

## Jogos Não-Cooperativos com Equações Diferenciais Parciais de Difusão <sup>★</sup>

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 quadratic noncooperative (duopoly) games with player actions subject to diffusion (heat) PDEs (Partial Differential Equations) with distinct diffusion coefficients and each player having access only to his own payoff value. The proposed approach employs extremum seeking, with sinusoidal perturbation signals employed to estimate the Gradient (first derivative) and Hessian (second derivative) of unknown quadratic functions. In order to compensate distinct diffusion 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) parabolic PDE-ODE loop as well as averaging theory in infinite dimensions 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 on an example of a two-player game under heat PDEs.

**Resumo:** Este artigo propõe uma estratégia baseada no não conhecimento do modelo para o estabelecimento do equilíbrio de Nash em jogos quadráticos não-cooperativos (duopólio) com ações de jogadores sujeitas a EDPs (Equações Diferenciais Parciais) de difusão (calor). Os coeficientes das equações diferenciais de difusão são distintos, com cada jogador tendo acesso apenas ao seu próprio *payoff*. A abordagem proposta emprega busca extremal, com perturbações senoidais empregadas para estimar o Gradiente (primeira derivada) e a Hessiana (segunda derivada) de funções quadráticas desconhecidas. A fim de compensar processos de difusão distintos nas entradas dos dois jogadores, emprega-se um controle de fronteira com estimativas baseadas em médias. Aplica-se uma análise de pequenos ganhos para a malha ISS composta por equações diferenciais parciais parabólicas e equações diferenciais ordinárias (EDP-EDO), bem como a teoria de média em dimensões infinitas a fim de se obter resultados de convergência para uma pequena vizinhança do equilíbrio de Nash. Quantificamos o tamanho desses conjuntos residuais e ilustramos os resultados teóricos numericamente em um exemplo de um duopólio sujeito à EDPs de calor.

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

**Palavras-chaves:** Busca extremal; Equilíbrio de Nash; Jogos Não-cooperativos; Equações do Calor; Realimentação por Preditor; Teoria da média em dimensões infinitas.

---

## 1. INTRODUÇÃO

A teoria dos jogos fornece uma estrutura teórica para representar situações sociais entre concorrentes com a utilização de modelos matemáticos de interação estratégica entre esses tomadores de decisões racionais (Fudenberg and Tirole, 1991). O emprego da teoria dos jogos para projetar, modelar e otimizar sistemas de engenharia, comportamentais e financeiros tornam este tópico de pesquisa uma ferramenta extremamente importante com uma ampla gama de aplicações (Han et al., 2019; Başar and Zaccour, 2018).

Os jogos podem ser classificados em duas categorias: *cooperativos* e *não-cooperativos* (Başar and Olsder, 1999). Um jogo é cooperativo se os jogadores forem capazes de formar compromissos vinculativos resultando em um lucro coletivo. Um jogo é não-cooperativo se os jogadores não puderem formar alianças, forçando os jogadores a estimarem as ações individuais de outros tomadores de decisões, *payoffs* e a analisarem o *equilíbrio de Nash* (Nash, 1951). O Equilíbrio de Nash assegura a impossibilidade de que qualquer jogador possa aumentar seu *payoff* tomando decisões unilateralmente (Başar and Olsder, 1999).

O desenvolvimento de algoritmos para alcançar a convergência ao equilíbrio de Nash tem sido foco de pesquisadores por várias décadas. Em particular, os autores Frihauf et al. (2012) estudaram o problema de alcance do Equilíbrio de Nash, em tempo real, em jogos estáticos não-cooperativos com  $N$  jogadores empregando uma abordagem baseada no não-conhecimento do modelo. Ao utilizar a busca extremal com perturbações sinusoidais (Krstić and Wang, 2000), os jogadores alcançam de maneira local e estável suas estratégias de Nash sem a necessidade de qualquer informação do modelo.

Por outro lado, atrasos são uns dos fenômenos mais comuns que surgem na prática da engenharia e indústria, envolvendo problemas em áreas como: virtualização de redes, redes definidas por *software*, computação em nuvem, internet das coisas, redes verdes e segurança (Han et al., 2019; Amina et al., 2013). Portanto, a motivação para empregar a busca extremal para otimização desses processos com atrasos, comumente modelados com a estrutura teórica dos jogos, é clara e justificada. Felizmente, a referência (Oliveira et al., 2020a) forneceu uma contribuição para o problema da Busca Extremal em jogos não-cooperativos jogadores sujeitos à atrasos constantes usando um preditor por realimentação (Krstić, 2009; Oliveira et al., 2017).

Este artigo generaliza os resultados da Busca Extremal obtidos em Oliveira et al. (2020a) para uma classe mais ampla de sistemas de dimensão infinita governados por Equações Diferenciais Parciais (EDPs) em vez de atrasos ou os resultados clássicos para Equações Diferenciais Ordinárias (EDOs) (Başar and Olsder, 1999). Isso é particularmente desafiador, uma vez que na literatura para compensação de EDP de múltiplas entradas (Krstić and Smyshlyaev, 2008), a maioria dos autores considera uma abordagem centralizada para o projeto de controle de fronteira (todos os vetores que multiplicam as entradas de controle precisam ser conhecidos), enquanto os jogos são descentralizados. Este artigo fornece uma contribuição

\* O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. O primeiro e segundo 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).

para a literatura sobre jogos não-cooperativos com dois jogadores atuando através de EDPs do calor, que podem representar comportamentos semelhantes aos humanos na teoria dos jogos, como movimentações aleatórias (decisões) e movimento browniano ou mesmo estarem conectadas com pequenas variantes da equação de Black-Scholes nos mercados financeiros (Lewicka and Manfredi, 2014). Nesse cenário, pode-se desenvolver um resultado para jogos com EDPs de calor com coeficientes de difusão distintos na malha de ação de cada jogador, onde os jogadores estimam apenas as entradas diagonais da matriz de Hessiana com o auxílio de seus próprios *payoffs* nesse problema de problema de busca, em tempo real, do Equilíbrio de Nash. Portanto, pode-se de dominar termos fora da diagonal, suficientemente pequenos, utilizando-se a teoria de pequenos ganhos KK:2018 para o sistema médio.

A análise de estabilidade apresenta uma sequência de etapas utilizando teoria da média em dimensão infinita (Hale and Lunel, 1990), funcional de Lyapunov Oliveira et al. (2020b), Teorema de pequenos ganho para sistemas dinâmicos formados por cascatas de EDO-EDP que apresentem propriedades ISS (Karafyllis and Krstić, 2018) no esquema não-cooperativo com o intuito de provar a estabilidade do sistema de malha fechada. Prova-se o alcance de uma pequena vizinhança do equilíbrio de Nash, mesmo na presença de EDPs de calor. Um exemplo numérico ilustra o resultados teóricos.

