \quad \forall i \in \mathcal{S}^1, \quad (4)$$

donde $\bar{p}(t) = \frac{1}{n^1} \sum_{j \in \mathcal{S}^1} p_j^1(t)$ es el pago promedio. A partir de esta ecuación, se puede observar que si la población que está jugando la i -ésima estrategia recibe un pago superior al promedio, entonces crecerá (si $p_i^1(t) > \bar{p}(t)$, entonces $\dot{x}_i^1 > 0$), es decir, más individuos jugarán esta estrategia. Por el contrario, si recibe un pago inferior al promedio, entonces decrecerá (si $p_i^1(t) < \bar{p}(t)$, entonces $\dot{x}_i^1 < 0$).

La Tabla 1 muestra los protocolos de revisión que han sido estudiados más ampliamente en la literatura. Al aplicar el *protocolo de comparación* obtenemos las *Smith dynamics*, mientras que el *protocolo de imitación proporcional* nos lleva a obtener los *replicator dynamics*.

²En la literatura existen abundantes pruebas de este teorema cuando en (1) se reemplaza ρ_{ij} por un protocolo de revisión específico, e.g., Weibull (1997). Sin embargo, el caso general no había sido estudiado.

Tabla 1: Algunos protocolos de revisión (Sandholm, 2010).

Nombre	Expresión
protocolo de imitación proporcional	$\rho_{ij} = \frac{x_j^1(t)}{m^1} [p_j^1(t) - p_i^1(t)]_+$
protocolo de comparación	$\rho_{ij} = [p_j^1(t) - p_i^1(t)]_+$
protocolo de comparación modificado	$\rho_{ij} = \frac{[p_j^1(t) - p_i^1(t)]_+}{x_i^1(t)}$

$[\cdot]_+ := \max(\cdot, 0)$.

2.2. Dinámicas Evolutivas

Tal y como ocurre con el *protocolo de imitación proporcional*, los protocolos de revisión moldean la forma como una población evoluciona. En otras palabras, la unión de las *mean dynamics* y un protocolo de revisión específico da lugar a unas dinámicas evolutivas (EDM, del inglés *evolutionary dynamics model*), las cuales definiremos a continuación.

Definición 1. La evolución temporal de la distribución estratégica $\mathbf{x}(t)$ se describe por un EDM de la forma

$$\dot{\mathbf{x}}(t) = \mathcal{V}(\mathbf{x}(t), \mathbf{p}(t)), \quad \mathbf{x}(0) \in \Delta, \quad (5)$$

donde $\mathcal{V} : \Delta \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ es Lipschitz continua y resulta de usar un protocolo de revisión específico en las *mean dynamics*. Se tiene que $\mathcal{V}(\mathbf{x}(t), \mathbf{p}(t)) \in T\Delta(\mathbf{x}(t))$, para todo $t \geq 0$ ($T\Delta(\mathbf{x}(t))$ denota el cono tangente de Δ en $\mathbf{x}(t)$).

Con base en la Definición 1, el EDM puede verse como un sistema dinámico en tiempo continuo cuya entrada es el vector de pagos $\mathbf{p}(t) = \text{col}(\mathbf{p}^1(t), \mathbf{p}^2(t), \dots, \mathbf{p}^N(t)) \in \mathbb{R}^n$, con $\mathbf{p}^k(t) = \text{col}(p_1^k(t), p_2^k(t), \dots, p_{n^k}^k(t)) \in \mathbb{R}^{n^k}$, para todo $k \in \mathcal{P}$. Aquí, $p_i^k(t) \in \mathbb{R}$ es el pago percibido por los agentes de la población $k \in \mathcal{P}$ que seleccionan la estrategia $i \in S^k$ en el tiempo t . En general, el vector de pagos $\mathbf{p}(t)$ es generado por un modelo de dinámicas de pago (PDM, del inglés *payoff dynamics model*) que se define de la siguiente forma (Park et al., 2019).

Definición 2. El vector de pagos $\mathbf{p}(t)$ es determinado por un PDM de la forma

$$\begin{aligned} \dot{\mathbf{q}}(t) &= \mathcal{W}(\mathbf{q}(t), \mathbf{x}(t)), & \mathbf{q}(0) &\in \mathbb{R}^d \\ \mathbf{p}(t) &= \mathcal{H}(\mathbf{q}(t), \mathbf{x}(t)), \end{aligned} \quad (6)$$

donde $\mathcal{W} : \mathbb{R}^d \times \Delta \rightarrow \mathbb{R}^d$ es Lipschitz continua, y $\mathcal{H} : \mathbb{R}^d \times \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^n$ es continuamente diferenciable y Lipschitz continua.

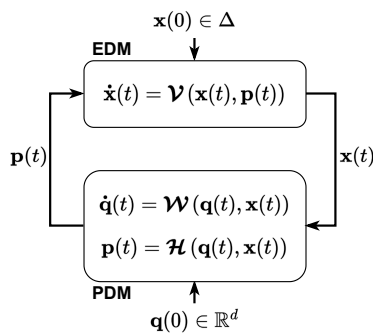


Figura 2: Sistema interconectado EDM-PDM.

De acuerdo con la Definición 1, la distribución estratégica de la sociedad, $\mathbf{x}(t)$, pertenece al conjunto Δ en todos los tiempos $t \geq 0$. Esta propiedad permite aplicar el esquema de juegos poblacionales en problemas de distribución dinámica de recursos, en los cuales ciertas cantidades de recursos, m^1, m^2, \dots, m^N , deben ser distribuidas en una serie de alternativas (estrategias). Los sistemas de respuesta a la demanda (Srikantha and Kundur, 2017), los sistemas de agua (Pashaie et al., 2017) y la coordinación de la carga de vehículos eléctricos (Martinez-Piazuelo et al., 2021) son algunos ejemplos de aplicaciones. En este tipo de aplicaciones, generalmente se estudian las condiciones necesarias para garantizar la convergencia de $\mathbf{x}(t)$ a un equilibrio de Nash (NE, del inglés *Nash equilibrium*). En pocas palabras, un NE es una distribución estratégica $\mathbf{x}^* \in \Delta$ tal que ningún agente puede mejorar su pago desviándose del equilibrio de manera unilateral. La definición formal de un NE es la siguiente.

Definición 3. Un NE es una distribución estratégica $\mathbf{x}^* \in \mathcal{N}$, donde

$$\mathcal{N} = \left\{ \mathbf{x} \in \Delta : x_i^k > 0 \Rightarrow p_i^k = \max_{j \in S^k} p_j^k, \forall k \in \mathcal{P} \right\}. \quad (7)$$

A su vez, el PDM permite incorporar dinámicas en la generación de pagos $p_i^k(t)$, para todo $i \in S^k$ y todo $k \in \mathcal{P}$. Como veremos más adelante, el PDM permitirá incluir las dinámicas de, por ejemplo, un sistema a controlar; o también será útil para incluir restricciones en procesos de optimización.

3. Relación con Juegos Matriciales

La teoría de juegos, que se formalizó en los años 40s por medio del libro seminal von Neumann and Morgenstern (1947), planteaba la interacción entre agentes a partir de nociones matriciales en juegos de dos jugadores con dos o tres estrategias. Estos juegos podían ser de suma cero (es decir, lo que gana uno es lo que pierde el otro), o juegos no cooperativos en los que múltiples ejemplos y dilemas se han discutido desde la guerra fría, periodo en el que esta teoría comenzó a consolidarse no sólo en el área económica sino en en otras áreas como las de conflictos sociales (Axelrod, 1984). Esta es la razón por la cual muchos lectores podrían estar preguntándose ¿cómo se relacionan este tipo de juegos con los conceptos que se tienen de la teoría de juegos clásica? Para ello, vamos a dar una breve descripción utilizando únicamente los *replicator dynamics*, ya que estas son las más conocidas y utilizadas en diferentes áreas del conocimiento.

Originalmente, cuando Taylor and Jonker (1978) introdujeron por primera vez los *replicator dynamics* inspirados en los comportamientos biológicos, su enfoque recalaba en la parte tradicional de los juegos matriciales. En otras palabras, Taylor y Jonker no tenían dinámicas en un PDM, y su modelo dinámico evolutivo se limitaba únicamente a la interacción que podría llegar a existir en términos de una matriz de pago como se hacía en la teoría de juegos clásica matricial. Por lo tanto, el PDM se convierte en $p_i(t) = \sum_{j=1}^n J_1^{ij} x_j$. En este caso, se asume que sólo se tiene una población, esto es, $N = 1$, y que el número de estrategias total es $n = n^1$. El juego corresponderá a tomar un

par de individuos de la población y enfrentarlos en un juego de 2×2 . Al no tener un PDM dinámico, se asume que los pagos están dados por una matriz cuadrada $J_1 \in \mathbb{R}^{n \times n}$, donde las filas corresponderán a las acciones de uno de los jugadores y las columnas al otro. De esta manera, las dinámicas estarán dadas por

$$\dot{x}_i(t) = x_i(t) \left[\sum_{j=1}^n J_1^{ij} x_j - \sum_{j=1}^n \sum_{k=1}^n J_1^{kj} x_j x_k \right], \quad (8)$$

para todo $i, j = 1, \dots, n$. La deducción de estas dinámicas se explicará a cabalidad en la Sección 4.1.

Para poder ilustrar el comportamiento de estas dinámicas, recurriremos a una matriz de pago del jugador P_1 dada por

$$J_1 = \begin{bmatrix} R & S \\ T & P \end{bmatrix}, \quad (9)$$

donde la primera fila (columna) corresponde a la estrategia cooperar, y la segunda a la estrategia no cooperar (para el jugador P_2 , la matriz de pago será $J_2 = J_1^T$). En este caso, dependiendo de la selección de las constantes R, S, T, P se pueden lograr diferentes tipos de juegos (dilemas), los cuales se describen a continuación.

1. **Dilema del prisionero:** en el dilema del prisionero, se tiene que $T > R$ y $P > S$ ($P - S > 0 > R - T$). El estado en el que ambos jugadores no cooperan es un único NE. Si ninguna otra condición se plantea, la no cooperación es una estrategia dominante. Sin embargo, para asegurar que la mutua cooperación es preferible sobre la mutua no cooperación, se requiere adicionalmente garantizar que $R > P$.
2. **Anti-coordinación, Snowdrift, Hawk-Dove o Chicken:** en este tipo de juegos, los jugadores se benefician cuando utilizan la estrategia opuesta de su contendor. En este caso, los pagos deberían satisfacer el hecho de tener $T > R$ y $S > P$ ($(P - S), (R - T) < 0$). De esta forma, emergen dos NE en términos de estrategias puras, que corresponden a que cada uno de los jugadores utilizan estrategias opuestas (es decir, un jugador coopera y el otro no). En el caso de juegos tipo *snowdrift*, se impone una condición adicional que sería $R > S$.
3. **Coordinación, Stag-Hunt o Batalla de los Sexos:** en este tipo de juegos de coordinación, los jugadores se benefician cuando utilizan la misma estrategia que su oponente. En el caso en el que $R > T$ y $P > S$ ($(P - S), (R - T) > 0$), los dos NE que emergen son los de mutua cooperación o mutua no cooperación. El juego de *sag-hunt*, que es un juego específico de coordinación requiere la condición adicional de $T > P$ para que se dé.
4. **Harmonía:** en un juego de armonía, similar al del dilema del prisionero, la cooperación es la que logra el mayor valor del juego. En este caso, se debe satisfacer que $R > T$ y $S > P$ por lo que el único NE que se tendría es el de la mutua cooperación.

Desde el punto de vista de los *replicator dynamics*, se tendrá una ecuación general de la forma

$$\dot{x}_1 = x_1(Rx_1 + Sx_2 - Rx_1^2 - Px_2^2 - (S + T)x_1x_2). \quad (10)$$

Como $x_1 + x_2 = 1$, al reemplazar se obtiene que

$$\dot{x}_1 = x_1(1 - x_1)(x_1((P - S) + (R - T)) - (P - S)). \quad (11)$$

Nótese que $\dot{x}_2 = -\dot{x}_1$. La Figura 3 muestra la evolución de las dinámicas para los cuatro casos previamente descritos. En este caso, se muestra la evolución de las dinámicas para diferentes condiciones iniciales. Como se puede observar, dependiendo el tipo de juego que se tenga y las condiciones iniciales, se llega a un punto de equilibrio diferente. En el caso del dilema del prisionero, la no cooperación es una estrategia dominante y es el punto de equilibrio al que se llega independientemente de dónde arranque la proporción de individuos. El caso contrario al dilema del prisionero es el de armonía en el que la mutua cooperación es el punto de equilibrio. Para los casos de anti-coordinación y coordinación, emergen tres puntos de equilibrio. Para el caso de anti-coordinación, los NE son inestables, siendo el punto de equilibrio un valor entre 0 y 1, independientemente de la condición inicial en la que arranque la proporción de individuos. El caso contrario se da en el caso de coordinación. Dependiendo de las condiciones iniciales, el valor tenderá o bien a 0 o a 1.

