

Jogos Não-Cooperativos com Equações Diferenciais Parciais de Transporte e Difusão [★]

Tiago Roux Oliveira ^{*} Victor Hugo Pereira Rodrigues ^{**}
Miroslav Krstić ^{***} Tamer Başar ^{****}

^{*} *Universidade do Estado do Rio de Janeiro (UERJ),
Rio de Janeiro, RJ, Brasil
(e-mail: tiagoroux@uerj.br)*

^{**} *Universidade Federal do Rio de Janeiro (UFRJ),
Rio de Janeiro, RJ, Brasil
(e-mail: rodrigues.vhp@gmail.com)*

^{***} *University of California at San Diego (UCSD),
San Diego – CA, USA
(e-mail: krstic@ucsd.edu)*

^{****} *University of Illinois at Urbana-Champaign,
Urbana – IL, USA
(e-mail: basar1@illinois.edu)*

Abstract: We propose a non-model based strategy for locally stable convergence to Nash equilibria in a quadratic noncooperative (duopoly) game with player actions subject to heterogeneous PDE dynamics. In this duopoly scenario where different players use different types of PDEs, one player compensates for a delay (transport PDE) and the other a heat (diffusion) PDE, each player having access only to his own payoff value. In order to compensate distinct PDE-modeled processes in the inputs of the two players, we employ boundary control with averaging-based estimates. We apply a small-gain analysis for the resulting Input-to-State Stable (ISS) coupled hyperbolic-parabolic PDE system as well as averaging theory in infinite dimensions, due to the infinite-dimensional state of the heat PDE and the delay, in order to obtain local convergence results to a small neighborhood of the Nash equilibrium. We quantify the size of these residual sets and illustrate the theoretical results numerically.

Resumo: Este artigo propõe uma estratégia de busca pelo equilíbrio de Nash baseada no não-conhecimento do modelo que garanta localmente a convergência estável em um jogo quadrático não-cooperativo (duopólio) com ações de jogadores sujeitas a dinâmicas governadas por EDPs heterogêneas. Neste cenário de duopólio onde diferentes jogadores usam diferentes tipos de EDPs, um jogador compensa um atraso (EDP de transporte) e o outro a EDP do calor (difusão), com cada jogador tendo acesso apenas ao seu próprio valor de *payoff*. A fim de se compensar processos distintos modelados por EDPs nas atuações dos jogadores, emprega-se o controle de fronteira com estimativas baseadas em média. Aplica-se uma análise de pequeno ganho para o sistema *Input-to-State Stable* (ISS) resultante do acoplamento de EDPs, bem como a teoria da média em dimensões infinitas, devido ao estado da EDP de calor e do atraso, para se obter resultados de convergência locais para uma pequena vizinhança do equilíbrio de Nash. Quantifica-se o tamanho desse conjunto residual e ilustra-se os resultados teóricos numericamente.

Keywords: Extremum seeking; Nash equilibrium; (Non)cooperative games; Transport equations; Heat equations; Predictor feedback; Averaging theory in infinite dimensions.

Palavras-chaves: Busca extremal; Equilíbrio de Nash; Jogos não-cooperativos; Equações de transporte; Equações de calor; Realimentação por preditor; Teoria da média em dimensões infinitas.

1. INTRODUÇÃO

A teoria dos jogos oferece uma base teórica eficaz para o emprego de modelos matemáticos envolvendo jogadores que interagem estrategicamente [1]. Esquemas de teoria dos jogos tornam este tópico de pesquisa relevante para diversos campos com diversas aplicações [2, 3]. Nesse sentido, podemos visualizar os jogos em duas categorias principais: cooperativos e não-cooperativos [4]. Diz-se que um jogo não é cooperativo se os jogadores não puderem formar alianças ou se todos os acordos precisarem ser autoaplicáveis, concentrando-se em prever ações e recompensas individuais de jogadores e analisar o *equilíbrio de Nash* [4]. O equilíbrio de Nash é um resultado que, uma vez alcançado, significa que nenhum jogador pode aumentar seu lucro alterando decisões unilateralmente [5].

Em particular, na referência [6] estudam o problema de encontrar, em tempo real, o equilíbrio de Nash em jogos estáticos não-cooperativos de N jogadores aplicando uma abordagem baseada no não-conhecimento do modelo via busca extremal [7]. A referência [8] forneceu uma contribuição para o problema de busca extremal de jogos não-cooperativos com dois jogadores sujeitos a atrasos constantes usando realimentação por preditores [9, 10] enquanto o artigo [11] estudou jogos não-cooperativos com dois jogadores atuando através da dinâmica EDP de calor, que pode representar comportamentos humanos na teoria dos jogos, como passos aleatórios (decisões) e movimento browniano, sendo conectado com pequenas variantes da equação Black-Scholes nos mercados financeiros [13, 14].

Neste artigo, generaliza-se os resultados de busca extremal obtidos em [8, 11] para uma classe mais ampla de sistemas de dimensão infinita governados por Equações Diferenciais Parciais (EDPs) heterogêneas, em vez de se restringir apenas a atrasos puros **ou** apenas um tipo de dinâmica EDP nas variáveis de decisão dos jogadores.

Neste cenário, pode-se desenvolver um resultado para jogos heterogêneos com EDPs distintas, de transporte e de difusão (calor), a serem compensadas simultaneamente na malha de ação de cada jogador. Além do mais, os jogadores estimam apenas as entradas diagonais da matriz Hessiana devido a estrutura de seus *payoffs* no problema de busca em tempo real do equilíbrio Nash. A estratégia é capazes de mostrar que os termos suficientemente pequenos fora da diagonal são dominados usando um argumento de ganho pequeno [15] para o sistema médio.

Análise de estabilidade apresenta uma sequência de etapas de cálculo de médias em dimensões infinitas adequadamente projetadas [16] e Teorema do Pequeno Ganho para múltiplas cascatas *Input-to-State Stable* (ISS) de sistemas EDP-EDO hiperbólicos/parabólicos [15] no jogo do duopólio heterogêneo para provar estabilidade em malha fechada com não-linearidades e perturbações periódicas. Uma pequena vizinhança do equilíbrio de Nash é alcan-

* O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Os autores também agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

çada, mesmo na presença concomitante de EDPs de transporte e de calor.

2. NOTAÇÃO E TERMINOLOGIA

Denota-se a derivada parcial de uma função $u(x, t)$ como $\partial_x u(x, t) = \partial u(x, t) / \partial x$, $\partial_t u(x, t) = \partial u(x, t) / \partial t$, ou, convenientemente, $u_x(x, t)$ e $u_t(x, t)$, respectivamente. A norma-2 (Euclidina) de um vetor de estado de dimensão finita $\vartheta(t)$ é denotada por barras simples, $|\vartheta(t)|$. Denota-se a norma espacial $\mathcal{L}_2[0, D]$ de $u(x, t)$ como $\|u(t)\|_{\mathcal{L}_2([0, D])}^2 := \int_0^D u^2(x, t) dx$. Para simplificar a notação, omite-se o índice $\mathcal{L}_2([0, D])$ e, portanto, $\|\cdot\| = \|\cdot\|_{\mathcal{L}_2([0, D])}$ [9]. Como definido em [17], uma função vetorial $f(t, \epsilon) \in \mathbb{R}^n$ é dita ser de ordem $\mathcal{O}(\epsilon)$ em um intervalo $[t_1, t_2]$, se $\exists k, \bar{\epsilon} : |f(t, \epsilon)| \leq k\epsilon, \forall \epsilon \in [0, \bar{\epsilon}]$ and $\forall t \in [t_1, t_2]$. Na maioria dos casos, não se fornece estimativas precisas para as constantes k e $\bar{\epsilon}$, e interpreta-se $\mathcal{O}(\epsilon)$ como um resíduo de ordem de magnitude ϵ suficientemente pequena. A definição de *Input-to-State Stability* (ISS) para sistemas baseados em EDO, bem como para sistemas baseados em EDP, é considerada como em [17] e [15], respectivamente. O operador de Laplace é representado por \mathcal{L} enquanto o de convolução é dado por $*$.