## 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,  $\|\cdot\| = \|\cdot\|_{\mathcal{L}_2([0, D])}$  (Krstić, 2009). Como definido em Khalil (2002), 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  $\epsilon$  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 Khalil (2002) e Karafyllis and Krstić (2018), respectivamente. O operador de Laplace é representado por  $\mathcal{L}$  enquanto o de convolução é dado por  $*$ .

## 3. DUOPÓLIO COM LUCROS QUADRÁTICOS E EDPS NAS DECISÕES

Em um duopólio, a otimalidade dos lucros (saídas) dos jogadores  $J_1$  e  $J_2$ , 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  $\theta_2(t) \in \mathbb{R}$ . Existe um perfil de sinais de entrada  $\theta^* = [\theta_1^*, \theta_2^*]^T \in \mathbb{R}^2$ , de modo que a estratégia de cada jogador seja uma entrada ideal para a função de lucro do outro jogador  $J_1$  e  $J_2$ . Quando cada jogador encontra sua entrada ideal,  $\theta(t) = \theta^*$ , diz-se que o equilíbrio de Nash é alcançado (Fudenberg and Tirole, 1991). Apesar do grande número de publicações sobre a Busca pelo Equilíbrio de Nash (Frihauf et al., 2012), o estudo do problema na presença de EDPs ainda está aberto.

Como mostrado nas Figuras 1 and Fig. 2, são consideradas equações do calor distintas (com atuação de Dirichlet) no vetor de ações dos jogo  $\theta(t) \in \mathbb{R}^2$ . Então, o vetor de atuação propagada  $\Theta(t) \in \mathbb{R}^2$  é dado por

$$\Theta_i(t) = \alpha_i(0, t), \quad \forall i \in \{1, 2\} \quad (1)$$

$$\partial_t \alpha_i(x, t) = \partial_{xx} \alpha_i(x, t), \quad x \in (0, D_i) \quad (2)$$

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

$$\alpha_i(D_i, t) = \theta_i(t), \quad (4)$$

onde  $\alpha_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}$  sendo conhecido cada comprimento de domínio  $D_i$ . A solução de (1)–(4) é dada por

$$\alpha_i(x, t) = \mathcal{L}^{-1} \left[ \frac{\cosh(x\sqrt{s})}{\cosh(D_i\sqrt{s})} \right] * \theta_i(t), \quad (5)$$

onde  $\mathcal{L}^{-1}(\cdot)$  representa a transformada inversa de Laplace e  $*$  o operador de convolução. Dada estas relações, define-se o *operador de difusão* para a EDP (1)–(4) com fronteira de entrada e medida dada por  $\mathcal{D} = \text{diag}\{D_1, D_2\}$  com

$$\mathcal{D}_i[\varphi(t)] = \mathcal{L}^{-1} \left[ \frac{1}{\cosh(D_i\sqrt{s})} \right] * \varphi(t), \quad \text{s.t. } \Theta(t) = \mathcal{D}[\theta(t)]. \quad (6)$$

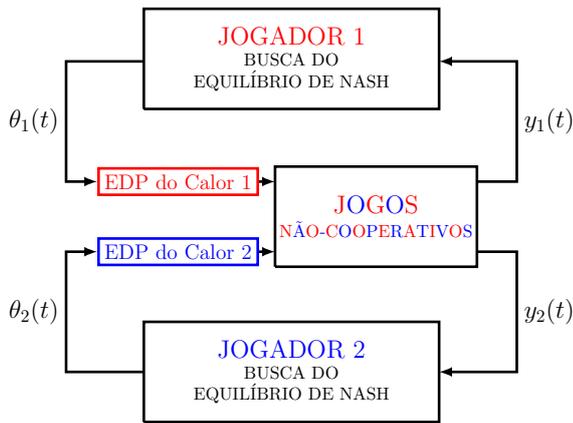


Figura 1. Diagrama em blocos da Busca pelo Equilíbrio de Nash em uma estrutura de mercado do tipo duopólio (dois jogadores,  $N = 2$ ) com ações propagadas através de EDPs do Calor.

Considera-se jogos nos quais a função de *payoff*  $y_i(t) = J_i(\mathcal{D}[\theta(t)]) = J_i(\Theta(t))$  de cada jogador é quadrática Başar and Olsder (1999), expressada como uma combinação das ações de todos jogadores propagados através de EDPs de calor distintas

$$J_1(\Theta(t)) = \frac{H_{11}^1}{2} \Theta_1^2(t) + \frac{H_{22}^1}{2} \Theta_2^2(t) + \epsilon H_{12}^1 \Theta_1(t) \Theta_2(t) + h_1^1 \Theta_1(t) + h_2^1 \Theta_2(t) + c_1, \quad (7)$$

$$J_2(\Theta(t)) = \frac{H_{11}^2}{2} \Theta_1^2(t) + \frac{H_{22}^2}{2} \Theta_2^2(t) + \epsilon H_{21}^2 \Theta_1(t) \Theta_2(t) + h_1^2 \Theta_1(t) + h_2^2 \Theta_2(t) + c_2, \quad (8)$$

onde  $J_1(\Theta), J_2(\Theta) : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $H_{jk}^i, h_j^i, c_i \in \mathbb{R}$  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, em primeiro lugar porque constituem aproximações de segunda ordem para outros tipos de funções de payoff não quadráticas e, em segundo lugar, porque são analiticamente tratáveis, levando a, em geral, soluções de equilíbrio de forma fechada que fornecem visão sobre as propriedades e

características do conceito de solução de equilíbrio em consideração (Başar and Olsder, 1999).

Para fins de completude, fornecemos aqui, em termos matemáticos, a definição do equilíbrio de Nash  $\Theta^* = [\Theta_1^*, \Theta_2^*]^T$  em um duopólio é:

$$J_1(\Theta_1^*, \Theta_2^*) \geq J_1(\Theta_1, \Theta_2^*) \text{ e } J_2(\Theta_1^*, \Theta_2^*) \geq J_2(\Theta_1^*, \Theta_2). \quad (9)$$

Portanto, nenhum jogador tem um incentivo para desviar unilateralmente sua ação de  $\Theta^*$ .