La diferencia con los análisis tradicionales que se hacen en la teoría de juegos convencional es que, a partir de las dinámicas que tenemos descritas por medio de la ecuación diferencial (10), podríamos hacer un análisis de los puntos de equilibrio que coincidiría con las descripciones hechas verbalmente. Para entender un poco cómo se haría un análisis de un sistema dinámico en el que las PDM tengan dinámicas y el comportamiento sea no lineal, presentaremos a continuación algunas aplicaciones en el diseño de sistemas de control.

4. Control

Esta sección expone el uso del par EDM-PDM (Definiciones 1 y 2) para controlar sistemas dinámicos. Específicamente, nos enfocaremos en la aplicación de los *replicator dynamics*, un tipo de EDM. Además, para que la exposición sea más sencilla, únicamente abordaremos juegos en los que interviene una única población. En este sentido, y con el fin de abreviar la notación, eliminaremos el superíndice que diferencia una población de otra en juegos con múltiples poblaciones. Por ejemplo, en lugar de escribir $x^1(t)$ para denotar la distribución estratégica de la única población del juego, escribiremos sólo $x(t)$.

Un punto clave relacionado con los conceptos presentados en la Sección 2 es que los jugadores involucrados en el juego eventualmente elegirán estrategias que produzcan pagos por encima del promedio. Como todas las EDM, los *replicator dynamics* modelan este tipo de comportamiento. El principio detrás del modelo de los *replicator dynamics* es simple: la tasa a la cual el porcentaje de la población que usa una estrategia cambia depende únicamente de qué tan grande (o pequeño) es el pago de esa estrategia comparado con los pagos de las demás estrategias. Si el pago de la estrategia considerada es mayor que el promedio, entonces el porcentaje de la población que usa esa estrategia crecerá, mientras que si el pago es menor que el promedio, entonces el porcentaje va a decrecer.

Los *replicator dynamics* fueron desarrollados en 1978 por Taylor y Jonker, y, desde entonces, sus propiedades han sido

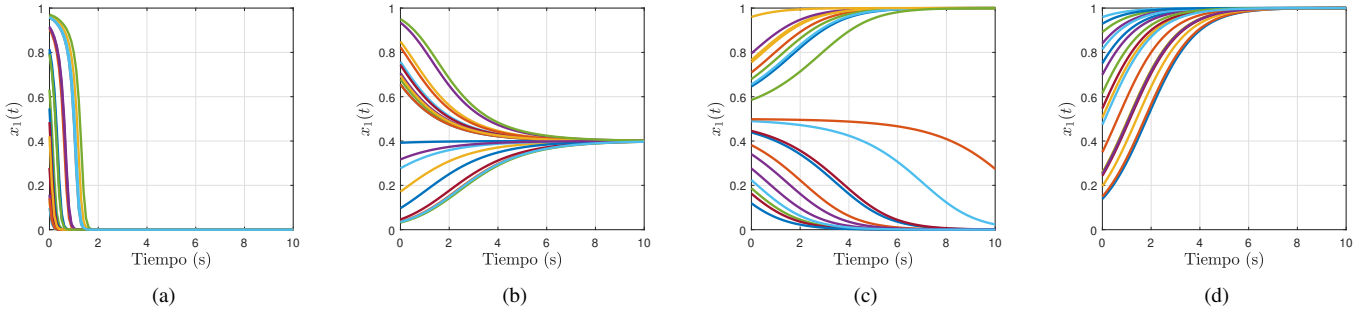


Figura 3: Evolución de la proporción de individuos $x_1(t)$ para múltiples condiciones iniciales para los diferentes dilemas y las dinámicas poblacionales basadas en los replicator dynamics: (a) dilema del prisionero, (b) anti-coordinación, (c) coordinación, (d) armonía.

ampliamente estudiadas. La mayor parte de la investigación inicial en torno a esta EDM estaba enfocada a explorar sus aplicaciones en biología y economía. Sin embargo, su uso no está restringido a esas áreas. En años recientes, algunos autores han demostrado el potencial que tiene la ecuación de los replicator dynamics para optimizar y controlar una variedad de sistemas. Por ejemplo, regular la temperatura dentro de un edificio de manera óptima o proporcionar niveles adecuados de servicio en los sistemas de distribución de agua.

A continuación, haremos una revisión de algunas propiedades de estabilidad del modelo de los replicator dynamics, en las que el concepto de pasividad³ juega un papel fundamental para que emerjan condiciones deseables cuando se implementan controladores basados en dicho modelo. Específicamente, la pasividad hace posible la aparición de puntos de equilibrio asintóticamente estables en el sistema en lazo cerrado.

4.1. Modelo de los Replicator Dynamics

Los replicator dynamics surgen del protocolo de imitación proporcional (Tabla 1). En concreto, al reemplazar dicho protocolo en las mean dynamics (1), obtenemos el siguiente conjunto de ecuaciones diferenciales:

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j \in S} x_j(t) \frac{x_i(t)}{m} [p_i(t) - p_j(t)]_+ \\ &\quad - x_i(t) \sum_{j \in S} \frac{x_j(t)}{m} [p_j(t) - p_i(t)]_+, \quad \forall i \in S, \\ &= \frac{1}{m} \sum_{j \in S} x_j(t) x_i(t) (p_i(t) - p_j(t)), \quad \forall i \in S, \\ &= \frac{x_i(t)}{m} \left(p_i(t) \sum_{j \in S} x_j(t) - \sum_{j \in S} x_j(t) p_j(t) \right), \quad \forall i \in S, \\ &= \frac{x_i(t) \sum_{j \in S} x_j(t)}{m} \left(p_i(t) \sum_{j \in S} x_j(t) - \frac{\sum_{j \in S} x_j(t) p_j(t)}{\sum_{j \in S} x_j(t)} \right), \quad \forall i \in S. \end{aligned} \tag{12}$$

³Decimos que un sistema es pasivo si sólo es capaz de disipar y/o almacenar energía, mas no de generarla. Formalmente, un sistema dinámico, $\dot{e} = f(e, u)$, $y = h(e, u)$, con entrada u y salida y se dice pasivo si existe una función positiva semidefinida $V(e)$ (conocida como función de almacenamiento) tal que $V(e) \leq u^T y$ (Khalil, 2002).

Finalmente, usando el hecho de que $\sum_{j \in S} x_j(t) = m$ (Teorema 2.1), se tiene la ecuación de los replicator dynamics

$$\dot{x}_i(t) = x_i(t) \left(p_i(t) - \frac{1}{m} \sum_{j \in S} x_j(t) p_j(t) \right), \quad \forall i \in S. \tag{13}$$

Cabe notar que la tasa de crecimiento $\dot{x}_i(t)$ es positiva si y sólo si la i -ésima función de pago, $p_i(t)$, es mayor que el promedio (calculado como $\sum_{j \in S} x_j(t) p_j(t) / m$). Es decir, las poblaciones más exitosas aumentan, mientras que las menos exitosas disminuyen. Este comportamiento está estrechamente relacionado con el proceso de selección natural, en el que los individuos más aptos tienen un mejor potencial para sobrevivir y reproducirse.

Como los replicator dynamics provienen de las mean dynamics, entonces satisfacen el Teorema 2.1. Es decir, el tamaño de la población permanece invariante. Esta característica es crucial en muchas aplicaciones, especialmente en aquellas relacionadas con la optimización y el control, donde la invarianza de la masa poblacional garantiza satisfacer ciertas restricciones como se ilustra en algunas de las aplicaciones descritas en este documento.

4.2. Propiedades Dinámicas de los Replicator Dynamics

Teorema 4.1. Pasividad: Sea $[x_1^*, \dots, x_n^*]^T \in \Delta$ un punto del sistema descrito por los replicator dynamics en (13). Dicho sistema es pasivo si se define su entrada como $u = [p_1(t), \dots, p_n(t)]^T$ y su salida como $y = [(x_1(t) - x_1^*), \dots, (x_n(t) - x_n^*)]^T$.

Demostración. La prueba que reproducimos a continuación se puede encontrar en Weibull (1997).

Usando la Divergencia de Kullback–Leibler,

$$V(x) = - \sum_{i=1}^n x_i^* \ln \left(\frac{x_i}{x_i^*} \right), \tag{14}$$

como la función de almacenamiento del sistema, tenemos que $\dot{V}(x) = \sum_{i=1}^n (x_i(t) - x_i^*) p_i(t) = u^T y$. Esto implica que la energía almacenada es igual a la potencia suministrada al sistema a lo largo del tiempo. Por lo tanto, el sistema es pasivo y sin pérdidas. \square

La pasividad es una característica clave porque permite la interconexión de un controlador basado en los *replicator dynamics* con cualquier planta que sea pasiva evitando problemas de inestabilidad. Esta propiedad se formaliza en el siguiente teorema (tomado de Khalil (2002), Teorema 6.1 y Lema 6.6).

Teorema 4.2. *Considere el sistema realimentado mostrado en la Figura 4.*

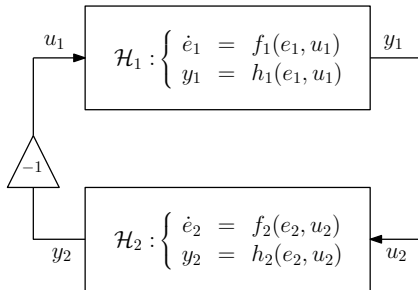


Figura 4: Realimentación negativa de dos sistemas. Se asume que \mathcal{H}_1 y \mathcal{H}_2 son pasivos.

Asuma que se satisfacen las siguientes condiciones:

- El origen es un punto de equilibrio tanto del sistema \mathcal{H}_1 como \mathcal{H}_2 .
- $h_1(0, 0) = h_2(0, 0) = 0$.
- \mathcal{H}_1 y \mathcal{H}_2 son pasivos.

Entonces el origen del sistema en lazo cerrado es estable.

El par EDM–PDM de la Figura 2 tiene la misma estructura del sistema realimentado de la Figura 4. Si lo pensamos como un sistema de control, las PDM modelan la planta controlada y las EDM (que en nuestro caso son los *replicator dynamics*) el controlador. Observe que, en este escenario, las funciones de pago dependen de la salida de un sistema dinámico (el sistema controlado) en lugar de depender directamente de la distribución de los jugadores entre las estrategias disponibles. La selección de funciones de pago es crucial para lograr los objetivos de control. Sin embargo, esta selección sigue siendo un problema de investigación abierto.

A continuación, presentamos dos ejemplos de uso de los *replicator dynamics* para controlar sistemas dinámicos. En el primer ejemplo, emplearemos los *replicator dynamics* como regulador de nivel en un sistema multi-tanque. En el segundo, los usaremos para controlar la temperatura en una edificación. En ambos casos haremos énfasis en el proceso de diseño y, en particular, en la selección apropiada de las funciones de pago.

4.3. Aplicación de los Replicator Dynamics a la Distribución de Agua en Sistemas Multi-Tanque

4.3.1. Definición del Problema

Imaginemos un sistema que tiene tres tanques dispuestos en la configuración mostrada en la Figura 5. De esos tres tanques, dos son fuente y uno receptor. Cada tanque fuente recibe agua del exterior y la vierte al tanque receptor a través de una válvula cuya apertura puede graduarse. Además, el tanque receptor tiene un sumidero por el que evacúa hacia el exterior el agua

que recibe. Supongamos también que nuestro objetivo es que, eventualmente, los tres tanques alberguen la misma cantidad de agua. Este sencillo problema que, de hecho, tiene aplicaciones relevantes en ingeniería (por ejemplo, la prevención de inundaciones en redes de drenaje urbano (Obando et al., 2022)), nos servirá para ilustrar los principios del uso de EDMs y PDMs como mecanismos de control de sistemas.

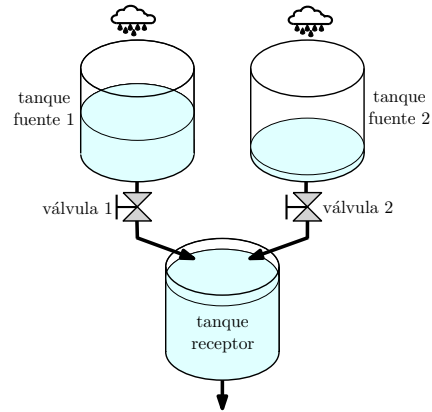


Figura 5: Sistema con tres tanques.

4.3.2. Diseño del Controlador

Consideremos una población de masa unitaria ($m = 1$), donde los jugadores que pueden elegir entre las tres estrategias que se describen en la Tabla 2 (esto es, $\mathcal{S} = \{1, 2, 3\}$). Así, entre más individuos elijan la estrategia 1, mayor será la apertura de la válvula 1. Lo mismo ocurre si muchos individuos eligen la válvula 2. Por otro lado, si muchos individuos eligen la estrategia 3, las válvulas 1 y 2 estarán casi cerradas.

Tabla 2: Estrategias para el problema de los tanques.