3. DUOPÓLIO COM PAYOFFS QUADRÁTICOS E EDPS HETEROGÊNEAS: FORMULAÇÃO GERAL

Em um jogo de duopólio, a otimalidade dos respectivos lucros (saídas) dos jogadores P_1 e P_2 , respectivamente $y_1(t) \in \mathbb{R}$ e $y_2(t) \in \mathbb{R}$, não dependem exclusivamente de suas próprias ações/estratégias (sinais de entrada) $\Theta_1(t) \in \mathbb{R}$ e $\Theta_2(t) \in \mathbb{R}$. Além disso, definindo $\Theta(t) := [\Theta_1(t), \Theta_2(t)]^T$, a função de *payoff* de cada jogador J_i depende também da ação Θ_j do outro jogador $j, j \neq i$. Diz-se que uma tupla de ações $\Theta^* = [\Theta_1^*, \Theta_2^*]^T$ está em equilíbrio de Nash, se nenhum jogador i puder melhorar sua recompensa desviando unilateralmente de Θ_i^* , sendo assim para todos $i \in \{1, 2\}$ [1]. Apesar do grande número de publicações sobre a busca do equilíbrio de Nash [6], seu estudo sob dinâmicas EDP heterogêneas ainda é um problema em aberto.

Conforme mostrado nas Fig. 1 e Fig. 2, EDPs distintas (transporte e calor) (com atuação Dirichlet) são consideradas no vetor de ações dos jogadores $\theta(t) \in \mathbb{R}^2$. Assim, o vetor atuação propagada $\Theta(t) \in \mathbb{R}^2$ é dada pelo seguinte EDP de transporte para o jogador P_1

$$\Theta_1(t) = \theta_1(t - D_1) = \alpha_1(0, t), \quad (1)$$

$$\partial_t \alpha_1(x, t) = \partial_x \alpha_1(x, t), \quad x \in (0, D_1) \quad (2)$$

$$\partial_x \alpha_1(0, t) = 0 \quad (3)$$

$$\alpha_1(D_1, t) = \theta_1(t), \quad (4)$$

e a próxima EDP de calor para o jogador P_2

$$\Theta_2(t) = \alpha_2(0, t), \quad (5)$$

$$\partial_t \alpha_2(x, t) = \partial_{xx} \alpha_2(x, t), \quad x \in (0, D_2) \quad (6)$$

$$\partial_x \alpha_2(0, t) = 0 \quad (7)$$

$$\alpha_2(D_2, t) = \theta_2(t), \quad (8)$$

nas quais $\alpha_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}, \forall i \in \{1, 2\}$, e sendo conhecido cada comprimento de domínio D_i .

A solução de (1)–(4) é

$$\alpha_1(x, t) = \theta_1(t + x - D_1), \quad (9)$$

que representa uma ação atrasada para o jogador P_1 em $x = 0$.

Por outro lado, a solução de (5)–(8) é dada por

$$\alpha_2(x, t) = \mathcal{L}^{-1} \left[\frac{\cosh(x\sqrt{s})}{\cosh(D_2\sqrt{s})} \right] * \theta_2(t), \quad (10)$$

onde $\mathcal{L}^{-1}(\cdot)$ denota a transformada inversa de Laplace e $*$ é o operador de convolução. Dadas essas relações, definimos o operador *heterogêneo transporte-difusão* $\mathcal{D} = \text{diag}\{\mathcal{D}_1, \mathcal{D}_2\}$ para as EDPs (1)–(4) e (5)–(8) com fronteiras de entradas e medições dadas por

$$\begin{aligned} \mathcal{D}_1[\varphi(t)] &= \varphi(t + x - D_1), \quad \text{s.t. } \Theta_1(t) = \mathcal{D}_1[\theta_1(t)], \\ \mathcal{D}_2[\varphi(t)] &= \mathcal{L}^{-1} \left[\frac{1}{\cosh(D_2\sqrt{s})} \right] * \varphi(t), \quad \text{s.t. } \Theta_2(t) = \mathcal{D}_2[\theta_2(t)]. \end{aligned} \quad (11)$$

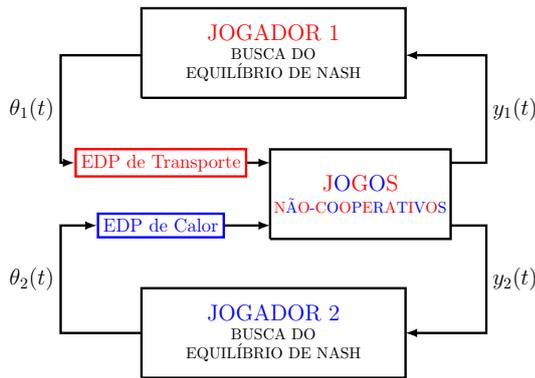


Figura 1. Busca do equilíbrio de Nash em um jogo heterogêneo não-cooperativo com jogadores atuando através de EDP de transporte-calor.

Considera-se jogos nos quais a função de *payoff* $y_i(t) = J_i(\Theta(t))$, $\forall i \in \{1, 2\}$, de cada jogador é quadrática [5], expressadas como uma combinação estritamente côncava de ações propagadas através das distintas EDPs de transporte-calor

$$J_1(\Theta(t)) = \frac{H_{11}^1}{2} \Theta_1^2(t) + \frac{H_{22}^1}{2} \Theta_2^2(t) + \epsilon H_{12}^1 \Theta_1(t) \Theta_2(t) + h_1^1 \Theta_1(t) + h_2^1 \Theta_2(t) + c_1, \quad (12)$$

$$J_2(\Theta(t)) = \frac{H_{11}^2}{2} \Theta_1^2(t) + \frac{H_{22}^2}{2} \Theta_2^2(t) + \epsilon H_{21}^2 \Theta_1(t) \Theta_2(t) + h_1^2 \Theta_1(t) + h_2^2 \Theta_2(t) + c_2, \quad (13)$$

nas quais $J_1(\Theta), J_2(\Theta) : \mathbb{R}^2 \rightarrow \mathbb{R}$, $H_{jk}^i, h_j^i, c_i \in \mathbb{R}$ são constantes, $H_{ii}^i < 0$, $\forall i, j, k \in \{1, 2\}$, e $\epsilon > 0$ sem perda de generalidade.

As funções de *payoff* quadráticas são de particular interesse na teoria dos jogos, primeiro porque constituem aproximações de segunda ordem para outros tipos de funções de *payoff* não quadráticas e, em segundo lugar, porque são analiticamente tratáveis, levando, em geral, a soluções de equilíbrio de forma fechada que fornecem insights sobre as propriedades e características do conceito de solução de equilíbrio em consideração [5].

Por uma questão de completude, fornecemos aqui, em termos matemáticos, a definição de um equilíbrio de Nash $\Theta^* = [\Theta_1^*, \Theta_2^*]^T$ em um jogo de 2-jogadores:

$$J_1(\Theta_1^*, \Theta_2^*) \geq J_1(\Theta_1, \Theta_2^*) \quad \text{e} \quad J_2(\Theta_1^*, \Theta_2^*) \geq J_2(\Theta_1^*, \Theta_2). \quad (14)$$

Portanto, nenhum jogador tem incentivo para desviar unilateralmente sua ação de Θ^* .