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

$$\begin{cases} H_{11}^1 \Theta_1^* + \epsilon H_{12}^1 \Theta_2^* + h_1^1 = 0 \\ \epsilon H_{21}^2 \Theta_1^* + H_{22}^2 \Theta_2^* + h_2^2 = 0 \end{cases}, \quad (10)$$

que pode ser escrito de forma matricial como

$$\begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^2 & H_{22}^2 \end{bmatrix} \begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix} = - \begin{bmatrix} h_1^1 \\ h_2^2 \end{bmatrix}. \quad (11)$$

Definindo-se a matrix Hessiana  $H$  e os vetores  $\Theta^*$  e  $h$  como

$$H := \begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^2 & H_{22}^2 \end{bmatrix}, \quad \Theta^* := \begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix}, \quad h := \begin{bmatrix} h_1^1 \\ h_2^2 \end{bmatrix}, \quad (12)$$

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

$$\begin{bmatrix} \Theta_1^* \\ \Theta_2^* \end{bmatrix} = - \begin{bmatrix} H_{11}^1 & \epsilon H_{12}^1 \\ \epsilon H_{21}^2 & H_{22}^2 \end{bmatrix}^{-1} \begin{bmatrix} h_1^1 \\ h_2^2 \end{bmatrix}. \quad (13)$$

Para mais detalhes, consultar (Başar and Olsder, 1999, Capítulo 4).

O *objetivo de controle* é projetar uma nova estratégia baseada na busca extremal para alcance do Equilíbrio de Nash em jogos (não)-cooperativos sujeitos a ações propagadas por distintas EDPs de calor.

Uma vez que o objetivo é encontrar a entrada ideal, e desconhecida,  $\Theta^*$  (e  $\theta^*$ ), define-se o erro de estimação

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^*, \quad \vartheta(t) = \hat{\Theta}(t) - \Theta^*, \quad (14)$$

onde os vetores  $\hat{\theta}(t)$  e  $\hat{\Theta}(t)$  são estimativas de  $\theta^*$  e  $\Theta^*$ . Para fazer (14) coerente com o otimizador do mapa estático  $\Theta^*$ , aplica-se o operador de difusão (6) a  $\tilde{\theta}$  em (14) e chega-se a

$$\vartheta_i(t) := \bar{\alpha}_i(0, t), \quad \forall i \in \{1, 2\} \quad (15)$$

$$\partial_t \bar{\alpha}_i(x, t) = \partial_{xx} \bar{\alpha}_i(x, t), \quad x \in (0, D_i) \quad (16)$$

$$\partial_x \bar{\alpha}_i(0, t) = 0 \quad (17)$$

$$\bar{\alpha}_i(D_i, t) = \tilde{\theta}_i(t), \quad (18)$$

onde  $\bar{\alpha}_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}$  e  $\vartheta(t) := \mathcal{D}[\tilde{\theta}(t)] = \hat{\Theta}(t) - \Theta^*$  é o erro de estimação  $\tilde{\theta}(t)$  propagado através do domínio de difusão. 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 (5), para uma entrada constante  $\theta = \theta_c$ , tem-se  $\mathcal{L}\{\theta_c\} = \theta_c/s$ , e aplicando-se o teorema do valor limite  $\lim_{t \rightarrow \infty} \mathcal{D}_i[\theta_{ci}] =$

$$\lim_{s \rightarrow 0} \left\{ \frac{\theta_{ci}}{\cosh(\sqrt{s}D_i)} \right\} = \theta_{ci}. \text{ No caso particular } \theta = \theta_c = \theta^*, \quad \Theta^* = \theta^*. \quad (19)$$

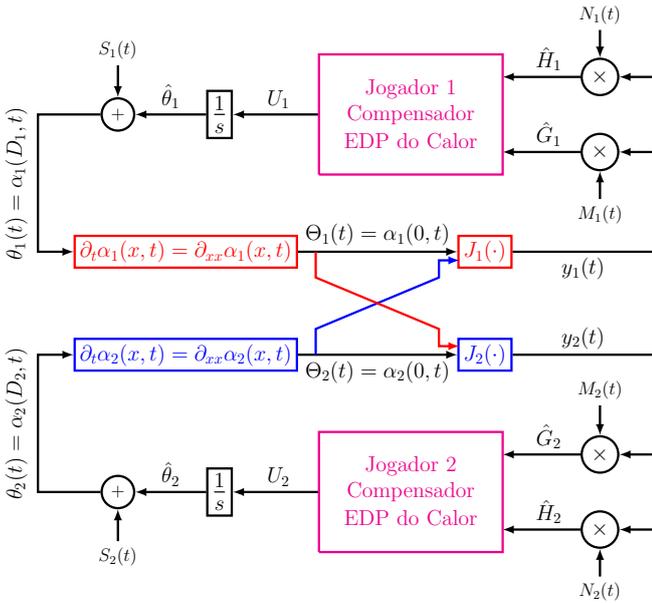


Figura 2. Diagrama em blocos ilustrando a Busca do Equilíbrio de Nash. Em magenta são apresentados os controladores de fronteira usados para compensar as EDPs de Calor no jogo não-cooperativo.

A Figura 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

$$\begin{cases} y_1(t) = J_1(\Theta(t)), \\ y_2(t) = J_2(\Theta(t)). \end{cases} \quad (20)$$

Os sinais de dither aditivos na presença de EDPs de Calor são definidos de acordo com (Feiling et al., 2018),

$$\begin{cases} S_1(t) = \frac{1}{2} a_1 e^{\sqrt{\frac{\omega_1}{2}} D_1} \sin\left(\omega_1 t + \sqrt{\frac{\omega_1}{2}} D_1\right) \\ \quad + \frac{1}{2} a_1 e^{-\sqrt{\frac{\omega_1}{2}} D_1} \sin\left(\omega_1 t - \sqrt{\frac{\omega_1}{2}} D_1\right), \\ S_2(t) = \frac{1}{2} a_2 e^{\sqrt{\frac{\omega_2}{2}} D_2} \sin\left(\omega_2 t + \sqrt{\frac{\omega_2}{2}} D_2\right) \\ \quad + \frac{1}{2} a_2 e^{-\sqrt{\frac{\omega_2}{2}} D_2} \sin\left(\omega_2 t - \sqrt{\frac{\omega_2}{2}} D_2\right) \end{cases}, \quad (21)$$

enquanto os multiplicativos são

$$\begin{cases} M_1(t) = \frac{2}{a_1} \sin(\omega_1 t) \\ M_2(t) = \frac{2}{a_2} \sin(\omega_2 t) \end{cases}, \quad (22)$$

com amplitudes  $a_1, a_2 > 0$  e frequências  $\omega_1 \neq \omega_2$ . Tais frequências  $\omega_i$  podem ser escolhidas como

$$\omega_i = \omega'_i \omega = \mathcal{O}(\omega), \quad i = 1 \text{ ou } 2, \quad (23)$$

onde  $\omega$  é uma constante positiva e  $\omega'_i$  é um número racional – uma escolha possível é dada em (Ghaffari et al., 2012).