Estrategia	Descripción
1	incrementar la apertura de la válvula del tanque fuente 1 (válvula 1)
2	incrementar la apertura de la válvula del tanque fuente 2 (válvula 2)
3	no incrementar la apertura de ninguna de las válvulas

Una vez definidas las estrategias de nuestro juego poblacional, determinemos los pagos de dichas estrategias, teniendo en mente que nuestro objetivo de control es que la cantidad de agua almacenada en cada tanque sea aproximadamente la misma. Una elección conveniente de las funciones de pago es $p_i(t) = \beta q_i(t)$, donde $i = 1, 2, 3$; β es una constante positiva; y $q_1(t)$, $q_2(t)$ y $q_3(t)$ denotan el volumen de agua alojado en el tanque fuente 1, en el tanque fuente 2 y en el tanque receptor, respectivamente. Esta elección implica que el pago que reciben los jugadores que eligen las estrategias 1 y 2 es directamente proporcional al agua alojada en dichos tanques. Entre tanto, el pago de la estrategia 3 es directamente proporcional al agua alojada en el tanque receptor.

Observemos que, si por ejemplo la cantidad de agua alojada en el tanque 1 es mayor a las cantidades que albergan el tanque 2 y el tanque receptor, entonces será más rentable para los jugadores elegir la estrategia 1, es decir, incrementar la apertura de la válvula 1. Esto hará que el agua del tanque 1 se desaloje y

eventualmente alcance el mismo nivel del agua de los otros dos tanques. La situación es similar con el tanque 2. Finalmente, si la mayor cantidad de agua está alojada en el tanque receptor, será más atractivo para los jugadores elegir la estrategia 3. Esto provocará que las válvulas 1 y 2 se cierren, el tanque receptor no incremente más su volumen de agua y los tanques fuente 1 y 2 empiecen a alojar más líquido. Nuevamente, bajo estas condiciones, los tres tanques alcanzarán el mismo nivel de agua, cumpliendo así el objetivo de control buscado. La Figura 6 representa la situación descrita sobre cómo el problema de los tanques puede abordarse usando juegos poblacionales.

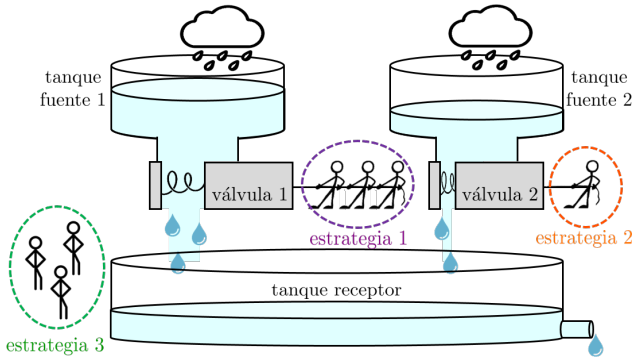


Figura 6: Analogía entre el problema de los tanques y un juego poblacional.

Establecidos los elementos clave del juego poblacional (jugadores, estrategias y pagos), es conveniente describir y contextualizar las dinámicas que regirán la evolución de la población de individuos (esto es, el EDM). Como veremos, estas dinámicas definirán las leyes de control que emplearemos. En este caso, el modelo de evolución escogido son los *replicator dynamics*; por lo tanto, tenemos que

$$\dot{x}_i(t) = x_i(t) \left(p_i(t) - \sum_{j=1}^3 x_j(t)p_j(t) \right), \quad i = 1, 2, 3. \quad (15)$$

Dado que $x_i(t)$ es la cantidad de individuos que eligen la i -ésima estrategia, observe que $x_1(t)$ corresponde a la proporción de jugadores que eligen abrir la válvula 1 (es decir, $x_1(t)$ es el porcentaje de apertura de la válvula 1); $x_2(t)$ es la proporción de jugadores que optan por abrir la válvula 2 (entonces, $x_2(t)$ es el porcentaje de apertura de la válvula 2); y $x_3(t)$ es la cantidad de jugadores que deciden no abrir ninguna de las válvulas. Así las cosas, nuestras leyes de control, que en este caso corresponden a la apertura de las válvulas, están gobernadas por el EDM en (15). Cabe anotar que la constante de proporcionalidad β de las funciones de pago es un parámetro de sintonización del controlador.

Por otra parte, notemos que las funciones de pago cambian dinámicamente puesto que dependen de los volúmenes de agua en cada tanque. Este comportamiento puede describirse a través del siguiente PDM, en el que hemos usado el modelo de Musingum (Chow, 1959) para caracterizar los cambios de agua en

los tanques:

$$\begin{aligned} \dot{q}_1(t) &= r_1 - a_1 x_1(t) q_1(t) \\ \dot{q}_2(t) &= r_2 - a_2 x_2(t) q_2(t) \\ \dot{q}_3(t) &= a_1 x_1(t) q_1(t) + a_2 x_2(t) q_2(t) - a_3 q_3(t) \\ p_1(t) &= \beta q_1(t) \\ p_2(t) &= \beta q_2(t) \\ p_3(t) &= \beta q_3(t), \end{aligned} \quad (16)$$

donde a_1 , a_2 y a_3 son parámetros de los tanques; y r_1 y r_2 es la cantidad de lluvia que reciben los tanques fuente 1 y 2, respectivamente.

Resumiendo, el problema de los tanques puede ser abordado con el par EDM-PDM dado por (15) y (16). A manera de ilustración, la Figura 7 muestra el esquema considerado para controlar el sistema de los tres tanques desde la perspectiva del par EDM-PDM.

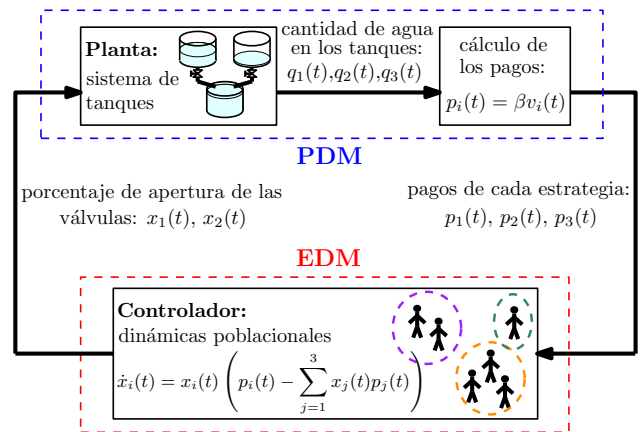


Figura 7: Sistema de tanques controlado con dinámicas poblacionales y modelo dinámico de pagos.

4.3.3. Diseño del Lazo Cerrado

La Figura 8 muestra que el objetivo de control se cumple. Es decir, que el agua en los tanques alcanza el mismo nivel cuando se utiliza el controlador basado en *replicator dynamics*. Para las simulaciones se utilizaron dos constantes de proporcionalidad $\beta = 0.01$ (línea continua) y $\beta = 0.1$ (línea punteada). Además, los parámetros $a_1 = a_2 = 0,1 \text{ s}^{-1}$, $a_3 = 0,05 \text{ s}^{-1}$, $r_1 = 0,5 \text{ m}^3/\text{s}$ y $r_2 = 0,25 \text{ m}^3/\text{s}$. Los resultados muestran que, independientemente de la constante de proporcionalidad empleada, el sistema de tres tanques controlado mediante dinámicas poblacionales actúa de acuerdo con la lógica utilizada en la etapa de diseño. Por ejemplo, si observamos la línea continua de la Figura 8(a), podemos notar que, desde el minuto 0 hasta el minuto 1, el tanque fuente 1 contiene la mayor cantidad de agua. Como se muestra en la Figura 8(b), esto aumenta la función de pago relacionada con abrir la válvula 1, lo que hace que el porcentaje de apertura de dicha válvula aumente durante el primer minuto de la simulación. Además, el tanque receptor tiene la mayor cantidad de carga desde el minuto 1 hasta el minuto 2, como se muestra en la Figura 8(a), por lo que es más rentable para los jugadores no abrir ninguna de las válvulas. De hecho, el porcentaje de apertura de las válvulas 1 y 2 disminuye a partir del minuto 1 de la simulación. El objetivo de control se logra por-

que el comportamiento descrito hace que los tanques tiendan a alcanzar el mismo volumen de agua.

Por otro lado, observe que la constante de proporcionalidad, β , influye en el comportamiento transitorio del sistema en lazo cerrado. Para la constante de proporcionalidad más alta (que en nuestro caso es $\beta = 0.1$ (línea punteada)), la planta se estabiliza más rápidamente a costa de mayores oscilaciones en las señales de control. Por lo tanto, notamos un compromiso entre desempeño y el tener acciones de control más favorables para los actuadores (al usar una constante de proporcionalidad menor, $\beta = 0.01$ en nuestro caso, tenemos señales de control más suaves). Esto hace que los actuadores se vean sometidos a menor estrés).

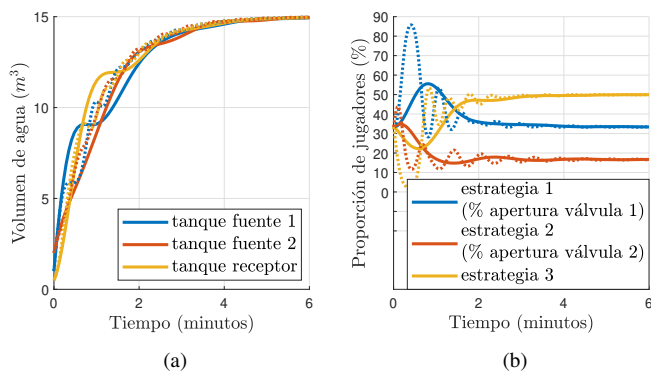


Figura 8: Simulación del sistema de tanques controlado con dinámicas poblacionales. (a) Volúmenes de agua en cada tanque. (b) Distribución de la población entre las estrategias del juego. Con línea continua se muestra el comportamiento para $\beta = 0,01$ y con línea punteada para $\beta = 0,1$.

Las dos principales ventajas de la estrategia de control basada en juegos poblacionales son las siguientes: no requiere modelos (note que para plantear el EDM que define la ley de control no se utiliza ningún modelo que describa la dinámica del agua alojada en los tanques), y, como se explica en Obando et al. (2016), existe una teoría sólida que garantiza que el objetivo de control se alcanza. Aunque no es posible abordar los detalles técnicos de la demostración de estabilidad del sistema controlado con *replicator dynamics* debido a limitaciones de espacio⁴, la idea de la prueba reside en verificar que el sistema multi-tanque es pasivo y luego aplicar el resultado del Teorema 4.2.

4.3.4. Comparación con un controlador difuso

Para establecer una línea base que permita comparar el desempeño del método basado en *replicator dynamics*, diseñamos una estrategia de control utilizando lógica difusa. La elección de esta técnica se fundamenta en su aplicación exitosa en problemas similares, tal es el caso de Li (2020).

El controlador difuso que gobierna la válvula asociada al tanque fuente 1 tiene dos entradas. La primera es la diferencia entre el volumen de agua alojado en el tanque fuente 1 y el tanque fuente 2, que denotamos como dif_f . La segunda corresponde a la diferencia entre el volumen del tanque fuente 1

y el tanque receptor, que llamamos dif_r . La salida de control es el porcentaje de apertura de la válvula 1, denotado como $valv_1$.

Ambas entradas comparten las mismas funciones de membresía: muy_inferior (MI), inferior (I), equilibrada (E), superior (S), y muy_superior (MS). Estas se detallan en la Figura 9(a). Por otro lado, las funciones de membresía de la salida: 0%, 25%, 50%, 75%, y 100%, se presentan en la Figura 9(b).

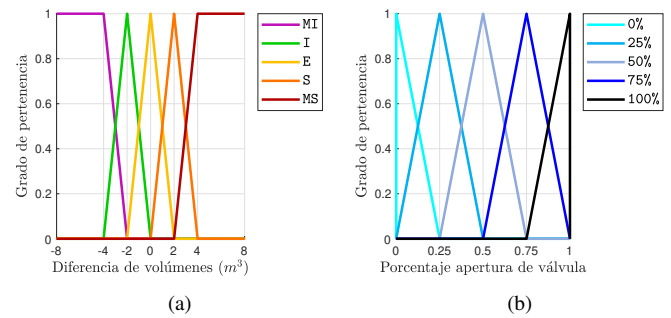


Figura 9: (a) Funciones de membresía de las entradas del controlador difuso dif_f y dif_r . (b) Funciones de membresía de la salida $valv_1$.

Observe, por ejemplo, que si $diferencia_receptor$ está alrededor de $-5 m^3$ es porque el volumen del tanque fuente 1 está muy por debajo del volumen del tanque receptor (tiene $5 m^3$ menos agua). Entonces, la función de membresía con mayor grado de pertenencia es la llamada MI. Por otro lado, si la diferencia de volúmenes entre los dos tanques fuente es cercana a 0, la función de membresía con el grado de pertenencia más alto para la variable dif_f es la nombrada E. Teniendo esto en mente, y con el objetivo de nivelar los volúmenes de los tres tanques de agua, definimos las reglas mostradas en la Tabla 3. Finalmente, el método de defusificación elegido es el del centro de gravedad⁵. Cabe mencionar que el diseño del controlador de la válvula del tanque fuente 2 sigue exactamente los mismos principios.