Para determinar a solução de equilíbrio de Nash em jogos quadráticos estritamente côncavos com dois jogadores deve-se diferenciar J_1 e J_2 , respectivamente, em relação a $\Theta_1(t)$ e $\Theta_2(t)$, igualando a zero as expressões resultantes e resolvendo o conjunto de equações assim obtido. Este conjunto de equações, que também fornece uma condição suficiente devido à estrita concavidade, é

$$\begin{cases} H_{11}^1 \Theta_1^* + \epsilon H_{12}^1 \Theta_2^* + h_1^1 = 0 \\ \epsilon H_{21}^1 \Theta_1^* + H_{22}^1 \Theta_2^* + h_2^1 = 0 \end{cases}, \quad (15)$$

que pode ser reescrita de forma matricial como

$$\begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^1 & H_{22}^1 \end{bmatrix} \begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix} = - \begin{bmatrix} h_1^1 \\ h_2^1 \end{bmatrix}. \quad (16)$$

Definindo a matriz Hessiana H e os vetores Θ^* e h por

$$H := \begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^1 & H_{22}^1 \end{bmatrix}, \quad \Theta^* := \begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix}, \quad h := \begin{bmatrix} h_1^1 \\ h_2^1 \end{bmatrix}, \quad (17)$$

existe um único equilíbrio de Nash dado por $\Theta^* = -H^{-1}h$, se H for inversível:

$$\begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix} = - \begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^1 & H_{22}^1 \end{bmatrix}^{-1} \begin{bmatrix} h_1^1 \\ h_2^1 \end{bmatrix}. \quad (18)$$

Para mais detalhes, consultar [5, Capítulo 4].

O *objetivo de controle* é projetar uma nova estratégia baseada em busca de extremal para alcançar o Equilíbrio de Nash em jogos não-cooperativos submetidos a atuação EDPs de transporte e de calor nas variáveis de decisão dos jogadores (sinais de entrada).

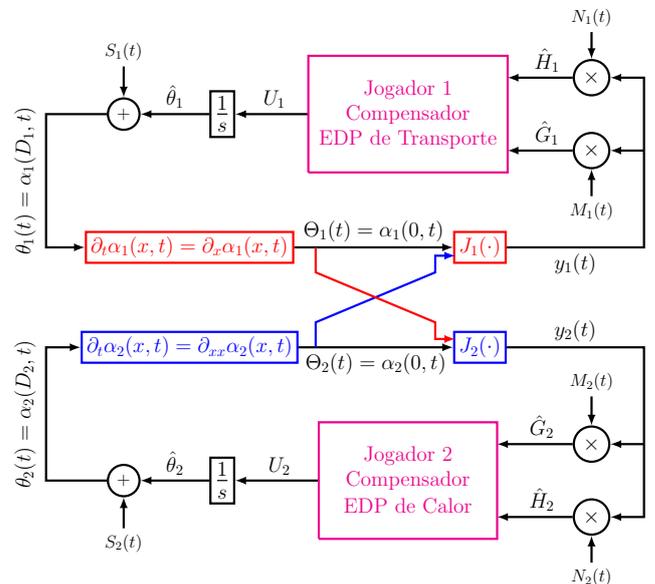


Figura 2. Diagrama de blocos ilustrando a estratégia de busca do Equilíbrio de Nash realizada por cada jogador. Na cor magenta estão os controladores por fronteira usados para compensar individualmente as EDPs de transporte e de calor no jogo heterogêneo não-cooperativo.

Como o objetivo é encontrar as entradas ótimas desconhecidas Θ^* (e θ^*), define-se os erros de estimativa

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^*, \quad \tilde{\Theta}(t) = \hat{\Theta}(t) - \Theta^*, \quad (19)$$

onde os vetores $\hat{\theta}(t)$ e $\hat{\Theta}(t)$ são as estimativas de θ^* e Θ^* . Para tornar (19) coerente com o otimizador do

mapa estático Θ^* , aplica-se o operador heterogêneo de transporte-difusão (11) a $\tilde{\theta}_i$ em (19) obtendo

$$\vartheta_1(t) = \tilde{\theta}_1(t - D_1) := \bar{\alpha}_1(0, t), \quad (20)$$

$$\partial_t \bar{\alpha}_1(x, t) = \partial_x \bar{\alpha}_1(x, t), \quad x \in (0, D_1) \quad (21)$$

$$\partial_x \bar{\alpha}_1(0, t) = 0 \quad (22)$$

$$\bar{\alpha}_1(D_1, t) = \tilde{\theta}_1(t), \quad (23)$$

e

$$\vartheta_2(t) := \bar{\alpha}_2(0, t), \quad (24)$$

$$\partial_t \bar{\alpha}_2(x, t) = \partial_{xx} \bar{\alpha}_2(x, t), \quad x \in (0, D_2) \quad (25)$$

$$\partial_x \bar{\alpha}_2(0, t) = 0 \quad (26)$$

$$\bar{\alpha}_2(D_2, t) = \tilde{\theta}_2(t), \quad (27)$$

onde $\bar{\alpha}_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}$, $\forall i \in \{1, 2\}$, e $\vartheta(t) := \mathcal{D}[\tilde{\theta}(t)] = \hat{\Theta}(t) - \Theta^*$ é o erro de estimação propagado $\tilde{\theta}(t)$ através do domínio das EDPs. Para $\lim_{t \rightarrow \infty} \theta(t) = \theta_c$, tem-se $\lim_{t \rightarrow \infty} \Theta(t) = \Theta_c = \theta_c$, onde o índice c indica um sinal constante. De fato, de (11), para uma entrada constante $\theta = \theta_c$, tem-se $\lim_{t \rightarrow \infty} \mathcal{D}_1[\theta_{c1}] = \theta_{c1}$ para o Jogador P_1 . Para o Jogador P_2 , tem-se $\mathcal{L}\{\theta_{c2}\} = \theta_{c2}/s$, e através do Teorema do valor final, $\lim_{t \rightarrow \infty} \mathcal{D}_2[\theta_{c2}] = \lim_{s \rightarrow 0} \left\{ \frac{\theta_{c2}}{\cosh(\sqrt{s}D_2)} \right\} = \theta_{c2}$. Então, no caso particular $\theta = \theta_c = \theta^*$, tem-se

$$\Theta^* = \theta^*. \quad (28)$$

A Fig. 2 mostra um diagrama esquemático que resume a política de busca do equilíbrio de Nash proposta para cada jogador, onde suas saídas são dadas por

$$\begin{cases} y_1(t) &= J_1(\Theta(t)), \\ y_2(t) &= J_2(\Theta(t)). \end{cases} \quad (29)$$

Os sinais de *dither* aditivos na presença de EDPs de calor e de transporte [10, 19] são definidos de acordo com

$$\begin{cases} S_1(t) = a_1 \sin(\omega_1(t + D_1)) \\ S_2(t) = \frac{1}{2} a_2 e^{\sqrt{\frac{\omega_2}{2}} D_2} \sin\left(\omega_2 t + \sqrt{\frac{\omega_2}{2}} D_2\right) \\ \quad + \frac{1}{2} a_2 e^{-\sqrt{\frac{\omega_2}{2}} D_2} \sin\left(\omega_2 t - \sqrt{\frac{\omega_2}{2}} D_2\right) \end{cases}, \quad (30)$$

e os *dither* multiplicativos são dados por

$$\begin{cases} M_1(t) &= \frac{2}{a_1} \sin(\omega_1 t) \\ M_2(t) &= \frac{2}{a_2} \sin(\omega_2 t) \end{cases}, \quad (31)$$

com amplitudes constantes e diferentes de zero $a_1, a_2 > 0$ nas frequências $\omega_1 \neq \omega_2$. Tais frequências de sondagem ω_i podem ser selecionadas como

$$\omega_i = \omega'_i \omega = \mathcal{O}(\omega), \quad i = 1 \text{ or } 2, \quad (32)$$

onde ω é uma constante positiva e ω'_i é um número racional – uma escolha possível é dada em [18].