Seguindo o paradigma de não compartilhamento de informações, apenas os elementos diagonais de  $H$  podem ser adequadamente recuperados, na média, pelos jogadores J1 ou J2. Dessa forma, os sinais  $N_1(t)$  e  $N_2(t)$  são definidos simplesmente como em (Ghaffari et al., 2012):

$$\begin{cases} N_1(t) = \frac{16}{a_1^2} \left( \sin^2(\omega_1 t) - \frac{1}{2} \right) \\ N_2(t) = \frac{16}{a_2^2} \left( \sin^2(\omega_2 t) - \frac{1}{2} \right). \end{cases} \quad (24)$$

Então, a versão média de

$$\begin{cases} \hat{H}_1(t) = N_1(t) y_1(t) \\ \hat{H}_2(t) = N_2(t) y_2(t) \end{cases} \quad (25)$$

é dada por

$$\begin{cases} \hat{H}_1^{\text{av}}(t) = [N_1(t) y_1(t)]_{\text{av}} = H_{11}^1 \\ \hat{H}_2^{\text{av}}(t) = [N_2(t) y_2(t)]_{\text{av}} = H_{22}^2 \end{cases}. \quad (26)$$

Considerando  $\hat{\theta}_1(t)$  e  $\hat{\theta}_2(t)$  como as estimativas de  $\theta_1^*$  e  $\theta_2^*$ , pode-se definir a partir de (14) as estimativas individuais do erro:

$$\begin{cases} \tilde{\theta}_1(t) = \hat{\theta}_1(t) - \theta_1^*, & \vartheta_1(t) = \hat{\Theta}_1(t) - \Theta_1^* \\ \tilde{\theta}_2(t) = \hat{\theta}_2(t) - \theta_2^*, & \vartheta_2(t) = \hat{\Theta}_2(t) - \Theta_2^* \end{cases}. \quad (27)$$

As estimativas do Gradiente das funções de *payoff* são dadas por

$$\begin{cases} \hat{G}_1(t) = M_1(t) y_1(t) \\ \hat{G}_2(t) = M_2(t) y_2(t) \end{cases}, \quad (28)$$

e que computando-se a média do sinal resultante, leva a

$$\begin{cases} \hat{G}_1^{\text{av}}(t) = H_{11}^1 \vartheta_1^{\text{av}}(t) + \epsilon H_{12}^1 \vartheta_2^{\text{av}}(t) \\ \hat{G}_2^{\text{av}}(t) = \epsilon H_{21}^2 \vartheta_1^{\text{av}}(t) + H_{22}^2 \vartheta_2^{\text{av}}(t) \end{cases}. \quad (29)$$

Adicionalmente, do diagrama de blocos na Figura 2,

$$\dot{\hat{\theta}}_i(t) = U_i(t), \quad \forall i \in \{1, 2\}, \quad (30)$$

e, conseqüentemente,

$$\dot{\hat{\theta}}_i(t) = U_i(t), \quad \forall i \in \{1, 2\}, \quad (31)$$

já que  $\dot{\hat{\theta}}(t) = \dot{\hat{\theta}}(t) + \theta^*$  é uma constante. Tomando-se a derivada temporal de (15)–(18) e com a ajuda de (14) e (31), a *dinâmica do erro propagado* é escrita como

$$\dot{\vartheta}_i(t) = u_i(0, t), \quad \forall i \in \{1, 2\} \quad (32)$$

$$\partial_t u_i(x, t) = \partial_{xx} u_i(x, t), \quad x \in (0, D_i) \quad (33)$$

$$\partial_x u_i(0, t) = 0 \quad (34)$$

$$u_i(D_i, t) = U_i(t), \quad (35)$$

onde  $u_i : [0, D_i] \times \mathbb{R}_+ \rightarrow \mathbb{R}$  e  $u_i(x, t) := \partial_t \tilde{\alpha}_i(x, t)$ .

Portanto, de (29) e (32)–(35), é possível encontrar uma forma compacta para a estimativa do gradiente média geral como

$$\hat{G}^{\text{av}}(t) = H \vartheta^{\text{av}}(t), \quad (36)$$

$$\hat{G}^{\text{av}}(t) = H \vartheta^{\text{av}}(t) = HD[U^{\text{av}}(t)], \quad (37)$$

na qual a Hessiana  $H$  é dada pela equação (12),  $\vartheta^{\text{av}}(t) := [\vartheta_1^{\text{av}}(t), \vartheta_2^{\text{av}}(t)]^T \in \mathbb{R}^2$ ,  $\hat{G}^{\text{av}}(t) := [\hat{G}_1^{\text{av}}(t), \hat{G}_2^{\text{av}}(t)]^T \in \mathbb{R}^2$  e  $U^{\text{av}}(t) := [U_1^{\text{av}}(t), U_2^{\text{av}}(t)]^T \in \mathbb{R}^2$  são as versões média 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 (políticas) de controle para cada jogador, a fim de atingir uma pequena vizinhança do ponto de Equilíbrio de Nash. Com este intuito, utiliza-se uma estratégia de busca extremal baseada no controle de fronteira para compensar o operador de difusão  $\mathcal{D}[\cdot]$  em (37) devido aos múltiplas e distintas EDPs de calor nas ações dos jogadores. Basicamente, as leis de controle devem ser

capazes de garantir a estabilização exponencial de  $\hat{G}^{av}(t)$  e, consequentemente, de  $\vartheta^{av}(t) = \hat{\Theta}^{av}(t) - \Theta^*$ . De (36), é claro que, se  $H$  é inversível,  $\vartheta^{av}(t) \rightarrow 0$  com  $\hat{G}^{av}(t) \rightarrow 0$ . Portanto, a convergência de  $\vartheta^{av}(t)$  para a origem resulta na convergência de  $\Theta(t)$  para uma pequena vizinhança de  $\Theta^*$  em (9) via teoria da média Hale and Lunel (1990).

#### 4. DUOPÓLIO NÃO-COOPERATIVO COM EDPS DO CALOR

Como mencionado anteriormente, no jogo não-cooperativo sujeito à dinâmica do calor proposto neste artigo o objetivo da busca extremal é alcançar o equilíbrio de Nash  $\Theta^* = \theta^*$ , ver (19), sem o compartilhamento de qualquer informação entre os jogadores. Cada jogador precisa medir apenas o seu próprio valor de *payoff* (20).

##### 4.1 Controle de fronteira de EDPs descentralizado usando apenas estimativas de termos diagonais da Hessiana

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

Inspirando-se em (Oliveira et al., 2020a), onde se considera a busca extremal para a compensação de EDPs mas não para jogos, propõe-se as seguintes leis de adaptação baseadas em fronteira,  $\hat{\theta}_i(t) = U_i(t)$ ,  $i \in \{1, 2\}$ :

$$\begin{cases} \dot{U}_1(t) = -c_1 U_1(t) \\ \quad + c_1 k_1 \left( \hat{G}_1(t) + \hat{H}_1(t) \int_0^{D_1} (D_1 - \tau) u_1(\tau, t) d\tau \right), \\ \dot{U}_2(t) = -c_2 U_2(t) \\ \quad + c_2 k_2 \left( \hat{G}_2(t) + \hat{H}_2(t) \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \right) \end{cases}, \quad (38)$$

com constantes  $k_1 > 0$ ,  $k_2 > 0$ ,  $c_1 > 0$  e  $c_2 > 0$ , para compensar as EDPs do Calor em (32)–(35).