Tabla 3: Reglas para la salida $valv_1$.

$dif_f \backslash dif_r$	MI	I	E	S	MS
MI	100 %	100 %	75 %	50 %	25 %
I	100 %	75 %	50 %	25 %	0 %
E	100 %	75 %	50 %	25 %	0 %
S	100 %	75 %	50 %	25 %	0 %
MS	75 %	50 %	25 %	0 %	0 %

La comparación entre la técnica que sigue el paradigma descrito en este artículo con el controlador difuso la realizamos en 100 escenarios en los que modificamos tanto el nivel inicial de los tanques como las condiciones de lluvia. El nivel de cada uno de los tanques es una variable aleatoria $\sim U[0, 5] m^3$, y la cantidad de lluvia que capta cada tanque fuente es otra variable aleatoria $\sim U[0.3, 0.7] m^3/s$. Además, los parámetros de los tanques son $a_1 = a_2 = a_3 = 0,1 s^{-1}$.

⁴Una demostración formal de estabilidad para este problema está reportada en Obando et al. (2022).

⁵La elección de los parámetros del controlador difuso se hizo usando prueba y error.

Dado que el problema tratado no establece volúmenes de referencia específicos sino que busca que los tanques alcancen un mismo nivel, hemos decidido emplear el criterio de desempeño dado en Clearwater et al. (1996). En dicho artículo, los autores miden el rendimiento de los controladores usando la desviación estándar de las señales analizadas. En nuestro caso, estas señales corresponden a los volúmenes de los tanques. Claramente, una menor desviación estándar implica que los volúmenes están más cerca unos de otros.

La Figura 10 muestra la evolución de la desviación estándar del nivel de los tanques para el controlador basado en *replicator dynamics* y para el controlador difuso. La línea sólida muestra el comportamiento del promedio de la desviación estándar para los 100 escenarios simulados, mientras la línea discontinua muestra el intervalo de confianza.

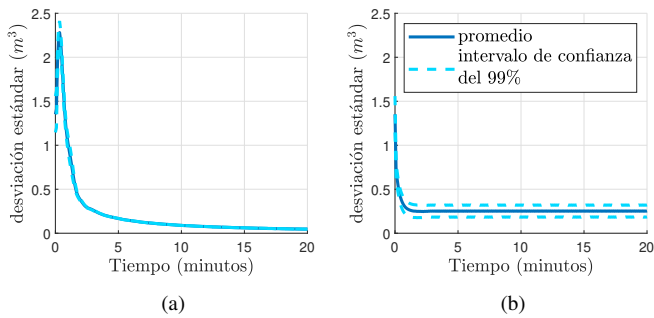


Figura 10: Evolución de la desviación estándar de los volúmenes de los tanques para el sistema controlado con (a) *replicator dynamics* y (b) lógica difusa.

Se puede observar que las dos técnicas logran disminuir la desviación estándar. Sin embargo, a diferencia del controlador difuso, el controlador basado en juegos poblacionales reduce la desviación estándar a cero a medida que transcurre el tiempo.

Por otro lado, puede notarse que el comportamiento de la desviación estándar varía menos de un escenario a otro al usar el controlador basado en *replicator dynamics* (puesto que el intervalo de confianza es más estrecho para dicho método). Esto indica que el controlador en cuestión tiene un comportamiento más robusto ante cambios en los niveles iniciales de los tanques y en los parámetros de lluvia.

4.4. Aplicación de los Replicator Dynamics al Control de Temperatura

El control de temperatura en edificaciones es un tema que ha sido abordado usando distintas estrategias. Por ejemplo, control clásico (PID), técnicas de horizonte móvil, y estrategias no convencionales (redes neuronales, lógica difusa). Entre los problemas tratados por las distintas técnicas, sobresale uno: los actuadores (que en este caso son sistemas de calefacción o aire acondicionado) pueden trabajar sólo hasta una potencia nominal. Esto se traduce en señales de control que están restringidas a no sobrepasar dicho valor. Ante esta situación, muchas de las estrategias de control empleadas se comportan de forma inadecuada. Los controladores PID presentan un comportamiento alejado del óptimo cuando las acciones de control están acopladas (como sucede en ciertas arquitecturas de sistemas de calefacción y aire acondicionado Ygge and Akkermans (1999)), puesto que realizan una distribución ineficiente de la potencia

disponible. Las técnicas de horizonte móvil, como el control predictivo basado en modelo (MPC, del inglés *model predictive control*), demandan una alta carga computacional. Finalmente, muchas de las metodologías no convencionales carecen de un sustento teórico que garantice la consecución del objetivo de control y además, en muchas de éstas se establecen condiciones de saturación en los controladores acarreado transitorios fuertes que pueden ser inconvenientes para los actuadores.

Con el fin de solucionar estos inconvenientes, nuestro objetivo es abordar el control de temperatura en edificaciones como un problema de asignación dinámica de recursos. Para ello, describiremos la implementación de un controlador basado en *replicator dynamics*.

4.4.1. Modelo Térmico de una Edificación

En primer lugar, establezcamos el modelo que describe la temperatura dentro un edificación compuesta por un cierto número de salones. Sea $q_i(t)$ la temperatura del i -ésimo salón en el instante t . Esta temperatura puede modelarse de la siguiente manera (Obando et al., 2013):

$$c_i \dot{q}_i(t) = x_i(t) + e_i(t) + \alpha_i (q_a(t) - q_i(t)) + \sum_{j \in \mathcal{V}_i} \sigma_{ij} (q_j(t) - q_i(t)), \tag{17}$$

donde c_i es la capacitancia térmica del salón; \mathcal{V}_i es el conjunto de salones adyacentes al i -ésimo; σ_{ij} es la conductancia térmica de la pared que separa el i -ésimo del j -ésimo salón (note que $\sigma_{ij} = \sigma_{ji}$); α_i es la conductancia térmica de las paredes que separan el i -ésimo salón del ambiente exterior; $q_a(t)$ es la temperatura ambiente; $x_i(t)$ es la señal proveniente de un actuador (por ejemplo, un aire acondicionado o calefactor); y $e_i(t)$ es una entrada de calor estocástica, debida por ejemplo a ventanas abiertas, personas o equipos presentes dentro del salón.

El desarrollo que se describe a continuación está basado en un modelo para cuatro salones dispuestos tal y como se muestra en la Figura 11. Asumiremos que cada salón cuenta con un calefactor y un sensor de temperatura. Además, supondremos que la potencia nominal del sistema de calefacción es de 10 kW. Es decir, debe satisfacerse que $\sum_{i=1}^4 x_i(t) \leq 10 \text{ kW}$, para todo $t \geq 0$.

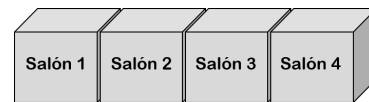


Figura 11: Configuración de los 4 salones que conforman la edificación considerada.

4.4.2. Control de Temperatura

Quando se habla de temperatura, hay variedad de preferencias. Por este motivo, es conveniente que en cada uno de los salones que conforman la edificación pueda fijarse una referencia térmica independiente. El objetivo de control será entonces mantener la temperatura de todas los salones en niveles cercanos a sus referencias establecidas.

Como se dijo antes, nos enfrentamos a un problema de asignación dinámica de recursos. Este tipo de problemas surge cuando los recursos con los que contamos son limitados (en nuestro caso, el recurso es la potencia de calefacción que, como vimos, está limitada a no superar una cierta potencia nominal) y se desea establecer una política de distribución de los mismos

para un número determinado de zonas (para el caso del control de temperatura en edificaciones, las zonas son los salones). Ahora, si pensamos en los recursos como los individuos de una población y las zonas como las estrategias disponibles, es natural suponer que un EDM puede cumplir el rol de la política de distribución requerida.

Siguiendo este orden de ideas, vamos a ajustar el par EDM-PDM al problema de control de temperatura en edificaciones. Para tal fin, es conveniente hacer la siguiente analogía: en el caso estudiado, los jugadores serán la potencia de calefacción disponible (por lo tanto la masa poblacional es $m = 10$ kW) y las estrategias serán los 4 salones (esto es, $\mathcal{S} = \{1, 2, 3, 4\}$). Además, con el propósito de que la calefacción sea asignada en mayor medida a los salones más fríos, debemos diseñar un PDM de tal forma que las funciones de pago asociadas a los salones con menor temperatura sean las más rentables. Bajo esta premisa, establecemos que el pago recibido por los jugadores que elijan el i -ésimo salón sea igual a la diferencia entre la referencia del i -ésimo salón, que denotaremos como \hat{q}_i , y su temperatura, es decir, $p_i(t) = \hat{q}_i - q_i(t)$. Observe que $p_i(t)$ será mayor (y por tanto, más atractivo para los jugadores) cuanto más frío esté el i -ésimo salón. Finalmente, siguiendo la analogía que hemos establecido, la cantidad de jugadores que ha elegido la i -ésima estrategia corresponde a la calefacción asignada al i -ésimo salón, esto es, $x_i(t)$.

Una vez definidos todos los ingredientes del juego poblacional que modela los detalles del problema de control de temperatura en una edificación, debemos elegir un EDM que gobierne la evolución de las poblaciones que juegan cada estrategia. Es decir, debemos definir la estrategia que determinará la distribución de las potencias de calefacción entre los diferentes salones. En nuestro caso, seleccionamos los *replicator dynamics* (13).

4.4.3. Limitación de la Potencia del Controlador Basado en Replicator Dynamics

Dada la analogía que hemos planteado para formular el EDM, la invarianza del simplex Δ (que, como se expuso en el Teorema 2.1, es una de las principales propiedades del todo EDM, incluyendo los *replicator dynamics*) implica que la suma de las potencias suministrada por todos los calefactores es $m = 10$ kW, y se preserva en el tiempo. Esta característica no es deseable en los sistemas de control de temperatura. Por ejemplo, si la temperatura ambiente sube en el transcurso del día, será necesario decrementar la potencia total de calefacción para mantener la temperatura de los salones en las referencias establecidas. Entonces, en lugar de requerir que la potencia suministrada por los actuadores sea constante, lo que de hecho necesitamos es que la potencia total de calefacción no sobrepase un determinado valor (su valor nominal). Para este propósito, se introduce un *salón ficticio* en el modelo, el cual denotamos como estrategia 5. Cabe resaltar que dicho salón no corresponde a ninguna habitación del edificio. Observe que si agregamos este *salón ficticio* al modelo de los *replicator dynamics*, la invarianza del simplex Δ se sigue satisfaciendo, pero ahora para 5 estrategias, es decir, $\sum_{i=1}^5 x_i(t) = m$, $\forall t \geq 0$, donde $x_5(t) \geq 0$ es la potencia "asignada" al *salón ficticio*. En realidad, $x_5(t)$ es simplemente una potencia que no se suministra a ninguna de las habitaciones de la edificación. La potencia real (potencia asignada a los salones) está dada por la expresión $\sum_{i=1}^4 x_i(t)$, que es

claramente menor o igual a la potencia nominal del sistema de calefacción, m .

Así, con la adición del *salón ficticio*, tenemos la limitación de potencia requerida:

$$\sum_{i=1}^4 x_i(t) \leq m, \quad x_i(t) \geq 0, \quad \forall t \geq 0. \quad (18)$$

4.4.4. Simulaciones del Controlador Basado en Replicator Dynamics

Para evaluar el desempeño del controlador basado en *replicator dynamics*, se simula la respuesta del sistema en lazo cerrado bajo las condiciones descritas en la Tabla 4.

Tabla 4: Parámetros térmicos del sistema de la Figura 11.

Parámetro	Valor
Referencias de los salones	22 °C, 21 °C, 20 °C y 19 °C
$q_a(t)$	1 °C
$e_i(t)$	0 W
c_i	$7,476 \times 10^4$ J/°C
α_i	123.6 W/°C
σ_{ij}	123.6 W/°C

Los resultados de la simulación se presentan en la Figura 12. Recordemos que la potencia nominal del sistema de calefacción es de 10 kW. En la misma figura, se muestra también la potencia total usada por el controlador. Se observa que con *replicator dynamics* es posible alcanzar la temperatura de referencia en todos los salones aún bajo la restricción de la potencia nominal. En la misma simulación, hemos introducido una perturbación alrededor de los 200 minutos, que corresponde a una pérdida de calor de 200 W en el salón 4 (por ejemplo, debida a la apertura de una ventana). Se puede notar que el sistema responde rápidamente a la perturbación y logra una recuperación satisfactoria.