Seguindo o paradigma de não-compartilhamento de informações, apenas os elementos diagonais de H podem ser recuperados adequadamente no sentido médio pelos jogadores P_1 ou P_2 . Nesse sentido, os sinais $N_1(t)$ e $N_2(t)$ são simplesmente definidos por [18]:

$$\begin{cases} N_1(t) = \frac{16}{a_1^2} \left(\sin^2(\omega_1 t) - \frac{1}{2} \right) \\ N_2(t) = \frac{16}{a_2^2} \left(\sin^2(\omega_2 t) - \frac{1}{2} \right). \end{cases} \quad (33)$$

Então, a versão média de

$$\begin{cases} \hat{H}_1(t) = N_1(t)y_1(t) \\ \hat{H}_2(t) = N_2(t)y_2(t) \end{cases} \quad (34)$$

é dada por

$$\begin{cases} \hat{H}_1^{\text{av}}(t) = [N_1(t)y_1(t)]_{\text{av}} = H_{11}^1 \\ \hat{H}_2^{\text{av}}(t) = [N_2(t)y_2(t)]_{\text{av}} = H_{22}^2 \end{cases}. \quad (35)$$

Considerando $\hat{\theta}_1(t)$ e $\hat{\theta}_2(t)$ como as estimativas de θ_1^* e θ_2^* , pode-se definir de (19) os erros de estimativa individuais:

$$\begin{cases} \tilde{\theta}_1(t) = \hat{\theta}_1(t) - \theta_1^*, & \vartheta_1(t) = \hat{\Theta}_1(t) - \Theta_1^* \\ \tilde{\theta}_2(t) = \hat{\theta}_2(t) - \theta_2^*, & \vartheta_2(t) = \hat{\Theta}_2(t) - \Theta_2^* \end{cases}. \quad (36)$$

As estimativas dos componentes do gradiente desconhecido das funções de *payoff* são dadas por

$$\begin{cases} \hat{G}_1(t) &= M_1(t)y_1(t) \\ \hat{G}_2(t) &= M_2(t)y_2(t) \end{cases}, \quad (37)$$

e calculando-se a média do sinal resultante, chega-se a

$$\begin{cases} \hat{G}_1^{\text{av}}(t) = H_{11}^1 \vartheta_1^{\text{av}}(t) + \epsilon H_{12}^1 \vartheta_2^{\text{av}}(t) \\ \hat{G}_2^{\text{av}}(t) = \epsilon H_{21}^2 \vartheta_1^{\text{av}}(t) + H_{22}^2 \vartheta_2^{\text{av}}(t) \end{cases}. \quad (38)$$

Além disso, a partir do diagrama de blocos na Fig. 2, tem-se

$$\dot{\hat{\theta}}_i(t) = U_i(t), \quad \dot{\hat{\theta}}_i(t) = U_i(t), \quad \forall i \in \{1, 2\}, \quad (39)$$

já que $\dot{\hat{\theta}}(t) = \hat{\dot{\theta}}(t)$, uma vez que θ^* é constante. Tomando a derivada temporal de (20)-(23) e (24)-(27), com a ajuda de (19) e (39), a *dinâmica do erro de propagação* é escrita como

$$\dot{\vartheta}_1(t) = U(t - D_1) = u_1(0, t), \quad (40)$$

$$\partial_t u_1(x, t) = \partial_x u_1(x, t), \quad x \in (0, D_1) \quad (41)$$

$$\partial_x u_1(0, t) = 0 \quad (42)$$

$$u_1(D_1, t) = U_1(t), \quad (43)$$

e

$$\dot{\vartheta}_2(t) = u_2(0, t), \quad (44)$$

$$\partial_t u_2(x, t) = \partial_{xx} u_2(x, t), \quad x \in (0, D_2) \quad (45)$$

$$\partial_x u_2(0, t) = 0 \quad (46)$$

$$u_2(D_2, t) = U_2(t), \quad (47)$$

onde $u_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}$, $u_i(x, t) := \partial_t \bar{\alpha}_i(x, t)$, $\forall i \in \{1, 2\}$, e $\bar{\alpha}_i(x, t) = \alpha_i(x, t) - \beta_i(x, t) - \Theta_i^*$. O termo $\beta_i(x, t)$ representa o estado da EDP no *problema de geração de trajetória* [12, Capítulo 12] resolvido para se obter $S_1(t) = \beta_1(D_1, t)$ e $S_2(t) = \beta_2(D_2, t)$ em (30) — para mais detalhes, consultar [22, Eqs. (19)–(22)].

Portanto, de (38), (40)–(43) e (44)–(47), é possível encontrar uma forma compacta para o gradiente médio global estimado de acordo com

$$\hat{G}^{\text{rav}}(t) = H \vartheta^{\text{av}}(t), \quad (48)$$

$$\hat{G}^{\text{rav}}(t) = H \dot{\vartheta}^{\text{av}}(t) = H \mathcal{D}[U^{\text{av}}(t)], \quad (49)$$

onde a H Hessiana é dada na equação (17), $\vartheta^{\text{av}}(t) := [\vartheta_1^{\text{av}}(t), \vartheta_2^{\text{av}}(t)]^T \in \mathbb{R}^2$, $\hat{G}^{\text{rav}}(t) := [\hat{G}_1^{\text{rav}}(t), \hat{G}_2^{\text{rav}}(t)]^T \in \mathbb{R}^2$ e $U^{\text{av}}(t) := [U_1^{\text{av}}(t), U_2^{\text{av}}(t)]^T \in \mathbb{R}^2$ são as versões médias de $U(t) := [U_1(t), U_2(t)]^T$, $\vartheta(t) := [\vartheta_1(t), \vartheta_2(t)]^T$ e $\hat{G}(t) := [\hat{G}_1(t), \hat{G}_2(t)]^T$, respectivamente.

Ao longo do artigo, a ideia principal é projetar leis de controle (políticas) para cada jogador, a fim de alcançar

uma pequena vizinhança do ponto de equilíbrio de Nash. Para tanto, utilizamos uma estratégia de busca extremal baseada no controle de fronteira para compensar o operador heterogêneo de transporte-difusão $\mathcal{D}[\cdot]$ em (49) devido às distintas EDPs (transporte-calor) nas ações dos jogadores. Basicamente, as leis de controle devem ser capazes de garantir a estabilização exponencial de $\hat{G}^{av}(t)$ e, conseqüentemente, de $\vartheta^{av}(t) = \hat{\Theta}^{av}(t) - \Theta^*$. De (48), fica claro que, se H for inversível, $\vartheta^{av}(t) \rightarrow 0$ como $\hat{G}^{av}(t) \rightarrow 0$. Assim, a convergência de $\vartheta^{av}(t)$ para a origem resulta na convergência de $\Theta(t)$ para uma pequena vizinhança de Θ^* em (14) via teoria da média [16].

4. JOGO HETEROGÊNIO NÃO-COOPERATIVO REGIDO POR EDPS DE TRANSPORTE-CALOR

Como mencionado anteriormente, no jogo heterogêneo não-cooperativo sujeito à dinâmica PDE transporte-calor, o objetivo da busca extremal é estimar o equilíbrio de Nash $\Theta^* = \theta^*$, veja (28), mas sem compartilhar nenhuma informação entre os jogadores. Cada jogador só precisa medir o valor de sua própria função de *payoff* (29).

4.1 Controle descentralizado de fronteira utilizando apenas os termos diagonais da Hessiana

Nesse sentido, pode-se formular o sistema em malha fechada de forma *descentralizada*, onde não é necessário nenhum conhecimento sobre o *payoff* ou ação do outro jogador.