Como discutido em (Feiling et al., 2018, Remark 2), a lei de controle por fronteira (38) pode ser reescrita como

$$\begin{aligned} \dot{U}_i(t) = & -c_i U_i(t) \\ & + c_i k_i \left[ \hat{G}_i(t) + \hat{H}_i(t) \left( \hat{\theta}_i(t) - \Theta_i(t) + a_i \sin(\omega_i t) \right) \right], \end{aligned} \quad (39)$$

usando as equações de difusão  $\partial_t \alpha_i(x, t) = \partial_{xx} \alpha_i(x, t)$ ,  $\forall i \in \{1, 2\}$ , e a integração por partes, associadas com (1)–(4), (14 e lembrando-se que  $\vartheta_i + a_i \sin(\omega_i t) = \Theta_i(t) - \Theta_i^*$ , análogo a (Feiling et al., 2018, Eq. (25)).

##### 4.2 Propriedades ISS para EDPs

Por uma questão de simplicidade, assume-se que  $c_1, c_2 \rightarrow +\infty$  em (38), levando a seguinte expressão:

$$U_i(t) = k_i \left( \hat{G}_i(t) + \hat{H}_i(t) \int_0^{D_i} (D_i - \tau) u_i(\tau, t) d\tau \right). \quad (40)$$

Relembrando-se (32)–(35), o sistema de malha fechada de dimensão infinita (37) e (40) em sua versão média pode ser escrito na forma de EPD, como

$$\dot{\hat{G}}^{av}(t) = H u^{av}(0, t), \quad (41)$$

$$\partial_t u^{av}(x, t) = D^{-2} \partial_{xx} u^{av}(x, t), \quad x \in (0, 1), \quad (42)$$

$$\partial_x u^{av}(0, t) = 0, \quad (43)$$

$$u^{av}(1, t) = U^{av}(t), \quad (44)$$

com  $D = \text{diag}\{D_1, D_2\}$ .

Na *abordagem por redução* (Artstein, 1982) (ou atribuição de espectro finito), utiliza-se a transformação (para  $i \in \{1, 2\}$ ):

$$\begin{aligned} \bar{G}_i^{av}(t) &= \hat{G}_i^{av}(t) + \sum_{j=1}^2 \epsilon_{ij}^i H_{ij}^i \int_0^{D_j} (D_j - \tau) u_j^{av}(\tau, t) d\tau \\ &= \hat{G}_i^{av}(t) + \sum_{j=1}^2 \epsilon_{ij}^i H_{ij}^i \int_0^1 D_j^2 (1 - \xi) u_j^{av}(\xi, t) d\xi \end{aligned} \quad (45)$$

onde  $\int_0^{D_j} (D_j - \tau) u_j^{av}(\tau, t) d\tau = \int_0^1 D_j^2 (1 - \xi) u_j^{av}(\xi, t) d\xi$ , enquanto  $\epsilon_{11}^1 = \epsilon_{22}^2 = 1$  e  $\epsilon_{12}^1 = \epsilon_{21}^2 = \epsilon$ .

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

$$\dot{\bar{G}}^{av}(t) = H U^{av}(t). \quad (46)$$

Agora, depois de adicionar e subtrair os próximos termos em azul e vermelho em (40), esta pode ser reescrita como:

$$\begin{aligned} U_1(t) &= k_1 \left( \hat{G}_1(t) + \hat{H}_1(t) \int_0^{D_1} (D_1 - \tau) u_1(\tau, t) d\tau \right. \\ &\quad \left. + \epsilon H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \right) \\ &\quad - k_1 \epsilon H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau, \end{aligned} \quad (47)$$

$$\begin{aligned} U_2(t) &= k_2 \left( \hat{G}_2(t) + \hat{H}_2(t) \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \right. \\ &\quad \left. + \epsilon H_{21}^2 \int_0^{D_1} (D_1 - \tau) u_1(\tau, t) d\tau \right) \\ &\quad - k_2 \epsilon H_{21}^2 \int_0^{D_1} (D_1 - \tau) u_1(\tau, t) d\tau, \end{aligned} \quad (48)$$

cujas versão média compacta é

$$U^{av}(t) = K \bar{G}^{av}(t) + \epsilon K \phi^{av}(D, t), \quad (49)$$

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

$$\begin{aligned} \phi(D, t) &:= - \begin{bmatrix} H_{12}^1 \int_0^{D_2} (D_2 - \tau) u_2(\tau, t) d\tau \\ H_{21}^2 \int_0^{D_1} (D_1 - \tau) u_1(\tau, t) d\tau \end{bmatrix}, \\ \phi(1, t) &:= - \begin{bmatrix} H_{12}^1 \int_0^1 D_2^2 (1 - \xi) u_2(\xi, t) d\xi \\ H_{21}^2 \int_0^1 D_1^2 (1 - \xi) u_1(\xi, t) d\xi \end{bmatrix}. \end{aligned} \quad (50)$$

Então, é possível encontrar uma forma compacta para o jogo médio a partir das equações (46) e (49),

$$\dot{\bar{G}}^{av}(t) = H K \bar{G}^{av}(t) + \epsilon H K \phi^{av}(1, t), \quad (51)$$

$$\partial_t u^{av}(x, t) = D^{-2} \partial_{xx} u^{av}(x, t), \quad x \in (0, 1), \quad (52)$$

$$\partial_x u^{av}(0, t) = 0, \quad (53)$$

$$u^{av}(1, t) = K \bar{G}^{av}(t) + \epsilon K \phi^{av}(1, t). \quad (54)$$

De (51), se  $HK$  é Hurwitz, é claro que a EDO que representa a variável de estado  $\bar{G}^{av}(t)$  é exponencialmente ISS (Karafyllis and Krstić, 2018) com respeito ao estado  $u(x, t)$  da EDP por meio da função  $\phi^{av}(1, t)$ . Além do mais, o subsistema EDP (52) é ISS (Karafyllis and Krstić, 2018) com respeito a  $\bar{G}^{av}(t)$  na condição de fronteira  $u^{av}(1, t)$ .

### 4.3 Análise de Estabilidade

Nesta seção, mostra-se que a *loop* parabólico EDP-EDO (51)–(54) contém um pequeno parâmetro  $\epsilon$  que pode levar a estabilidade do sistema de malha fechada si o mesmo for suficientemente pequeno. Dessa forma, assume-se a seguinte hipótese para o jogo não-cooperativo (Frihauf et al., 2012).