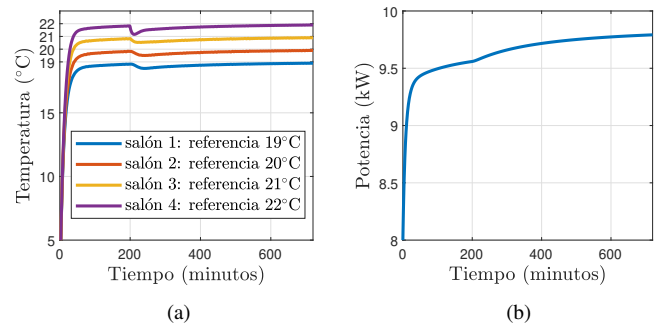


Figura 12: (a) Respuesta del sistema controlado usando los *replicator dynamics*. (b) Potencia usada por el controlador.

En conclusión, el controlador basado en *replicator dynamics* hace una distribución eficiente de la calefacción, puesto que admite la inclusión de restricciones en la potencia suministrada.

4.4.5. Comparación con un controlador MPC

Con el fin de comparar la respuesta obtenida, implementamos un controlador MPC lineal, cuya formulación limita la potencia de los actuadores. Los parámetros del controlador MPC empleado son: tiempo de muestreo de 1 minuto, horizontes de

predicción y control de 10 minutos. Además, la función de costo del problema de optimización que se resuelve en cada iteración del controlador es cuadrática y penaliza el error entre la temperatura de cada salón y su referencia (no se considera penalización alguna para las señales de control). Finalmente, se incluye la limitación de potencia dada en (18) dentro de las restricciones del problema de optimización mencionado.

La Figura 13 muestra el desempeño del controlador basado en *replicator dynamics* y del MPC. En este caso, el escenario de simulación tiene los mismos parámetros que hemos usado anteriormente, a excepción de la potencia disponible, que se reduce a 8 kW. Dicha reducción permite evaluar el rendimiento de los controladores cuando es imposible alcanzar las referencias de temperatura establecidas en cada salón.

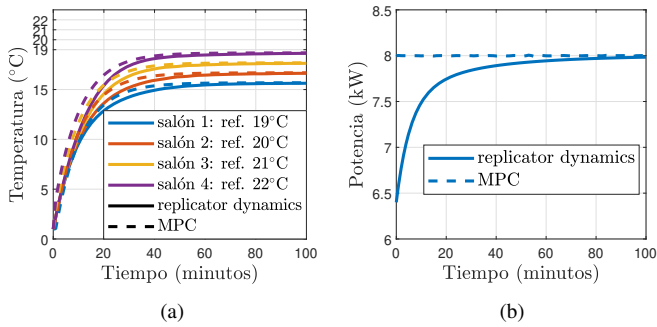


Figura 13: Respuesta del sistema controlado con *replicator dynamics* (línea continua) y MPC: (a) temperatura de los salones; (b) potencia total usada por el controlador.

Observe que los dos controladores alcanzan el mismo estado estable, es decir, asignan la misma potencia de calefacción a cada salón. Este resultado es notable puesto que los *replicator dynamics* emplean muchos menos recursos computacionales en comparación al MPC (que necesita resolver un problema de optimización en cada iteración, mientras los *replicator dynamics* sólo requieren la solución de un conjunto de ecuaciones diferenciales de primer orden). Para sustentar esta afirmación, proponemos nuevos escenarios de simulación en los que variamos el número de salones que conforman la edificación desde 10 hasta 100. En todos los casos, los salones están dispuestos uno al lado del otro respetando la configuración mostrada en la Figura 11. Cada escenario se simula 10 veces, eligiendo las referencias de los salones aleatoriamente del siguiente conjunto {19 °C, 20 °C, 21 °C, 22 °C}. El tiempo de simulación es de 1 hora. Además, la potencia de calefacción nominal se asume igual $2n$ kW, donde n es el número de salones. Todos los demás parámetros del experimento son los que se muestran en la Tabla 4. Finalmente, los índices de desempeño escogidos para la comparación son el promedio de la integral del error absoluto (IAE, del inglés *integral absolute error*), cuyo comportamiento se presenta en la Tabla 5; y el promedio del tiempo que tardan las simulaciones, que se muestra en la Figura 14. Cabe señalar que los experimentos se ejecutaron usando MATLAB en un computador equipado con un procesador Intel Core i7–8700K de 3.7 GHz y una RAM de 32 GB. Para resolver las ecuaciones diferenciales que modelan tanto las temperaturas de los salones como los *replicator dynamics* se usa el método de Runge–Kutta con un tiempo de muestreo fijo de 2 s. Además, para la solución

del problema de optimización asociado al controlador MPC se emplea programación cuadrática con un método de punto interior.

Tabla 5: Comportamiento del IAE para *replicator dynamics* y MPC.

número de salones	<i>replicator dynamics</i> IAE (°C·s)	MPC IAE (°C·s)	número de salones	<i>replicator dynamics</i> IAE (°C·s)	MPC IAE (°C·s)
10	84632	84588	60	194750	194662
20	120484	120428	70	209717	209596
30	140017	139929	80	230080	230003
40	160151	160078	90	239468	239353
50	176491	176419	100	252738	252602

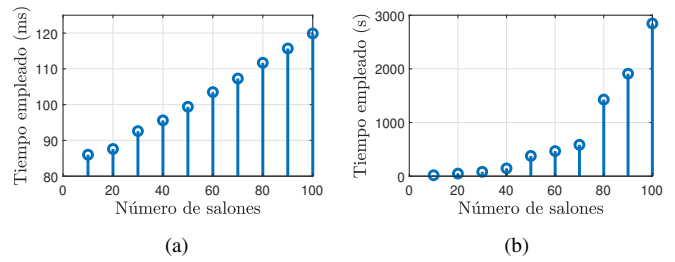


Figura 14: Promedio del tiempo empleado en las simulaciones para (a) *replicator dynamics* y (b) MPC.

En primer lugar, encontramos que el IAE del controlador MPC es menor, aunque la diferencia con *replicator dynamics* está lejos de ser crítica. En porcentaje, la diferencia del IAE está entre 0.04 % y 0.05 % a favor de MPC. En el peor de los casos (cuando el número de salones es 100), la diferencia de IAE's es de 137 °C·s durante el tiempo de simulación, que es de 1 hora. Esto implica que, durante esa hora, las temperaturas de los salones controlados con *replicator dynamics* están, en promedio, 0.04 °C más alejadas de sus referencias que al usar MPC.

En cuanto al tiempo empleado en las simulaciones, confirmamos que *replicator dynamics* es más eficiente. Por un lado, en promedio, *replicator dynamics* se simuló 7000 veces más rápido que MPC. Por otro lado, *replicator dynamics* escala mejor, ya que el crecimiento del tiempo de las simulaciones respecto al número de salones es lineal, mientras que para MPC es exponencial.

5. Optimización

El par EDM–PDM no solo puede emplearse para abordar problemas de control. Este paradigma también resulta útil para solucionar problemas de optimización. Por ejemplo, en Pantoja et al. (2019) los autores abordan el problema de despacho económico en redes eléctricas, o en Tembine et al. (2010) los autores asignan de forma óptima los canales de un sistema de comunicación. En este caso, los PDMs pueden ser tanto estáticos (o *memoryless*, si seguimos la denominación empleada en la teoría de pasividad) como dinámicos. Aquí el término “estático” puede llegar a generar confusiones. Cuando decimos PDM estático no hacemos referencia a que los pagos son constantes en el tiempo, sino a que no corresponden a la salida de un sistema dinámico (en contraste con las aplicaciones en control,

donde los pagos dependen de, por ejemplo, los niveles de agua en tanques tal y como se expuso en el caso de estudio de la Sección 4.3.2).

Para facilitar la presentación, nos enfocaremos en describir el uso del par EDM–PDM en la solución de problemas de optimización donde el PDM es estático. En este caso, los PDMs son funciones exclusivamente de los estados poblacionales, implicando que $\mathbf{W}(\mathbf{q}(t), \mathbf{x}(t)) = 0$ en (6). Sin embargo, al final de la sección, presentaremos también un ejemplo en el que se requiere un PDM dinámico para modelar la competencia en mercados energéticos. Por otro lado, centraremos nuestra discusión en las *Smith Dynamics*, una clase de EDM con propiedades similares a las de los *replicator dynamics*, pero con una mejor tasa de convergencia (Barreiro-Gomez et al., 2017a).

5.1. Smith Dynamics

Propuestas en 1984 por M. Smith para modelar las decisiones que toman los conductores en sistemas de tráfico urbano (Smith, 1984), las *Smith Dynamics* son otra EDM que ha sido objeto de un gran número de estudios. Particularmente, algunos artículos han explotado las propiedades que tienen estas dinámicas para resolver problemas de optimización (Mandal and Tallapragada, 2022; Chow et al., 2019; Zhu et al., 2012). Antes de enunciar y analizar dichas propiedades, describamos la obtención de las *Smith Dynamics* a partir de la metodología expuesta en la Sección 2.1.

Al reemplazar el *protocolo de comparación por parejas* dado en la Tabla 1 en las *mean dynamics*, obtenemos las *Smith dynamics*. Si consideramos la k -ésima población, estas dinámicas están dadas por el sistema de ecuaciones diferenciales que se enuncia a continuación: Al reemplazar el *protocolo de comparación por parejas* dado en la Tabla 1 en las *mean dynamics*, obtenemos las *Smith dynamics*. Si consideramos la k -ésima población, estas dinámicas están dadas por

$$\dot{x}_i^k(t) = \sum_{j \in S} x_j^k(t) [p_i^k(t) - p_j^k(t)]_+ - x_i^k(t) \sum_{j \in S} [p_j^k(t) - p_i^k(t)]_+, \tag{19}$$

donde $i \in S^k$. Las *Smith dynamics* tienen las propiedades enunciadas a continuación, las cuales están estrechamente ligadas con el planteamiento y solución de problemas de optimización.

Teorema 5.1. Descenso: sea $V : \Delta^k \mapsto \mathbb{R}$ una función C^1 , $\mathbf{x}^k(t)$ una solución de (19), y $t_2 \geq t_1$. Si elegimos el vector de funciones de pago como $\mathbf{p}^k(t) = -\nabla V(\mathbf{x}^k(t))$, entonces $V(\mathbf{x}^k(t_2)) \leq V(\mathbf{x}^k(t_1))$.

Demostración. Sea $\dot{V}(\mathbf{x}^k(t))$ la derivada de V a lo largo de $\mathbf{x}^k(t)$. Basta con demostrar que si $\mathbf{x}^k(t)$ es una solución de las *Smith dynamics*, entonces $\dot{V}(\mathbf{x}^k(t)) \leq 0$.

Supongamos que $\mathbf{x}^k(t)$ es solución de (19) y que $\mathbf{p}^k(t) = -\nabla V(\mathbf{x}^k(t))$. Note que

$$\begin{aligned} \dot{V}(\mathbf{x}^k(t)) &= \nabla^T V(\mathbf{x}^k(t)) \dot{\mathbf{x}}^k(t) \\ &= -\mathbf{p}^{k,T}(t) \dot{\mathbf{x}}^k(t) \\ &= -\sum_{i \in S^k} \sum_{j \in S^k} p_i^k(t) x_j^k(t) [p_i^k(t) - p_j^k(t)]_+ \\ &\quad + \sum_{i \in S^k} \sum_{j \in S^k} p_i^k(t) x_i^k(t) [p_j^k(t) - p_i^k(t)]_+. \end{aligned} \tag{20}$$

Usando la función signo, $\text{sgn}()$, y definiendo

$$\begin{aligned} \xi_{ij}(t) &:= (1 + \text{sgn}(p_i^k(t) - p_j^k(t))) x_j^k(t) \\ &\quad + (1 - \text{sgn}(p_i^k(t) - p_j^k(t))) x_i^k(t), \end{aligned} \tag{21}$$

podemos reescribir la expresión en (20) como

$$\begin{aligned} \dot{V}(\mathbf{x}^k(t)) &= -\frac{1}{2} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij}(t) p_i^k(t) (p_i^k(t) - p_j^k(t)) \\ &= -\frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij}(t) (p_i^{k2}(t) - p_j^k(t) p_i^k(t)) \\ &\quad - \frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij}(t) (p_i^{k2}(t) - p_j^k(t) p_i^k(t)). \end{aligned} \tag{22}$$

Note que $\xi_{ij}(t) \geq 0$ puesto que $\mathbf{x}^k(t) \in \Delta^k$.

Finalmente, intercambiando los índices y el orden de las dos últimas sumatorias en (22), tenemos

$$\begin{aligned} \dot{V}(\mathbf{x}^k(t)) &= -\frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij}(t) (p_i^{k2}(t) - p_j^k(t) p_i^k(t)) \\ &\quad - \frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ji}(t) (p_j^{k2}(t) - p_j^k(t) p_i^k(t)) \\ &= -\frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij}(t) (p_i^{k2}(t) - 2p_j^k(t) p_i^k(t) + p_j^{k2}(t)), \end{aligned} \tag{23}$$

puesto que $\xi_{ij}(t) = \xi_{ji}(t)$. En conclusión, $\dot{V}(\mathbf{x}^k(t)) = -\frac{1}{4} \sum_{i \in S^k} \sum_{j \in S^k} \xi_{ij} (p_i^k(t) - p_j^k(t))^2 \leq 0$, quedando demostrado el teorema. \square

Corolario 5.1. Satisfacción de Restricciones : sea m^k un real positivo y $\mathbf{x}^k(t)$ una solución de (19). Suponga que $\sum_{i=1}^{n^k} x_i^k(0) = m^k$ y $x_i^k(0) \geq 0$, para todo $i = 1, \dots, n^k$. Entonces, para todo tiempo $t \geq 0$, $\sum_{i=1}^{n^k} x_i^k(t) = m^k$ y $x_i^k(t) \geq 0$, para todo $i = 1, \dots, n^k$.