Inspirado por [10, 19], onde a busca extremal foi considerada com compensação de EDP, mas não no contexto de jogos, propõe-se as seguintes leis de atualização baseadas em fronteiras $\hat{\theta}_i(t) = U_i(t)$, $i \in \{1, 2\}$:

$$\begin{cases} \dot{U}_1(t) = -c_1 U_1(t) \\ \quad + c_1 k_1 \left(\hat{G}_1(t) + \hat{H}_1(t) \int_0^{D_1} u_1(\tau, t) d\tau \right) \\ \dot{U}_2(t) = -c_2 U_2(t) \\ \quad + c_2 k_2 \left(\hat{G}_2(t) + \hat{H}_2(t) \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \right) \end{cases}, \quad (50)$$

para constantes positivas $k_1 > 0$, $k_2 > 0$, $c_1 > 0$ e $c_2 > 0$, com intuito de compensar as EDPs de calor e transporte em (40)–(43) e (44)–(47).

A lei de controle por fronteira (50) pode ser reescrita como

$$\begin{aligned} \dot{U}_1(t) &= -c_1 U_1(t) \\ &\quad + c_1 k_1 \left(\hat{G}_1(t) + \hat{H}_1(t) \int_{t-D_1}^t U_1(\tau, t) d\tau \right) \\ \dot{U}_2(t) &= -c_2 U_2(t) \\ &\quad + c_2 k_2 \left[\hat{G}_2(t) + \hat{H}_2(t) \left(\hat{\theta}_2(t) - \Theta_2(t) + a_2 \sin(\omega_2 t) \right) \right], \end{aligned} \quad (51)$$

usando a relação $u_1(x, t) = U_1(t + x - D_1)$ para a EDP de transporte e a equação de difusão $\partial_t \alpha_2(x, t) = \partial_{xx} \alpha_2(x, t)$ bem como a integração por partes, associada a (5)–(8), (19) e lembrando que $\vartheta_2 + a_2 \sin(\omega_2 t) = \Theta_2(t) - \Theta_2^*$, analogamente a [19, Eq. (25)].

4.2 Propriedades do tipo ISS para sistemas representados por EDPs

Para simplificar, assume-se $c_1, c_2 \rightarrow +\infty$ em (50), resultando na seguinte expressão geral:

$$\begin{aligned} U_1(t) &= k_1 \left(\hat{G}_1(t) + \hat{H}_1(t) \int_0^{D_1} u_1(\tau, t) d\tau \right), \\ U_2(t) &= k_2 \left(\hat{G}_2(t) + \hat{H}_2(t) \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \right). \end{aligned} \quad (52)$$

Relembrando (40)–(43) e (44)–(47), a malha fechada do sistema de dimensão infinita (49) e (52) em sua versão média pode ser escrita na forma de EDP, dada por

$$\dot{\hat{G}}_1^{av}(t) = H_{11}^1 u_1^{av}(0, t) + \epsilon H_{12}^1 u_2^{av}(0, t), \quad (53)$$

$$\partial_t u_1^{av}(x, t) = D_1^{-1} \partial_x u_1^{av}(x, t), \quad x \in (0, 1), \quad (54)$$

$$\partial_x u_1^{av}(0, t) = 0, \quad (55)$$

$$u_1^{av}(1, t) = U_1^{av}(t), \quad (56)$$

e

$$\dot{\hat{G}}_2^{av}(t) = \epsilon H_{21}^2 u_1^{av}(0, t) + H_{22}^2 u_2^{av}(0, t), \quad (57)$$

$$\partial_t u_2^{av}(x, t) = D_2^{-2} \partial_{xx} u_2^{av}(x, t), \quad x \in (0, 1), \quad (58)$$

$$\partial_x u_2^{av}(0, t) = 0, \quad (59)$$

$$u_2^{av}(1, t) = U_2^{av}(t). \quad (60)$$

Na *abordagem por redução* [20] (ou atribuição de espectro finito), utiliza-se as seguintes transformações:

$$\begin{aligned} \bar{G}_1^{av}(t) &= \hat{G}_1^{av}(t) + \epsilon_{11}^1 H_{11}^1 \int_0^{D_1} u_1^{av}(\tau, t) d\tau \\ &\quad + \epsilon_{12}^1 H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2^{av}(\tau, t) d\tau \end{aligned} \quad (61)$$

e

$$\begin{aligned} \bar{G}_2^{av}(t) &= \hat{G}_2^{av}(t) + \epsilon_{22}^2 H_{22}^2 \int_0^{D_2} (D_2 - \tau) u_2^{av}(\tau, t) d\tau \\ &\quad + \epsilon_{21}^2 H_{21}^2 \int_0^{D_1} u_1^{av}(\tau, t) d\tau, \end{aligned} \quad (62)$$

onde $\epsilon_{11}^1 = \epsilon_{22}^2 = 1$ e $\epsilon_{12}^1 = \epsilon_{21}^2 = \epsilon$.

Com algumas manipulações matemáticas, não é difícil verificar que \bar{G}^{av} satisfaz

$$\dot{\bar{G}}^{av}(t) = H U^{av}(t). \quad (63)$$

Agora, após adicionar e subtrair os próximos termos em azul e vermelho em (52), a mesma pode ser reescrita como:

$$U_1(t) = k_1 \left(\hat{G}_1(t) + \hat{H}_1(t) \int_0^{D_1} u_1(\tau, t) d\tau + \epsilon H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau - k_1 \epsilon H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau, \right) \quad (64)$$

$$U_2(t) = k_2 \left(\hat{G}_2(t) + \hat{H}_2(t) \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau + \epsilon H_{21}^2 \int_0^{D_1} u_1(\tau, t) d\tau - \epsilon H_{21}^2 \int_0^{D_1} u_1(\tau, t) d\tau \right), \quad (65)$$

cuja forma compacta média é

$$U^{av}(t) = K \bar{G}^{av}(t) + \epsilon K \phi^{av}(\mathcal{D}, t), \quad (66)$$

onde a matriz $K := \text{diag}\{k_1, k_2\}$ com entradas $k_1 > 0$, $k_2 > 0$ e a variável auxiliar $\phi(\mathcal{D}, t)$ é definida como

$$\phi(\mathcal{D}, t) := - \begin{bmatrix} H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \\ H_{21}^2 \int_0^{D_1} u_1(\tau, t) d\tau \end{bmatrix}, \quad \phi(1, t) := - \begin{bmatrix} H_{12}^1 \int_0^1 D_2^2 (1 - \xi) u_2(\xi, t) d\xi \\ H_{21}^2 \int_0^1 D_1 u_1(\xi, t) d\xi \end{bmatrix}, \quad (67)$$

já que $\int_0^{D_j} (D_j - \tau) u_j^{av}(\tau, t) d\tau = \int_0^1 D_j^2 (1 - \xi) u_j^{av}(\xi, t) d\xi$, para $j \in \{1, 2\}$. Então, é possível encontrar uma forma média compacta para o jogo (63) e (66), como

$$\dot{\bar{G}}^{av}(t) = HK \bar{G}^{av}(t) + \epsilon HK \phi^{av}(1, t), \quad (68)$$

$$\partial_t u_1^{av}(x, t) = D_1^{-1} \partial_x u_1^{av}(x, t), \quad x \in (0, 1), \quad (69)$$

$$\partial_t u_2^{av}(x, t) = D_2^{-2} \partial_{xx} u_2^{av}(x, t), \quad x \in (0, 1), \quad (70)$$

$$\partial_x u^{av}(0, t) = 0, \quad (71)$$

$$u^{av}(1, t) = K \bar{G}^{av}(t) + \epsilon K \phi^{av}(1, t). \quad (72)$$

De (68), se HK for Hurwitz, fica claro que a dinâmica da variável de estado EDO $\bar{G}^{av}(t)$ é exponencialmente ISS [15] em relação ao estado da EDP $u^{av}(x, t) = [u_1^{av}(x, t), u_2^{av}(x, t)]^T$ por meio da função $\phi^{av}(1, t)$. Além disso, o subsistema EDP (69) e (70) é ISS [15] em relação a $\bar{G}^{av}(t)$ na condição de contorno $u^{av}(1, t)$.