**Hipótese 1.** A matriz Hessiana  $H$  dada por (12) é estritamente diagonal dominante, ou seja,

$$\begin{cases} |\epsilon H_{12}^1| < |H_{11}^1| \\ |\epsilon H_{21}^2| < |H_{22}^2| \end{cases} \quad (55)$$

Pela Hipótese 1, o Equilíbrio de Nash,  $\Theta^*$ , existe e é único, pois matrizes estritamente diagonais dominantes são não-singulares pelo Teorema de Levy-Desplanques (Horn and Johnson, 1985).

O próximo teorema resume as propriedades de estabilidade e convergência da busca extremal no sistema de malha fechada de um jogo não-cooperativo de dois jogadores com EDPs do Calor.

**Teorema 1.** Considere o sistema de malha fechada (41)–(44) sob a Hipótese 1, múltiplas EDPs de calor (1)–(4) com distintos coeficientes de difusão  $D_1$  e  $D_2$  em um jogo do tipo duopólio com funções de payoff (7) e (8) e leis de controle  $U_i(t)$  definidas em (38). Existem constantes  $c_1 > 0$ ,  $c_2 > 0$  e  $\omega > 0$  suficientemente grandes bem como  $\epsilon > 0$  suficientemente pequena tal que o sistema de malha fechada com estado  $\vartheta_i(t)$ ,  $u_i(x, t)$ ,  $\forall i \in \{1, 2\}$ , possui apenas uma solução periódica exponencialmente estável em  $t$  de período  $\Pi := 2\pi \times \text{LCM}\{1/\omega_i\}$ , com  $\omega_i$  em (23) de ordem  $\mathcal{O}(\omega)$ , denotada por  $\vartheta_i^\Pi(t)$ ,  $u_i^\Pi(x, t)$  e satisfazendo,  $\forall t \geq 0$ :

$$\left( \sum_{i=1}^2 [\vartheta_i^\Pi(t)]^2 + \int_0^{D_i} [u_i^\Pi(x, t)]^2 dx \right)^{1/2} \leq \mathcal{O}(1/\omega). \quad (56)$$

Além do mais,

$$\limsup_{t \rightarrow +\infty} |\Theta(t) - \Theta^*| = \mathcal{O}(|a| + 1/\omega), \quad (57)$$

$$\limsup_{t \rightarrow +\infty} |\theta(t) - \theta^*| = \mathcal{O}\left(|a|e^{\max(D_i)\sqrt{\omega/2}} + 1/\omega\right), \quad (58)$$

onde  $a = [a_1 \ a_2]^T$  e  $\theta^* = \Theta^*$  é o único (e desconhecido) Equilíbrio de Nash dado por (13).

**Prova.** A prova do Teorema segue etapas semelhantes às empregadas para provar os resultados sobre a busca extremal na presença de atrasos (Oliveira et al., 2017). Nesse sentido, simplesmente apontaremos as principais diferenças para o caso dos jogos (não a busca clássica por extremos), em vez de fornecer a prova detalhadamente.

Primeiro, considere a representação parabólica EDP-EDO equivalente (51)–(54) reescrita para cada Jogador  $i$ ,  $i \in \{1, 2\}$ :

$$\dot{\bar{G}}_i^{av}(t) = H_{ii}^i k_i \bar{G}_i^{av}(t) + \epsilon H_{ii}^i k_i \phi_i^{av}(1, t), \quad (59)$$

$$\partial_t u_i^{av}(x, t) = D_i^{-2} \partial_{xx} u_i^{av}(x, t), \quad x \in (0, 1), \quad (60)$$

$$\partial_x u_i^{av}(0, t) = 0, \quad (61)$$

$$u_i^{av}(1, t) = k_i \bar{G}_i^{av}(t) + \epsilon k_i \phi_i^{av}(1, t). \quad (62)$$

Portanto, o sistema médio de malha fechada 59)–(62) satisfaz todas as hipóteses (H1)–(H7) do Teorema dos Pequenos Ganhos (Karafyllis and Krstić, 2018, Teorema 8.2, p. 205) para o *loop* parabólico EDP-EDO com  $p(z) = 1$ ,  $r(z) = D_i^2$ ,  $q(z) = 0$ ,  $F(\bar{G}_i^{av}, u, 0) = H_{ii}^i k_i \bar{G}_i^{av} + \epsilon H_{ii}^i k_i \phi_i^{av}(1)$ ,  $g(x, \bar{G}_i^{av}, u) = 0$ ,  $f(x, t) = 0$ ,  $\varphi_0(0, u_i, \bar{G}_i^{av}) = b_1 u_i(0, t)$ ,  $b_1 < 0$ ,  $b_2 = 1$ ,  $\varphi_1(0, u, \bar{G}_i^{av}) = k_i \bar{G}_i^{av} + \epsilon k_i \phi_i^{av}(1)$ ,  $a_1 = 1$ ,  $a_2 = 0$ ,  $L = \max(|H_{ii}^i| k_i, \frac{1}{\sqrt{3}} \epsilon |H_{ii}^i| k_i k_H D_j^2)$ ,  $K_0 = 1$ ,  $B_0 = C_0 = 0$ ,  $\gamma_0$  é de ordem  $\mathcal{O}(\epsilon)$ ,  $K_1 = \frac{1}{\sqrt{3}} \epsilon k_i k_H D_j^2$ ,  $B_1 = k_i$ ,  $C_1 = 0$ ,  $\gamma_1$  é de ordem  $\mathcal{O}(1)$ ,  $K_2 = B_2 = 0$  e  $i \neq j$ . A constante  $k_H > 0$  a seguir, logo após a desigualdade (63). A Hipótese (H6) é satisfeita com  $M = 1$ ,  $\sigma = |H_{ii}^i| k_i$  como pode-se verificar rapidamente por meio da fórmula da variação das constantes