Demostración. La prueba sigue inmediatamente del resultado dado en el Teorema 2.1 y del hecho de que las *Smith Dynamics* son un caso particular de las *mean dynamics*. \square

Observe que los resultados anteriores tienen las siguientes implicaciones. Por un lado, el Teorema 5.1 establece la posibilidad de diseñar funciones de pago que, al ser usadas en las *Smith dynamics*, generan trayectorias que reducen el valor de una función objetivo dada. Por otro lado, el Corolario 5.1 garantiza que las soluciones de las *Smith dynamics* satisfacen ciertas restricciones. En conclusión, las *Smith dynamics* pueden emplearse como algoritmo de solución del siguiente problema de optimización, que en la literatura se conoce como de distribución de recursos:

$$\begin{aligned} \text{mín } V(\mathbf{z}) \\ \text{sujeto a: } \sum_{i=1}^n z_i &= Z \\ z_i &\geq 0, \quad i = 1, \dots, n, \end{aligned} \tag{24}$$

donde $Z \in \mathbb{R}_{>0}$ generalmente modela el recurso que se va a distribuir entre n zonas, z_i es el recurso alojado en la i -ésima zona, y $\mathbf{z} = [z_1, \dots, z_n]^T$. Esta propiedad se escribe formalmente en el Teorema 5.2.

Teorema 5.2. Algoritmo de Optimización: considere el problema de distribución de recursos (24). Asuma que V es estrictamente convexa y que \mathbf{z}^* es la solución óptima de dicho problema. Además, suponga que $\mathbf{x}^k(t)$ es la solución de las Smith dynamics dadas en (19) con los siguientes parámetros: $\mathbf{p}^k(t) = -\nabla V(\mathbf{x}^k(t))$; $\mathcal{S}^k = \{1, \dots, n\}$; $m^k = \mathbf{Z}$; $\sum_{i \in \mathcal{S}} x_i^k(0) = m^k$; y $x_i^k(0) > 0$, para todo $i \in \mathcal{S}$. Se cumple que $\lim_{t \rightarrow \infty} \mathbf{x}^k(t) = \mathbf{z}^*$.

Demostración. Note que, bajo las suposiciones hechas, el Corolario 5.1 garantiza que $\mathbf{x}^k(t)$ satisface las restricciones de (24) para todo $t \geq 0$. Además, el Teorema 5.1 certifica que la función objetivo $V(\mathbf{x}^k(t))$ decrece a medida que transcurre el tiempo. Aún así, no tenemos garantía de que $\mathbf{x}^k(t)$ converja a la solución óptima de (24). Sin embargo, si usamos $V(\mathbf{x}^k(t)) - V(\mathbf{z}^*)$ como función candidata de Lyapunov, y siguiendo el razonamiento mostrado en la prueba del Teorema 5.1, podemos concluir que \mathbf{z}^* es asintóticamente estable bajo las Smith dynamics. Esto concluye la prueba. \square

Con el fin de ilustrar el uso de las Smith dynamics para resolver problemas de optimización, a continuación presentamos dos aplicaciones ligadas a sistemas eléctricos.

5.2. Despacho Económico de Generadores Eléctricos

En el problema de despacho económico se asume un conjunto de n generadores eléctricos que tienen que satisfacer una demanda D . Cada generador produce energía a un determinado costo que depende tanto de la cantidad de potencia generada como de la eficiencia de la unidad. El objetivo del despacho económico es determinar qué proporción de la demanda D será cubierta por cada generador de tal forma que el costo global se minimice.

Para formular este problema matemáticamente, es conveniente asociar a los generadores funciones de costo que modelen el gasto de recursos en el que se debe incurrir para producir cierta cantidad de potencia. En nuestro caso, usaremos funciones cuadráticas puesto que son las más empleadas en la literatura (Pantoja et al., 2019), esto es,

$$c_i(g_i) = \alpha_i g_i^2 + \beta_i g_i + \gamma_i, \quad (25)$$

donde $g_i \geq 0$ es la potencia producida por el i -ésimo generador; $c_i(g_i)$ es el costo de generación; y α_i, β_i y γ_i son parámetros que dependen de la eficiencia de la unidad y de su tecnología.

Dado que, como se dijo, el despacho económico contempla la reducción del costo total de generación, el objetivo es minimizar la función $\sum_{i=1}^n c_i(g_i)$. Además, como el total de la demanda debe ser cubierta por los generadores, entonces tenemos que $\sum_{i=1}^n g_i = G$. En resumen, el problema de despacho económico es el siguiente:

$$\begin{aligned} & \min_{g_1, \dots, g_n} \sum_{i=1}^n c_i(g_i) \\ \text{sujeto a: } & \sum_{i=1}^n g_i = G \\ & g_i \geq 0, \quad i = 1, \dots, n. \end{aligned} \quad (26)$$

Note que este problema tiene la misma forma del planteado en (24). Adicionalmente, la función objetivo es convexa, por

lo tanto, puede solucionarse usando las Smith dynamics con las instrucciones dadas en el Teorema 5.2. Al hacerlo, emerge la siguiente analogía: en primer lugar, dado que se trata de un único problema de optimización, necesitaremos una sola población, es decir, $\mathcal{P} = \{1\}$, y por lo tanto $k = 1$. El tamaño de dicha población será igual a la demanda que debe suplirse, esto es, $m^1 = G$. Además, el conjunto de estrategias corresponderá al conjunto de generadores disponibles, es decir, $\mathcal{S}^1 = \{1, \dots, n\}$. Así las cosas, la población que está jugando la i -ésima estrategia representará la porción de demanda que debe suplir el i -ésimo generador, esto es, $x_i^1(t) = g_i$. Finalmente, las funciones de pago serán los costos marginales de cada uno de los generadores. En particular, para el i -ésimo generador tenemos que $p_i^1(t) = \frac{dc_i(x_i^1(t))}{dx_i^1}$.

Para ilustrar el desempeño del algoritmo de optimización basado en las Smith dynamics se plantea un escenario con 50 generadores y una demanda de 10 MW. Los parámetros de las funciones de costo de los generadores se eligen así: $\alpha_i \sim U[10, 14] \text{ \$/kW}^2$; $\beta_i \sim U[-4, 4] \text{ \$/kW}$; y $\gamma_i \sim U[10, 20] \text{ \$}$. Los resultados bajo estas condiciones se ilustran en la Figura 15. Puede observarse que la función de costo decrece constantemente (de acuerdo con el resultado del Teorema 5.1). Además, la demanda es cubierta por los generadores durante todo el tiempo (conforme a lo expuesto en el Corolario 5.1). Finalmente, los costos marginales de los generadores se igualan en estado estable. Esto, obedeciendo las condiciones de Karush–Kuhn–Tucker, confirma que el algoritmo converge al óptimo del problema planteado (como lo certifica el Teorema 5.2).

5.3. Aplicación en Mercados Energéticos

Considere ahora un juego de mercado energético⁶ donde $N \in \mathbb{Z}_{\geq 1}$ jugadores (sistemas de manejo de energía) compiten para comprar energía a lo largo de un horizonte de $T \in \mathbb{Z}_{\geq 1}$ instantes de tiempo. En particular, empleemos $\mathcal{P} = \{1, 2, \dots, N\}$ para denotar el conjunto de jugadores, $\mathcal{T} = \{1, 2, \dots, T\}$ para denotar el horizonte de tiempo a considerar, $\mathcal{S}^k \subseteq \mathcal{T}$ para representar el conjunto de instantes de tiempo en los que el jugador $k \in \mathcal{P}$ compete, y m^k para denotar el total de energía que el jugador k requiere comprar a lo largo del horizonte. Además, se asume que $|\mathcal{S}^k| = n^k \geq 2$, para todo $k \in \mathcal{P}$, que $\mathcal{T} = \cup_{k \in \mathcal{P}} \mathcal{S}^k$ (lo cual implica que $n = \sum_{k \in \mathcal{P}} n^k \geq T$), y que para cada instante de tiempo $i \in \mathcal{T}$ hay un límite superior sobre la energía disponible que se denota $e_i \in \mathbb{R}_{>0}$. Finalmente, se considera que cada sistema de manejo de energía tiene una función de utilidad $\varphi^k : \mathbb{R}^n \rightarrow \mathbb{R}$ que busca maximizar (se asume que $\varphi^k(\cdot)$ es cóncava y doblemente continuamente diferenciable para todo $k \in \mathcal{P}$). En general, el valor de dicha función depende de las decisiones de los demás jugadores (por ejemplo, a través del precio de la energía si este último incrementa con la demanda).

Para modelar el mercado energético descrito como un juego poblacional, se propone la siguiente analogía. Cada jugador (sistema de manejo de energía) se modela como una población con masa m^k y conjunto de estrategias \mathcal{S}^k . De esta manera, x_i^k representa la energía a ser comprada por el jugador $k \in \mathcal{P}$ en

⁶Un juego de mercado energético puede verse como una competencia de Cournot, y varias aplicaciones de control pueden modelarse bajo dicha abstracción (Grammatico, 2017; Wang et al., 2021; Yi and Pavel, 2019).

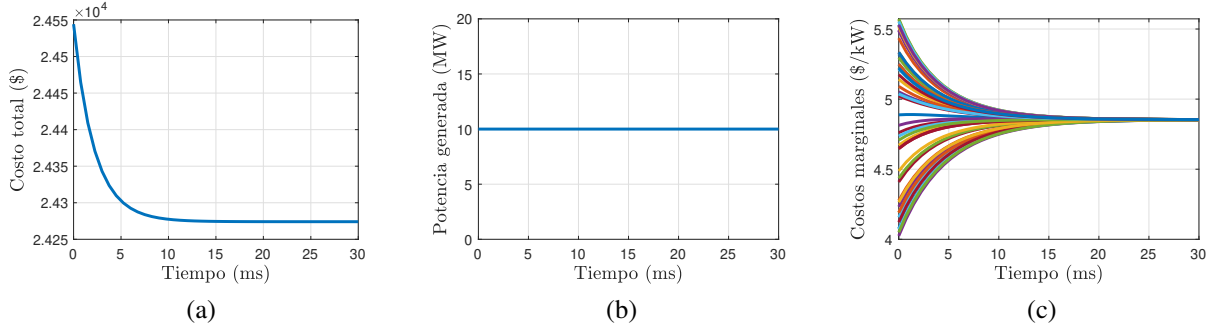


Figura 15: Resultados del despacho económico usando *Smith dynamics*: (a) costo total de la potencia producida; (b) demanda cubierta por los generadores; (c) costos marginales.

el instante de tiempo $i \in \mathcal{S}^k \subseteq \mathcal{T}$. Bajo esta analogía, cualquier distribución estratégica $\mathbf{x} \in \Delta$ garantiza que todos los jugadores (poblaciones) compran exactamente la energía requerida. Por otro lado, las restricciones de energía disponible se pueden modelar como restricciones de la forma $\sum_{k \in \mathcal{P}_i} x_i^k \leq e_i$, para todo $i \in \mathcal{T}$, donde $\mathcal{P}_i = \{k \in \mathcal{P} : i \in \mathcal{S}^k\}$ corresponde al conjunto de jugadores que compiten en el instante de tiempo i . Con esto en mente, la región factible del problema será $\mathcal{X} = \{\mathbf{x} \in \Delta : \sum_{k \in \mathcal{P}_i} x_i^k - e_i \leq 0, \forall i \in \mathcal{T}\}$, y se asume que \mathcal{X} es no vacío. Note que en este caso, el conjunto Δ no es suficiente para caracterizar todas las restricciones del problema.

El objetivo es entonces diseñar un PDM que lleve a la sociedad de agentes a un NE que cumpla las restricciones del problema (esto se conoce como un equilibrio generalizado de Nash). Es decir, llevar a la sociedad a una distribución estratégica $\mathbf{x}^* \in \mathcal{N} \cap \mathcal{X}$ (bajo las suposiciones del problema se cumple que $\mathcal{N} \cap \mathcal{X}$ es un conjunto no vacío y compacto).

Para estudiar el problema, consideramos entonces un EDM caracterizado por las *Smith dynamics* en (19) para todo $i, j \in \mathcal{S}^k$ y todo $k \in \mathcal{P}$, y con $\mathbf{x}(0) \in \Delta$. En este caso $\mathcal{V}(\mathbf{x}(t), \mathbf{p}(t)) = \text{col}(\dot{x}_1^1(t), \dots, \dot{x}_n^N(t))$, y es posible verificar que $\mathcal{V}(\mathbf{x}(t), \mathbf{p}(t)) \in T\Delta(\mathbf{x}(t))$, para todo $t \geq 0$.