4.3 Análise de Estabilidade

Nesta seção, mostra-se que a malha EDP-EDO hiperbólica-parabólica (68)–(72) contém um parâmetro ϵ que pode levar à estabilidade em malha fechada se for suficientemente pequeno. Para tanto, considera-se a seguinte condição para o jogo não-cooperativo [6].

Hipótese 1. A matriz Hessiana H dada por (17) é estritamente diagonal dominante, ou seja,

$$\begin{cases} |\epsilon H_{12}^1| < |H_{11}^1| \\ |\epsilon H_{21}^2| < |H_{22}^2| \end{cases}. \quad (73)$$

Por Hipótese 1, o Equilíbrio de Nash Θ^* existe e é único uma vez que matrizes estritamente diagonalmente dominantes são não-singulares pelo Teorema de Levy-Desplanques [21].

O próximo teorema reúne as propriedades de estabilidade/convergência para o sistema de malha fechada utilizando a realimentação por busca extremal proposta para alcance do equilíbrio de Nash no jogo não-cooperativo de 2-jogadores e com ações propagadas através de EDPs de transporte-calor.

Teorema 1. Considere o sistema de malha fechada (40)–(47) sob EDPs de calor-transporte (1)–(8) de coeficientes de transporte-difusão distintos D_1 e D_2 para um jogo quadrático de duopólio heterogêneo com funções de *payoff* (29) dadas em (12), (13) satisfazendo a Hipótese 1 e as leis de controle $U_i(t)$ definidas em (50) ou (51), com $\hat{G}_1(t)$, $\hat{G}_2(t)$ e $\hat{H}_1(t)$, $\hat{H}_2(t)$ dados por (34) e (37), respectivamente. Os sinais de perturbação ES são dados em (30) e (31). Assim, existem constantes $c_1 > 0$, $c_2 > 0$ e $\omega > 0$ suficientemente grandes, bem como $\epsilon > 0$ suficientemente pequenos para que o sistema do erro em malha fechada com estado $\vartheta_i(t)$, $u_i(x, t)$, $\forall i \in \{1, 2\}$, tem uma única solução periódica localmente exponencialmente estável em t de período $\Pi := 2\pi \times \text{MMC}^1\{1/\omega_i\}$, com ω_i em (32) de ordem $\mathcal{O}(\omega)$, denotado por $\vartheta_i^\Pi(t)$, $u_i^\Pi(x, t)$ e satisfatório, $\forall t \geq 0$:

$$\left(\sum_{i=1}^2 [\vartheta_i^\Pi(t)]^2 + \int_0^{D_i} [u_i^\Pi(x, t)]^2 dx \right)^{1/2} \leq \mathcal{O}(1/\omega). \quad (74)$$

Além do mais,

$$\limsup_{t \rightarrow +\infty} |\Theta(t) - \Theta^*| = \mathcal{O}(|a| + 1/\omega), \quad (75)$$

$$\limsup_{t \rightarrow +\infty} |\theta_1(t) - \theta_1^*| = \mathcal{O}(a_1 + 1/\omega), \quad (76)$$

$$\limsup_{t \rightarrow +\infty} |\theta_2(t) - \theta_2^*| = \mathcal{O}\left(a_2 e^{D_2 \sqrt{\omega/2}} + 1/\omega\right), \quad (77)$$

onde $a = [a_1 \ a_2]^T$ e $\theta^* = \Theta^*$ é o único e desconhecido equilíbrio de Nash dado por (18).

Prova. Primeiro, considere a representação EDP-EDO hiperbólica/parabólica equivalente (68)–(72) reescrita para cada Jogador P_i , $i \in \{1, 2\}$:

$$\dot{\bar{G}}_i^{av}(t) = H_{ii}^i k_i \bar{G}_i^{av}(t) + \epsilon H_{ii}^i k_i \phi_i^{av}(1, t), \quad (78)$$

$$\partial_t u_1^{av}(x, t) = D_1^{-1} \partial_x u_1^{av}(x, t), \quad x \in (0, 1), \quad (79)$$

$$\partial_t u_2^{av}(x, t) = D_2^{-2} \partial_{xx} u_2^{av}(x, t), \quad x \in (0, 1), \quad (80)$$

$$\partial_x u_i^{av}(0, t) = 0, \quad (81)$$

$$u_i^{av}(1, t) = k_i \bar{G}_i^{av}(t) + \epsilon k_i \phi_i^{av}(1, t). \quad (82)$$

Para o jogador P_1 , o sistema de malha fechada médio, dado por (78)–(79) e (81)–(82), satisfaz ambas as Hipóteses **(H1)** e **(H2)** do Teorema do Pequeno Ganho [15, Teorema 8.1, p. 198] para a cascata EDP-EDO hiperbólica com $n = 1$, $x(t) = \bar{G}_1^{av}(t)$, $F(\bar{G}_1^{av}(t), u_1(x, t), v(t)) = H_{11}^1 k_1 \bar{G}_1^{av}(t) + \epsilon v(t)$ com $v(t) = H_{11}^1 k_1 \phi_1^{av}(1, t)$, $c = 1/D_1$, $a(x) = 0$, $g(x, \bar{G}_1^{av}(t), u_1(x, t)) = 0$, $f(x, t) = 0$, $\varphi(d(t), u_1(x, t), \bar{G}_1^{av}(t)) = k_1 \bar{G}_1^{av}(t) + \epsilon d(t)$ onde $d(t) = k_1 \phi_1^{av}(1, t)$, $N = k_1$, $L = |H_{11}^1| k_1$, $B = 0$, $\gamma_2 = k_1$, $b_2 = \epsilon$ e $A = \gamma_1 = 0$.

Note que a Hipótese **(H1)** é assegurada com $M = 1$, $\gamma_3 = 0$, $\sigma = |H_{11}^1| k_1$ e $b_3 > 0$ sendo uma constante apropriada de

¹ A sigla MMC significa o mínimo múltiplo comum.

ordem $\mathcal{O}(\epsilon)$ como pode ser facilmente verificado por meio da fórmula de variação de constantes (para $i = 1$)

$$\begin{aligned} \bar{G}_i^{\text{av}}(t) &= \exp(-|H_{ii}^i|k_i t)\bar{G}_i^{\text{av}}(0) + \\ &+ \int_0^1 \exp(-|H_{ii}^i|k_i(t+s))\epsilon H_{ii}^i k_i \phi_i^{\text{av}}(1, s) ds, \end{aligned} \quad (83)$$

e da aplicação da desigualdade de Cauchy-Schwarz ao termo $\phi_1^{\text{av}}(1, t)$ na equação (67):

$$\begin{aligned} \phi_1^{\text{av}}(1, t) &\leq |H_{12}^1| D_2^2 \left(\int_0^1 (1-\tau)^2 d\tau \right)^{\frac{1}{2}} \times \left(\int_0^1 [u_2^{\text{av}}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}} \\ &\leq \frac{1}{\sqrt{3}} k_H D_2^2 \left(\int_0^1 [u_2^{\text{av}}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}}, \end{aligned} \quad (84)$$

já que $|H_{ij}^i| < k_H < \frac{1}{\epsilon} |H_{ii}^i|$ de acordo com a Hipótese 1, onde k_H é uma constante positiva de ordem $\mathcal{O}(1)$. Portanto, se $0 < \epsilon < 1$ for suficientemente pequeno, o [15, Teorema 8.1, p. 198] nos permite concluir que existem constantes $\delta_1, \Delta_1 > 0$ tais que para todo $u_{1,0}^{\text{av}} \in C^0([0, 1])$, $\bar{G}_{1,0}^{\text{av}} \in \mathbb{R}$, a única solução generalizada deste problema de valor de limite inicial, com $u_1^{\text{av}}(x, 0) = u_{1,0}^{\text{av}}$ e $\bar{G}_1^{\text{av}}(0) = \bar{G}_{1,0}^{\text{av}}$, satisfaz a seguinte estimativa:

$$\begin{aligned} |\bar{G}_1^{\text{av}}(t)| + \|u_1^{\text{av}}(t)\|_{\infty} &\leq \Delta_1 (|\bar{G}_{1,0}^{\text{av}}| + \|u_{1,0}^{\text{av}}\|_{\infty}) \exp(-\delta_1 t) \\ &+ \bar{\gamma}_1 \epsilon \max_{0 \leq s \leq t} (\|u_2^{\text{av}}(s)\|_{\infty}), \quad \forall t \geq 0, \end{aligned} \quad (85)$$

para alguma constante adequada $\bar{\gamma}_1 > \max(|H_{11}^1|k_1, k_1)$.