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

e da aplicação da desigualdade de Cauchy-Schwarz ao termo  $\phi_i^{av}(1, t)$  na equação (50):

$$\begin{aligned} \phi_i^{av}(1, t) &\leq |H_{ij}^i| D_j^2 \left( \int_0^1 (1-\tau)^2 d\tau \right)^{\frac{1}{2}} \times \left( \int_0^1 [u_j^{av}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}} \\ &\leq \frac{1}{\sqrt{3}} k_H D_j^2 \left( \int_0^1 [u_j^{av}(\xi, t)]^2 d\xi \right)^{\frac{1}{2}}, \quad (63) \end{aligned}$$

uma vez 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  $\mathcal{O}(1)$ . Tem-se que a condição de baixo ganho (Karafyllis and Krstić, 2018, Inequality (8.3.24)):

$$\begin{aligned} \max(\gamma_0 K_0, \gamma_1 K_1) + \sigma^{-1} K_2 &< 1, \\ \gamma_3 \max(\gamma_0 B_0, \gamma_1 B_1) + \gamma_3 \sigma^{-1} B_2 &< 1 \quad (64) \end{aligned}$$

assegurada para  $0 < \epsilon < 1$  suficientemente pequeno. Portanto, se tal condição é satisfeita, então (Karafyllis and Krstić, 2018, Theorem 8.2, p. 205) permite que se assegure a existência de constantes  $\delta, \Delta > 0$  tal que para cada  $u_0^{av} \in C^0([0, 1])$ ,  $\bar{G}_0^{av} \in \mathbb{R}^n$ , a única solução generalizada do problema de valor-fronteira inicial, com  $u^{av}(x, 0) = u_0^{av}$  e  $\bar{G}^{av}(0) = \bar{G}_0^{av}$ , satisfaz a seguinte estimativa:

$$|\bar{G}^{av}(t)| + \|u^{av}(t)\|_\infty \leq \Delta (|\bar{G}_0^{av}| + \|u_0^{av}\|_\infty) \exp(-\delta t). \quad (65)$$

Então, de (36) e (45), conclui-se os mesmo resultados em norma  $\left( \sum_{i=1}^2 [\vartheta_i^{av}(t)]^2 + \int_0^{D_i} [u_i^{av}(x, t)]^2 dx \right)^{1/2}$  já que  $H$  é não singular, i.e.,  $|\vartheta_i^{av}(t)| \leq |H^{-1}| |\hat{G}^{av}(t)|$ . Como desenvolvido em Feiling et al. (2018), os próximos passos da prova poderiam ser a aplicação do teoria da média local para sistemas de dimensão infinita em (Hale and Lunel, 1990, Sec. 2) e (Feiling et al., 2018, Appendix A), mostrando que a solução periódica satisfaz (56) para  $\omega$  suficientemente grande, e então conclusão da atratividade do Equilíbrio de Nash  $\Theta^*$  de acordo com (57). O conjunto residual final para o erro  $\theta(t) - \theta^*$  em (58) depende de  $|a| e^{\max(D_i)\sqrt{\omega/2}}$  devido a amplitude dos sinais de dither  $S_i(t)$  em (21).  $\square$

## 5. SIMULAÇÕES COM UM JOGO DO TIPO DUOPÓLIO

Para exemplificar um jogo não-cooperativo com dois jogadores que empregam a estratégia de busca extremal com compensação de atraso proposta, considera-se as seguintes funções de *payoff* (7) e (8) sujeitas a EDPs de calor com coeficientes de difusão distintos  $D_1 = 1$  e  $D_2 = 3$  nas decisões dos jogadores,  $i \in \{1, 2\}$ :

$$J_1(\Theta(t)) = -5 \Theta_1^2(t) + 5 \epsilon \Theta_1(t) \Theta_2(t) + 250 \Theta_1(t) - 150 \Theta_2(t) - 3000, \quad (66)$$

$$J_2(\Theta(t)) = -5 \Theta_2^2(t) + 5 \epsilon \Theta_1(t) \Theta_2(t) - 150 \Theta_1(t) + 150 \Theta_2(t) + 2500, \quad (67)$$

que, de acordo com (13), gera um único Equilíbrio de Nash em

$$\Theta_1^* = \theta_1^* = \frac{100 + 30\epsilon}{4 - \epsilon^2}, \quad \Theta_2^* = \theta_2^* = \frac{60 + 50\epsilon}{4 - \epsilon^2}. \quad (68)$$

Exceto para os EDPs de calor nos sinais de entrada dos jogadores, a planta e os parâmetros do controlador foram escolhidos de forma semelhante a Frihauf et al. (2012):  $\epsilon = 1$ ,  $a_1 = 0.075$ ,  $a_2 = 0.05$ ,  $k_1 = 2$ ,  $k_2 = 5$ ,  $\omega_1 = 26.75$  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 foram definidas como  $c_1 = c_2 = 100$ . As Figuras 3(a) e 3(b) mostram uma evolução notável na busca do Equilíbrio de Nash e a compensação simultânea dos EDPs de calor em nosso jogo não cooperativo. Além disso, a Figura 4 apresenta a evolução do estado de dimensão infinita  $\alpha_1(x, t)$  e  $\alpha_2(x, t)$  das EDPs do Calor.

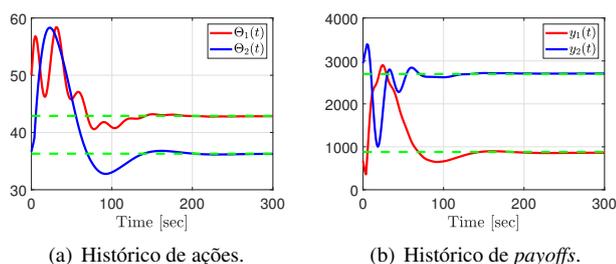


Figura 3. (a) histórico de ações e (b) histórico de payoffs para  $J_1$  e  $J_2$ . As linhas tracejadas representam os valores do Equilíbrio de Nash,  $\Theta_1^* = 43.33$  e  $\Theta_2^* = 36.67$  (with  $J_1(\Theta^*) = 889$  e  $J_2(\Theta^*) = 2722$ ).

## 6. CONCLUSÃO

Este artigo introduziu uma abordagem via busca extremal e controle de fronteira para calcular de forma distribuída o equilíbrio de Nash em duopólios não-cooperativos com funções de *payoff* quadráticas desconhecidas em um cenário em que as ações dos jogadores são propagadas através de EDPs de calor distintas. Dessa forma, qualquer 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 local e a convergência são garantidas por meio da teoria da média em dimensão infinita e análise de pequenos ganhos para cascatas de EDOs-EDPs. Simulações conduzidas para um jogo de dois jogadores sob EDPs de calor com coeficientes de difusão distintos ilustram os resultados teóricos apresentados.

## REFERÊNCIAS

Amina, S., Schwartz, G.A., and Sastry, S.S. (2013). Security of interdependent and identical networked control systems. *Automatica*, 49, 186–192.

Artstein, Z. (1982). Linear systems with delayed controls: A reduction. *IEEE Transactions on Automatic Control*, 27, 869–879.

Başar, T. and Olsder, G.J. (1999). *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, Philadelphia.

Başar, T. and Zaccour, G. (2018). *Handbook of Dynamic Game Theory*, volume 2. Springer International Publishing.

Feiling, J., Koga, S., Krstić, M., and Oliveira, T.R. (2018). Gradient extremum seeking for static maps with actuation dynamics governed by diffusion PDEs. *Automatica*, 95, 197–206.

Frihauf, P., Krstić, M., and Başar, T. (2012). Nash equilibrium seeking in noncooperative games. *IEEE Transactions on Automatic Control*, 57, 1192–1207.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. The MIT Press, Cambridge–Massachusetts.