Por otro lado, para el PDM se proponen las dinámicas dadas por

$$\begin{aligned} g_j(t) &= \sum_{k \in \mathcal{P}_j} x_j^k(t) - e_j \\ \dot{q}_j(t) &= \max(g_j(t), 0) - q_j(t) \max(-g_j(t), 0) \\ p_i^k(t) &= \frac{\partial \varphi^k(\mathbf{x}(t))}{\partial x_i^k} - q_i(t), \end{aligned} \quad (27)$$

para todo $j \in \mathcal{T}$, todo $i \in \mathcal{S}^k$ y todo $k \in \mathcal{P}$, y con $\mathbf{q}(0) \in \mathbb{R}_{\geq 0}^T$. La intuición detrás del PDM propuesto está basada en ideas de optimización primal-dual (las variables q_j se interpretan como los multiplicadores de Lagrange asociados a las restricciones $g_j \leq 0$). En particular, note que $\dot{q}_j(t) = 0$ si y sólo si $g_j(t) \leq 0$ y $q_j(t)g_j(t) = 0$. Dichos requisitos corresponden a las condiciones de factibilidad del problema primal y de holgura complementaria, respectivamente. Además, es posible verificar que las dinámicas propuestas garantizan que $q_j(t) \geq 0$ para todo $t \geq 0$ (condición de no negatividad de los multiplicadores de Lagrange).

Bajo las suposiciones del mercado energético, el EDM-PDM considerado garantiza la estabilidad asintótica del conjunto $\mathcal{N} \cap \mathcal{X}$. Este resultado está respaldado formalmente por Martínez-Piazuelo et al. (2022d, Teorema 3).

A manera de ilustración, considere el escenario con $N = 20$, $T = 40$, \mathcal{S}^k generado aleatoriamente, $m^k \sim U[3, 4]$, y $e_i \sim U[2, 2.5]$, para todo $k \in \mathcal{P}$ y todo $i \in \mathcal{T}$ (garantizando las suposiciones del problema). Además, para las funciones de utilidad se emplean funciones de la forma considerada en Martínez-Piazuelo et al. (2022d, Sección 6). Por referencia, la instancia considerada en la simulación resulta en un espacio dimensional con $n = 403$. En la Figura 16 se presentan los resultados de simulación para el escenario considerado. Efectivamente, se verifica que la sociedad converge asintóticamente a un NE en el cual se satisfacen todas las restricciones de energía del problema.

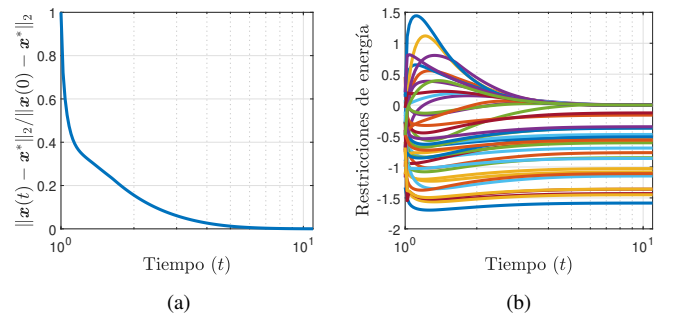


Figura 16: Simulación del mercado energético considerado. (a) Trayectoria del índice de desempeño $\|\mathbf{x}(t) - \mathbf{x}^*\|_2 / \|\mathbf{x}(0) - \mathbf{x}^*\|_2$, donde $\mathbf{x}^* \in \mathcal{NE} \cap \mathcal{X}$. (b) Trayectorias de las restricciones $g_i(t) = \sum_{k \in \mathcal{P}_i} x_i^k(t) - e_i$, para todo $i \in \mathcal{T}$.

A partir de los dos ejemplos de optimización descritos, es importante resaltar que cuando se considera un problema que sólo involucra restricciones tipo simplex que pueden ser caracterizadas por el conjunto Δ (como las que intervienen en el ejemplo del despacho económico), se emplea un PDM estático definido por el gradiente de la función objetivo del problema. Por otro lado, si el problema de optimización considera restricciones adicionales a las capturadas por Δ (como en la aplicación en mercados energéticos), el PDM se modela como un sistema dinámico. En este caso, las variables de estado del PDM representan los multiplicadores de Lagrange asociados a las restricciones adicionales, mientras que la salida corresponde al gradiente del Lagrangiano respectivo.

6. Discusión

Los juegos poblacionales y las dinámicas evolutivas han sido empleados como controladores y optimizadores en diversos campos como las ciencias sociales, la ecología y la ingeniería. Estos enfoques buscan orientar a las poblaciones o sus agentes hacia los equilibrios deseados y promover la cooperación en escenarios como el dilema del prisionero. Se han propuesto controladores de ganancias adaptativas para regular la dinámica de los juegos de la población cambiando de forma adaptativa la matriz de pagos del juego Govaert et al. (2022). Además, el uso de estructuras de comunicación gráfica ha ampliado la aplicabilidad de la dinámica de juego de la población en el diseño de protocolos de aprendizaje y control en circunstancias distribuidas Govaert et al. (2022). Por otra parte, la integración de la retroalimentación ambiental y la consideración de mutaciones en la dinámica replicador-mutador han llevado a la identificación de nuevos comportamientos dinámicos y la exploración de opciones de control Tan et al. (2021). También se ha realizado el análisis de la dinámica evolutiva estocástica del juego, permitiendo una comprensión cuantitativa del comportamiento a largo plazo de estos sistemas Gong et al. (2022). Otros trabajos, como el de Chow et al. (2019) utilizan las dinámicas evolutivas para los juegos de población con un conjunto de estrategias discretas, inspirada en la teoría de juegos de transporte óptimo y de campo medio (*mean field games*), para resolver problemas de transporte óptimo utilizado en aplicaciones de optimización distribuida.

Sin embargo, el paradigma presentado en este artículo tiene características que lo hacen destacar sobre otras técnicas a la hora de abordar problemas de control y optimización. Por un lado, el proceso de diseño es relativamente simple puesto que la analogía entre juegos y problemas de ingeniería es evidente en una gran variedad de casos. Por ejemplo, un problema de optimización puede verse como un juego en el que las variables de decisión están asociadas con estrategias y la función objetivo se expresa en términos de funciones de pago. Además, las dinámicas del sistema de control o variables de decisión adicionales, por ejemplo, los multiplicadores de Lagrange, pueden ser modeladas mediante el formalismo del PDM. Por otro lado, el buen número de resultados formales que existe en torno a las EDMs (por ejemplo, los teoremas presentados en las Secciones 2, 4.2 y 5.1) permite realizar análisis matemáticos rigurosos que garantizan ciertas propiedades clave, como la estabilidad de lazo cerrado en problemas de control o la convergencia al óptimo cuando se habla de optimización, así como la invarianza del simplex Δ a lo largo del tiempo. Cabe anotar que muchos métodos, especialmente aquellos basados en heurísticas, carecen de esta característica, que juega un rol fundamental en problemas sensibles o que requieran de cierto grado de seguridad (Rass et al., 2020). En particular, note que la invarianza de Δ garantizada por el EDM (ver Definición 1) resulta atractiva para aplicaciones de control u optimización dinámica donde las restricciones tipo simplex Δ deben cumplirse en todos los instantes de tiempo, y no sólo de manera asintótica al converger a la solución.

Cuando hablamos de aplicaciones de control, como las expuestas en la Sección 4, una de las ventajas de usar el paradigma basado en EDMs y PDMs es que los controladores resultantes sólo requieren la evaluación de un conjunto de expresiones algebraicas para actualizar sus acciones. Por lo tanto,

en comparación con técnicas basadas en optimización, como MPC, nuestro enfoque disminuye la carga computacional. De hecho, en Obando et al. (2022) se presenta una comparación de los tiempos de simulación de un controlador basado en el par EDM-PDM y un MPC para regular un sistema de drenaje urbano de mediana escala. Los resultados muestran que el MPC necesita un 380 % más de tiempo. Lo anterior es un factor esencial que debe considerarse al diseñar estrategias de control para sistemas con dinámicas rápidas, como las redes eléctricas.

Cabe resaltar que la implementación de los métodos expuestos en este documento conlleva algunos retos que no fueron discutidos por limitaciones de espacio, pero que han sido abordados en otros trabajos. Por ejemplo, la influencia de retardos en la transferencia y procesamiento de la información (Park and Leonard, 2021), la discretización de las EDMs para ser implementadas en computadoras (Martinez-Piazuelo et al., 2022a), o su robustez ante errores inducidos por aproximaciones numéricas (Obando et al., 2016). Finalmente, destacamos que existen múltiples líneas de trabajo futuro. Entre las que podemos mencionar, está el desarrollo de métodos de sintonización de controladores basados en juegos poblacionales (parte de este tema se aborda en Obando et al. (2013)), o la extensión de los algoritmos de optimización discutidos en la Sección 5 para resolver problemas de optimización multicriterio (Mahmoudi et al., 2021; Villanueva et al., 2023).

7. Conclusiones

A través de diversos casos de estudios ampliamente citados en la literatura, este trabajo ha demostrado cómo las dinámicas que modelan la evolución de una población de jugadores pueden emplearse para abordar problemas de ingeniería desde una perspectiva diferente a la tradicional. La teoría de juegos evolutiva se utiliza para crear modelos, políticas de decisión óptima y estrategias de control para una variedad de sistemas. El objetivo es obtener resultados adecuados bajo las limitaciones impuestas por las variables subyacentes de los modelos y sistemas.

La aplicación del paradigma descrito en este documento puede potenciarse al abordar ciertos retos que emergen a la hora de implementar los métodos basados en juegos poblacionales en entornos reales. En este contexto, se propone como trabajo futuro analizar la influencia de los retardos causados por el *hardware* utilizado para procesar la información, así como los errores de aproximación inherentes a los métodos numéricos empleados para calcular tanto las acciones de control como las iteraciones de los algoritmos de optimización.

Agradecimientos

Los autores agradecen al proyecto PID2020-115905RB-C21 (L-BEST) financiado por MCIN/AEI/10.13039/5011000110, al proyecto TED2021-129927B-I00 (MASHED) financiado por "European Union NextGenerationEU/PRTR", y al proyecto BPIN 2021000100499 financiado por el CTel - SGR y MinCiencias, Colombia. Juan Martinez-Piazuelo agradece a la Universitat Politècnica de Catalunya y al Banco Santander por el soporte financiero de su beca predoctoral FPI-UPC.