Para o jogador P_2 , o sistema médio de malha fechada (78) e (80)–(82) satisfaz todas as hipóteses **(H1)** a **(H7)** do Teorema do Pequeno Ganho [15, Teorema 8.2, p. 205] para o laço EDP-EDO parabólico com $n = 1$, $p(x) = 1$, $r(x) = D_2^2$, $q(x) = 0$, $F(\bar{G}_2^{\text{av}}(t), u_2(x, t), v(t)) = H_{22}^2 k_2 \bar{G}_2^{\text{av}}(t) + \epsilon v(t)$, $v(t) = H_{22}^2 k_2 \phi_2^{\text{av}}(1, t)$, $g(x, \bar{G}_2^{\text{av}}(t), u_2(x, t)) = 0$, $f(x, t) = 0$, $\varphi_0(d(t), u_2(x, t), \bar{G}_2^{\text{av}}(t)) = b_1 u_2(0, t)$, $b_1 < 0$, $b_2 = 1$, $\varphi_1(d(t), u_2(x, t), \bar{G}_2^{\text{av}}(t)) = k_2 \bar{G}_2^{\text{av}}(t) + \epsilon d(t)$, $d(t) = k_2 \phi_2^{\text{av}}(1, t)$, $a_1 = 1$, $a_2 = 0$, $L = |H_{22}^2|k_2$, $K_0 = |b_1|$, $B_0 = C_0 = 0$, γ_0 é de ordem $\mathcal{O}(1)$, $K_1 = 0$, $B_1 = k_2$, $C_1 = \epsilon$, γ_1 é de ordem $\mathcal{O}(1)$ e $K_2 = B_2 = 0$. A Hipótese **(H6)** é assegurada com $M = 1$, $\gamma_3 = 0$, $\sigma = |H_{22}^2|k_2$ e $b_3 > 0$ sendo de ordem $\mathcal{O}(\epsilon)$, como pode ser facilmente verificado por meio da fórmula da variação das constantes (83), para $i = 2$, e da aplicação da desigualdade de Cauchy-Schwarz ao termo $\phi_2^{\text{av}}(1, t)$ na equação (67):

$$\begin{aligned} \phi_2^{\text{av}}(1, t) &\leq |H_{21}^2| \left(\int_0^1 D_1^2 d\tau \right)^{\frac{1}{2}} \times \left(\int_0^1 [u_1^{\text{av}}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}} \\ &\leq k_H D_1 \left(\int_0^1 [u_1^{\text{av}}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}}, \end{aligned} \quad (86)$$

com o mesmo $k_H > 0$ definido logo após (84). Portanto, segue-se que a condição de pequeno ganho em [15, Inequality (8.3.24)] mantém desde que $0 < \epsilon < 1$ seja suficientemente pequeno. Por isso, [15, Teorema 8.2, p. 205] nos permite concluir que existem constantes $\delta_2, \Delta_2 > 0$ tais que

$$\begin{aligned} |\bar{G}_2^{\text{av}}(t)| + \|u_2^{\text{av}}(t)\|_{\infty} &\leq \Delta_2 (|\bar{G}_{2,0}^{\text{av}}| + \|u_{2,0}^{\text{av}}\|_{\infty}) \exp(-\delta_2 t) \\ &+ \bar{\gamma}_2 \epsilon \max_{0 \leq s \leq t} (\|u_1^{\text{av}}(s)\|_{\infty}), \quad \forall t \geq 0, \end{aligned} \quad (87)$$

para uma constante adequada $\bar{\gamma}_2 > \max(|H_{22}^2|k_2, k_2)$, $u_2^{\text{av}}(x, 0) = u_{2,0}^{\text{av}}$ e $\bar{G}_2^{\text{av}}(0) = \bar{G}_{2,0}^{\text{av}}$. Como as desigualdades (85) e (87) são semelhantes às encontradas em [15, Teorema 11.2, p. 269] — veja desigualdades (11.2.23) e (11.2.24) — podemos finalmente invocar [15, Teorema 11.5, p. 277], sob a condição de $0 < \epsilon < 1$ suficientemente pequeno, para concluir que

$$\begin{aligned} &|\bar{G}_1^{\text{av}}(t)| + |\bar{G}_2^{\text{av}}(t)| + \|u_1^{\text{av}}(t)\|_{\infty} + \|u_2^{\text{av}}(t)\|_{\infty} \leq \\ &\Delta (|\bar{G}_{1,0}^{\text{av}}| + |\bar{G}_{2,0}^{\text{av}}| + \|u_{1,0}^{\text{av}}\|_{\infty} + \|u_{2,0}^{\text{av}}\|_{\infty}) \exp(-\delta t), \quad \forall t \geq 0, \end{aligned} \quad (88)$$

para algum $\delta > 0$ e $\Delta > 0$. Portanto, concluímos que a origem do sistema médio de malha fechada (68)–(72) é exponencialmente estável sob a suposição de $0 < \epsilon < 1$ sendo suficientemente pequeno. Então, a partir de (48), (61) e (62), concluí-se os mesmos resultados na norma

$$\left(\sum_{i=1}^2 [\vartheta_i^{\text{av}}(t)]^2 + \int_0^{D_i} [u_i^{\text{av}}(x, t)]^2 dx \right)^{1/2} \quad (89)$$

já que H não é singular, ou seja, $|\vartheta_i^{\text{av}}(t)| \leq |H^{-1}| |\hat{G}^{\text{av}}(t)|$.

Conforme desenvolvido em [19], os próximos passos da prova seriam a aplicação da teoria da média para sistemas de dimensão infinita em [16, Sec. 2] e [19, Apêndice A], mostrando que as soluções periódicas satisfazem (74) para ω suficientemente grande, e então a conclusão da atratividade do Equilíbrio de Nash Θ^* de acordo com (75). Os conjuntos de resíduos finais para os erros $\theta_i(t) - \theta_i^*$ em (76) e (77) dependem de a_1 e $a_2 e^{D_2 \sqrt{\frac{\omega}{2}}}$ devido à amplitude dos pontilhados aditivos $S_i(t)$ em (30), para $i \in \{1, 2\}$. \square

5. SIMULAÇÕES DO DUOPÓLIO HETEROGÊNEO COM EDPS DE TRANSPORTE-CALOR

Para um exemplo de um jogo heterogêneo não-cooperativo com 2 jogadores que empregam a estratégia proposta de busca extremal para compensação de EDPS, revisita-se o exemplo em [11] e considerando as seguintes funções de *payoff* (12) e (13) sujeito a EDPS de calor-transporte (1)–(4) e (5)–(8) com coeficientes de transporte-difusão distintos $D_1 = 30$ e $D_2 = 3$ nas decisões dos jogadores, $i \in \{1, 2\}$:

$$\begin{aligned} J_1(\Theta(t)) &= -5 \Theta_1^2(t) + 5 \epsilon \Theta_1(t) \Theta_2(t) + \\ &+ 250 \Theta_1(t) - 150 \Theta_2(t) - 3000, \end{aligned} \quad (90)$$

$$\begin{aligned} J_2(\Theta(t)) &= -5 \Theta_2^2(t) + 5 \epsilon \Theta_1(t) \Theta_2(t) + \\ &- 150 \Theta_1(t) + 150 \Theta_2(t) + 2500, \end{aligned} \quad (91)$$

que, de acordo com (18), produz o único equilíbrio de Nash

$$\Theta_1^* = \theta_1^* = \frac{100 + 30\epsilon}{4 - \epsilon^2}, \quad \Theta_2^* = \theta_2^* = \frac{60 + 50\epsilon}{4 - \epsilon^2}. \quad (92)$$

Para atingir (92), os jogadores implementam a estratégia de otimização em tempo real baseada no não-conhecimento do modelo da Seção 4 atuando através das EDPS de transporte-calor (veja Fig. 2).

Para fins de comparação, exceto para os EDPS de transporte de calor nos sinais de entrada dos jogadores, a planta e os parâmetros do controlador foram escolhidos de forma semelhante ao [6] nos exemplos de simulação: $\epsilon = 1$, $a_1 = 0.075$, $a_2 = 0.05$, $k_1 = 2$, $k_2 = 5$, $\omega_1 = 26.75$ rad/s, $\omega_2 = 22$ rad/s e $\theta_1(0) = \hat{\theta}_1(0) = 50$, $\theta_2(0) = \hat{\theta}_2(0) = \theta_2^* = 110/3$.

Além disso, as constantes de tempo dos filtros do controlador por fronteira foram definidas como $c_1 = c_2 = 100$.

As Figuras. 3(a) e 3(b) mostram que, mesmo em um cenário adversário de forte acoplamento entre os jogadores com $\epsilon = 1$, o esquema baseado em controle por fronteira proposto é capaz de garantir a alcance de uma pequena vizinhança do equilíbrio de Nash compensando simultaneamente o efeito das EDPs de transporte-calor no jogo heterogêneo e não-cooperativo. Isso sugere que nossa análise de estabilidade pode ser conservadora e a hipótese teórica $0 < \epsilon < 1$ pode ser relaxada devido ao desempenho do sistema de malha fechada.

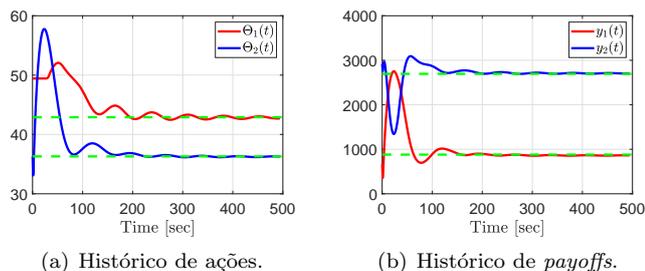


Figura 3. (a) histórico de ações e (b) histórico de payoffs para P_1 e P_2 , para $\epsilon = 1$. As linhas tracejadas representam os valores do Equilíbrio de Nash, $\Theta_1^* = 43.33$ e $\Theta_2^* = 36.67$ (with $J_1(\Theta^*) = 889$ and $J_2(\Theta^*) = 2722$).

6. CONCLUSÕES

Neste artigo introduz-se uma abordagem baseada no não-conhecimento do modelo via busca extremal e controle por fronteiras para calcular de forma distribuída o equilíbrio de Nash em jogos não-cooperativos do tipo duopólio com funções de *payoff* quadráticas desconhecidas e com jogadores agindo através de EDPs heterogêneas de transporte-calor. Um jogador pode atingir de forma estável seu equilíbrio de Nash medindo apenas o valor de sua função de *payoff* (nenhuma outra informação sobre o jogo é necessária). A estabilidade e convergência local são garantidas por meio da teoria da média em dimensões infinitas e uma análise de pequeno ganho para os ramos formados por EDO-EDP. Simulações realizadas para um jogo de dois jogadores sob EDOs hiperbólico-parabólicas com coeficientes de transporte-difusão distintos suportam os resultados teóricos.

REFERÊNCIAS

- [1] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, Cambridge–Massachusetts, 1991.
- [2] Z. Han, D. Niyato, W. Saad and T. Başar. *Game Theory for Next Generation Wireless and Communication Networks: Modeling, Analysis, and Design*. Cambridge University Press, 2019.
- [3] S. Amina, G. A. Schwartz and S. S. Sastry. *Security of interdependent and identical networked control*. *Automatica*, 49: 186–192, 2013.
- [4] J. F. Nash. *Noncooperative games*. *Ann. Math.*, 54: 286–295, 1951.
- [5] T. Başar, G. J. Olsder. *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, Philadelphia, 1999.

- [6] P. Frihauf, M. Krstić and T. Başar. *Nash equilibrium seeking in noncooperative games*. *IEEE Trans. Autom. Control* 57: 1192–1207, 2012.
- [7] M. Krstić and H. H. Wang. *Stability of extremum seeking feedback for general dynamic systems*. *Automatica*, 36: 595–601, 2000.
- [8] T. R. Oliveira, V. H. P. Rodrigues, M. Krstić and T. Başar. *Nash equilibrium seeking with arbitrarily delayed player actions*. *Conference on Decision and Control (CDC)*, 150–155, Jeju Island, KR, 2020.
- [9] M. Krstić. *Delay Compensation for Nonlinear, Adaptive, and PDE Systems*. Birkhauser, Boston, 2009.
- [10] T. R. Oliveira, M. Krstić and D. Tsubakino. *Extremum seeking for static maps with delays*. *IEEE Trans. Autom. Control*, 62: 1911–1926, 2017.
- [11] T. R. Oliveira, V. H. P. Rodrigues, M. Krstić and T. Başar. *Nash equilibrium seeking with players acting through heat PDE dynamics*. *American Control Conference (ACC)*, 684–689, New Orleans, USA, 2021.
- [12] M. Krstić and A. Smyshlyaev. *Boundary Control of PDEs: A Course on Backstepping Designs*. SIAM, 2008.
- [13] E. O. Roxin. *Differential games with partial differential equations*. In: Hagedorn P., Knobloch H. W., Olsder G. J. (eds). *Differential Games and Applications*. *Lect. Notes Control. Inf. Sci*, vol. 3, 157–168, Springer, Berlin, Heidelberg, 1977.
- [14] M. Lewicka and J. J. Manfredi. *Game theoretical methods in PDEs*. *Bollettino dell’Unione Matematica Italiana* 7, 211–216, 2014.
- [15] I. Karafyllis and M. Krstić. *Input-to-State Stability for PDEs*. Springer, Springer Nature Switzerland AG, 2018.
- [16] J. K. Hale and S. M. V. Lunel. *Averaging in infinite dimensions*. *Journal of Integral Equations and Applications* 2, 463–494, 1990.
- [17] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.
- [18] A. Ghaffari, M. Krstić and D. Nesic. *Multivariable Newton-based extremum seeking*. *Automatica*, 48: 1759–1767, 2012.
- [19] J. Feiling, S. Koga, M. Krstić and T. R. Oliveira. *Gradient extremum seeking for static maps with actuation dynamics governed by diffusion PDEs*. *Automatica*, 95: 197–206, 2018.
- [20] Z. Artstein. *Linear systems with delayed controls: A reduction*. *IEEE Trans. Autom. Control*, 27: 869–879, 1982.
- [21] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge Press, 1985.
- [22] T. R. Oliveira, J. Feiling, S. Koga and M. Krstić. *Multivariable extremum seeking for PDE dynamic systems*. *IEEE Trans. Autom. Control*: 65, 4949–4956, 2020.