Referencias

- Ananduta, W., Barreiro-Gomez, J., Ocampo-Martinez, C., Quijano, N., 2018. Mitigation of communication failures in distributed model predictive control strategies. *IET Control Theory & Applications* 12, 2507–2515. doi:10.1049/iet-cta.2018.5044.
- Axelrod, R., 1984. *The evolution of cooperation*. Basic Books.
- Bacci, G., Lasaulce, S., Saad, W., Sanguinetti, L., 2016. Game theory for networks: A tutorial on game-theoretic tools for emerging signal processing applications. *IEEE Signal Processing Magazine* 33, 94–119. doi:10.1109/MSP.2015.2451994.
- Barreiro-Gomez, J., Mas, I., Giribet, J., Moreno, P., Ocampo-Martinez, C., Sanchez-Pena, R., Quijano, N., 2021. Distributed data-driven UAV-formation control via evolutionary games: Experimental results. *Journal of The Franklin Institute* 358, 5334–5352. doi:10.1016/j.jfranklin.2021.05.002.
- Barreiro-Gomez, J., Obando, G., Quijano, N., 2017a. Distributed population dynamics: Optimization and control applications. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47, 304–314. doi:10.1109/TSMC.2016.2523934.
- Barreiro-Gomez, J., Ocampo-Martinez, C., Quijano, N., 2017b. Dynamical tuning for multi-objective model predictive control based on population games. *ISA Transactions* 69, 175–186. doi:10.1016/j.isatra.2017.03.027.
- Bhattacharya, A., 2021. *The man from the future: The visionary life of John von Neumann*. Penguin UK.
- Brown, J.S., Cunningham, J.J., Gatenby, R.A., 2016. Aggregation effects and population-based dynamics as a source of therapy resistance in cancer. *IEEE Transactions on Biomedical Engineering* 64, 512–518. doi:10.1109/TBME.2016.2623564.
- Chow, S.N., Li, W., Lu, J., Zhou, H., 2019. Population games and discrete optimal transport. *Journal of Nonlinear Science* 29, 871–896. doi:10.1007/s00332-018-9507-5.
- Chow, V.T., 1959. *Open-channel hydraulics*. The Blackburn Press.
- Clearwater, S.H., Costanza, R., Dixon, M., Schroeder, B., 1996. Saving energy using market-based control, in: Clearwater, S.H. (Ed.), *Market-based control: A paradigm for distributed resource allocation*. World Scientific, pp. 253–273.
- Domingos, E.F., Santos, F.C., Lenaerts, T., 2023. EGTtools: Evolutionary game dynamics in python. *iScience* 26, 106419. doi:10.1016/j.isci.2023.106419.
- Gao, B., Pavel, L., 2021. On passivity, reinforcement learning, and higher order learning in multiagent finite games. *IEEE Transactions on Automatic Control* 66, 121–136. doi:10.1109/TAC.2020.2978037.
- Genis-Mendoza, F., Konstantopoulos, G., Bauso, D., 2022. Online pricing for demand-side management in a low-voltage resistive micro-grid via a Stackelberg game with incentive strategies. *IET Smart Grid* 5, 76–89. doi:10.1049/stg2.12053.
- Gong, L., Yao, W., Gao, J., Cao, M., 2022. Limit cycles analysis and control of evolutionary game dynamics with environmental feedback. *Automatica* 145, 110536. doi:10.1016/j.automatica.2022.110536.
- Govaert, A., Zino, L., Tegling, E., 2022. Population games on dynamic community networks. *IEEE Control Systems Letters* 6, 2695–2700. doi:10.1109/LCSYS.2022.3174916.
- Grammatico, S., 2017. Dynamic control of agents playing aggregative games with coupling constraints. *IEEE Transactions on Automatic Control* 62, 4537–4548. doi:10.1109/TAC.2017.2672902.
- Grammatico, S., Parise, F., Colombino, M., Lygeros, J., 2016. Decentralized convergence to Nash equilibria in constrained deterministic mean field control. *IEEE Transactions on Automatic Control* 61, 3315–3329. doi:10.1109/TAC.2015.2513368.
- Groot, N., De Schutter, B., Hellendoorn, H., 2014. Toward system-optimal routing in traffic networks: A reverse Stackelberg game approach. *IEEE Transactions on Intelligent Transportation Systems* 16, 29–40. doi:10.1109/TITS.2014.2322312.
- Han, Z., Niyato, D., Saad, W., Başar, T., 2019. *Game theory for next generation wireless and communication networks: Modeling, analysis, and design*. Cambridge University Press.
- Horstmann, D., 2005. On some cross-diffusion models in population dynamics and their connections to well-posed filters in signal enhancement processes. *IMA journal of applied mathematics* 70, 386–399. doi:10.1093/imamat/hxh036.
- Izquierdo, L.R., Izquierdo, S.S., Sandholm, W.H., 2019. An introduction to ABED: Agent-based simulation of evolutionary game dynamics. *Games and Economic Behavior* 118, 434–462. doi:10.1016/j.geb.2019.09.014.
- Jackson, M.O., et al., 2008. *Social and economic networks*. volume 3. Princeton University Press.
- Jaleel, H., Shamma, J.S., 2020. Distributed optimization for robot networks: From real-time convex optimization to game-theoretic self-organization. *Proceedings of the IEEE* 108, 1953–1967. doi:10.1109/JPROC.2020.3028295.
- Kara, S., Martins, N.C., Arcaç, M., 2022. Population games with Erlang clocks: Convergence to Nash equilibria for pairwise comparison dynamics, in: 2022 IEEE Conference on Decision and Control, Cancun, MX, IEEE. pp. 7688–7695. doi:10.1109/CDC51059.2022.9993228.
- Khalil, H.K., 2002. *Nonlinear systems*; 3rd ed. Prentice-Hall.
- Li, J., 2020. A data-driven improved fuzzy logic control optimization-simulation tool for reducing flooding volume at downstream urban drainage systems. *Science of the Total Environment* 732, 138931. doi:10.1016/j.scitotenv.2020.138931.
- Lu, Z., Cai, F., Liu, J., Yang, J., Zhang, S., Wu, S., 2022. Evolution of water resource allocation in the river basin between administrators and managers. *Hydrology Research* 53, 716–732. doi:10.2166/nh.2022.128.
- Mahmoudi, A., Deng, X., Javed, S.A., Yuan, J., 2021. Large-scale multiple criteria decision-making with missing values: project selection through topsis-ops. *Journal of Ambient Intelligence and Humanized Computing* 12, 9341–9362. doi:10.1007/s12652-020-02649-w.
- Mandal, N., Tallapragada, P., 2022. Dynamics of a stratified population of optimum seeking agents on a network—part i: Modeling and convergence analysis. *IEEE Transactions on Control of Network Systems* 9, 1104–1115. doi:10.1109/TCNS.2021.3122528.
- Martínez, S.C., Montero, J.J.G., Gómez, M.M., Martín, F.A., Salichs, M.Á., 2023. Aplicación de estrategias psicológicas y sociales para incrementar el vínculo en interacción humano-robot. *Revista Iberoamericana de Automática e Informática industrial* 20, 199–212. doi:10.4995/riai.2023.18739.
- Martinez-Piazuelo, J., Ananduta, W., Ocampo-Martinez, C., Grammatico, S., Quijano, N., 2023. Population games with replicator dynamics under event-triggered payoff provider and a demand response application. *IEEE Control Systems Letters* 7, 3417–3422. doi:10.1109/LCSYS.2023.3285532.
- Martinez-Piazuelo, J., Diaz-Garcia, G., Quijano, N., Giraldo, L.F., 2022a. Discrete-time distributed population dynamics for optimization and control. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 52, 7112–7122. doi:10.1109/TSMC.2022.3151042.
- Martinez-Piazuelo, J., Ocampo-Martinez, C., Quijano, N., 2022b. Generalized Nash equilibrium seeking in population games under the Brown-von Neumann-Nash dynamics, in: 2022 European Control Conference, London, UK, IEEE. pp. 2161–2166. doi:10.23919/ECC55457.2022.9838437.
- Martinez-Piazuelo, J., Quijano, N., Ocampo-Martinez, C., 2021. Decentralized charging coordination of electric vehicles under feeder capacity constraints. *IEEE Transactions on Control of Network Systems* 9, 1600–1610. doi:10.1109/TCNS.2021.3128498.
- Martinez-Piazuelo, J., Quijano, N., Ocampo-Martinez, C., 2022c. Nash equilibrium seeking in full-potential population games under capacity and migration constraints. *Automatica* 141, 110285. doi:10.1016/j.automatica.2022.110285.
- Martinez-Piazuelo, J., Quijano, N., Ocampo-Martinez, C., 2022d. A payoff dynamics model for generalized Nash equilibrium seeking in population games. *Automatica* 140, 110227. doi:10.1016/j.automatica.2022.110227.
- Martins, N.C., Certorio, J., La, R.J., 2023. Epidemic population games and evolutionary dynamics. *Automatica* 153, 111016. doi:10.1016/j.automatica.2023.111016.
- Maynard-Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18. doi:10.1038/246015a0.
- Mojica-Nava, E., Macana, C.A., Quijano, N., 2013. Dynamic population games for optimal dispatch on hierarchical microgrid control. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 44, 306–317. doi:10.1109/TSMCC.2013.2266117.
- Muros, F., 2021. El control coalicional en el marco de la teoría de juegos cooperativos. *Revista Iberoamericana de Automática e Informática Industrial* 18, 97–112. doi:10.4995/riai.2020.13456.
- von Neumann, J., 1928. Zur theorie der gesellschaftsspiele. *Mathematische Annalen* 100, 295–320. doi:10.1007/BF01448847.
- von Neumann, J., Morgenstern, O., 1947. *Theory of games and economic behavior*. Princeton university press.
- Obando, G., Pantoja, A., Quijano, N., 2013. Building temperature control based on population dynamics. *IEEE Transactions on Control Systems Technology* 22, 404–412. doi:10.1109/TCST.2013.2252175.
- Obando, G., Poveda, J.I., Quijano, N., 2016. Replicator dynamics under perturbations and time delays. *Mathematics of Control, Signals, and Systems* 28, 1–32. doi:10.1007/s00498-016-0170-9.

- Obando, G., Quijano, N., Ocampo-Martinez, C., 2022. Decentralized control for urban drainage systems using replicator dynamics. *IEEE Access* 10, 56740–56762. doi:10.1109/ACCESS.2022.3177631.
- Ochoa, D.E., Poveda, J.I., Uribe, C.A., Quijano, N., 2021. Robust optimization over networks using distributed restarting of accelerated dynamics. *IEEE Control Systems Letters* 5, 301–306. doi:10.1109/LCSYS.2020.3001632.
- Pantoja, A., Obando, G., Quijano, N., 2019. Distributed optimization with information-constrained population dynamics. *Journal of the Franklin Institute* 356, 209–236. doi:10.1016/j.jfranklin.2018.10.016.
- Park, S., Barreiro-Gomez, J., 2023. Payoff mechanism design for coordination in multi-agent task allocation games. arXiv preprint arXiv:2306.02278 doi:10.48550/arXiv.2306.02278.
- Park, S., Leonard, N.E., 2021. KL divergence regularized learning model for multi-agent decision making, in: 2021 American Control Conference, New Orleans, US, IEEE. pp. 4509–4514. doi:10.23919/ACC50511.2021.9483414.
- Park, S., Martins, N.C., Shamma, J.S., 2019. From population games to payoff dynamics models: A passivity-based approach, in: 2019 IEEE Conference on Decision and Control, Nice, FR, IEEE. pp. 6584–6601. doi:10.1109/CDC40024.2019.9029756.
- Pashaie, A., Pavel, L., Damaren, C.J., 2017. A population game approach for dynamic resource allocation problems. *International Journal of Control* 90, 1957–1972. doi:10.1080/00207179.2016.1231422.
- Pawlick, J., Zhu, Q., 2021. Game theory for cyber deception: From theory to applications. Springer-Verlag.
- Quijano, N., Ocampo-Martinez, C., Barreiro-Gomez, J., Obando, G., Pantoja, A., Mojica-Nava, E., 2017. The role of population games and evolutionary dynamics in distributed control systems. *IEEE Control Systems Magazine* 37, 70–97. doi:10.1109/MCS.2016.2621479.
- Rass, S., Schauer, S., König, S., Zhu, Q., 2020. Cyber-security in critical infrastructures: A game-theoretic approach. Springer.
- Sánchez-Amores, A., Martínez-Piázuolo, J., Maestre, J.M., Ocampo-Martinez, C., Camacho, E.F., Quijano, N., 2023. Coalitional model predictive control of parabolic-trough solar collector fields with population-dynamics assistance. *Applied Energy* 334, 120740. doi:10.1016/j.apenergy.2023.120740.
- Sandholm, W.H., 2010. Population games and evolutionary dynamics. MIT Press.
- Smith, M.J., 1984. The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov. *Transportation Science* 18, 245–252. doi:10.1287/trsc.18.3.245.
- Srikantha, P., Kundur, D., 2017. Resilient distributed real-time demand response via population games. *IEEE Transactions on Smart Grid* 8, 2532–2543. doi:10.1109/TSG.2016.2526651.
- Tan, S., Fang, Z., Wang, Y., Lü, J., 2022. Consensus-based multipopulation game dynamics for distributed nash equilibria seeking and optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53, 813–823. doi:10.1109/TSMC.2022.3188266.
- Tan, S., Wang, Y., Vasilakos, A.V., 2021. Distributed population dynamics for searching generalized nash equilibria of population games with graphical strategy interactions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 52, 3263–3272. doi:10.1109/TSMC.2021.3062827.
- Taylor, P.D., Jonker, L.B., 1978. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40, 145–156. doi:10.1016/0025-5564(78)90077-9.
- Tembine, H., Altman, E., El-Azouzi, R., Hayel, Y., 2010. Evolutionary games in wireless networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 40, 634–646. doi:10.1109/TSMCB.2009.2034631.
- Villanueva, P., Bona, S., Lostado-Lorza, R., Veiga, F., 2023. Morphological design of a bicycle propulsion component using the hierarchical analysis process (AHP). *Applied Sciences* 13, 7792. doi:10.3390/app13137792.
- Wang, Z., Liu, F., Ma, Z., Chen, Y., Jia, M., Wei, W., Wu, Q., 2021. Distributed generalized Nash equilibrium seeking for energy sharing games in prosumers. *IEEE Transactions on Power Systems* 36, 3973–3986. doi:10.1109/TPWRS.2021.3058675.
- Weibull, J.W., 1997. Evolutionary game theory. MIT press.
- Ygge, F., Akkermans, H., 1999. Decentralized markets versus central control: A comparative study. *Journal of artificial intelligence research* 11, 301–333. doi:10.1613/jair.627.
- Yi, P., Pavel, L., 2019. An operator splitting approach for distributed generalized Nash equilibria computation. *Automatica* 102, 111–121. doi:10.1016/j.automatica.2019.01.008.
- Zhu, Q., Tembine, H., Başar, T., 2012. Evolutionary games for multiple access control, in: *Advances in Dynamic Games: Theory, Applications, and Numerical Methods for Differential and Stochastic Games*. Springer, pp. 39–71. doi:10.1007/978-0-8176-8355-9_3.