Ghaffari, A., Krstić, M., and Nesić, D. (2012). Multivariable Newton-based extremum seeking. *Automatica*, 48, 1759–1767.

Hale, J.K. and Lunel, S.M.V. (1990). Averaging in infinite dimensions. *Journal of Integral Equations and Applications*, 2, 463–494.

Han, Z., Niyato, D., Saad, W., and Başar, T. (2019). *Game Theory for Next Generation Wireless and Communication Networks: Modeling, Analysis, and Design*. Cambridge University Press.

Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge Univ. Press.

Karafyllis, I. and Krstić, M. (2018). *Input-to-State Stability for PDEs*. Springer, Switzerland.

Khalil, H.K. (2002). *Nonlinear Systems*. Prentice Hall.

Krstić, M. (2009). *Delay Compensation for Nonlinear, Adaptive, and PDE Systems*. Birkhauser, Boston.

Krstić, M. and Smyshlyayev, A.S. (2008). *Boundary Control of PDEs: A Course on Backstepping Designs*. SIAM, Philadelphia–Pennsylvania.

Krstić, M. and Wang, H.H. (2000). Stability of extremum seeking feedback for general dynamic systems. *Automatica*, 36, 595–601.

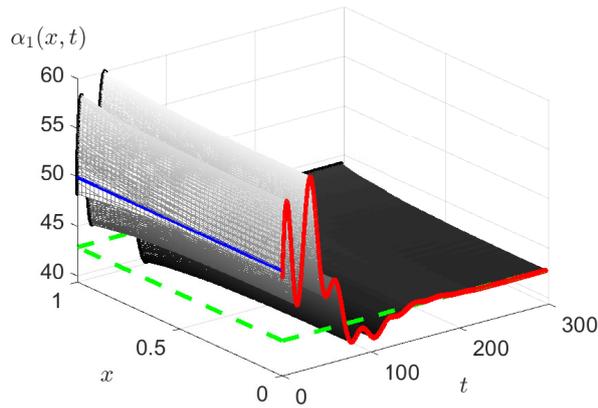
Lewicka, M. and Manfredi, J.J. (2014). Game theoretical methods in PDEs. *Bollettino dell'Unione Matematica Italiana*, 7, 211–216.

Nash, J.F. (1951). Noncooperative games. *Annals of Mathematics*, 54, 286–295.

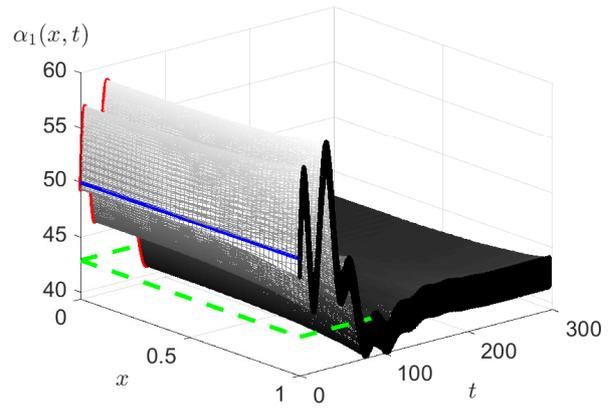
Oliveira, T.R., Krstić, M., and Tsubakino, D. (2017). Extremum seeking for static maps with delays. *IEEE Transactions on Automatic Control*, 62, 1911–1926.

Oliveira, T.R., Rodrigues, V.H.P., Krstić, M., and Başar, T. (2020a). Nash equilibrium seeking with arbitrarily delayed player actions. In *Conference on Decision and Control (CDC)*, 150–155. Jeju Island, KR.

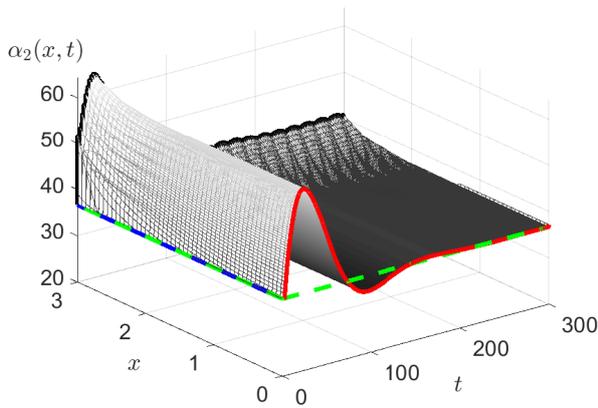
Oliveira, T.R., Tsubakino, D., and Krstić, M. (2020b). A simplified multivariable gradient extremum seeking for distinct input delays with delay-independent convergence rates. In *American Control Conference (ACC)*, 608–613. Denver - CO, USA.



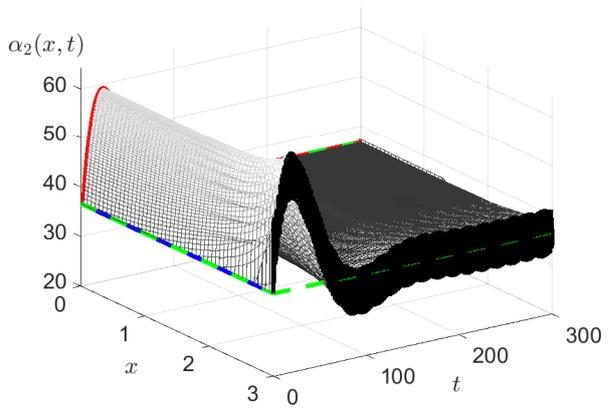
(a) Parâmetro  $\Theta_1(t)$  (vermelho) convergindo para um resíduo  $\mathcal{O}(|a| + 1/\omega)$  da vizinhança de  $\Theta_1^*$  (tracejado-verde) de acordo com (58).



(b) Parâmetro  $\theta_1(t)$  (preto) convergindo para um resíduo  $\mathcal{O}(|a|e^{\max(D_i)\sqrt{\omega/2}} + 1/\omega)$  da vizinhança de  $\theta_1^*$  (tracejado-verde) de acordo com (57).



(c) Parâmetro  $\Theta_2(t)$  (vermelho) convergindo para um resíduo  $\mathcal{O}(|a| + 1/\omega)$  da vizinhança de  $\Theta_2^*$  (tracejado-verde) de acordo com (58).



(d) Parâmetro  $\theta_2(t)$  (preto) convergindo para um resíduo  $\mathcal{O}(|a|e^{\max(D_i)\sqrt{\omega/2}} + 1/\omega)$  da vizinhança de  $\theta_2^*$  (tracejado-verde) de acordo com (57).

Figura 4. Evolução do estado de dimensão infinita  $\alpha_1(x, t)$  e  $\alpha_2(x, t)$  das EDPs do Calor em jogo do tipo duopólio com atuação de fronteira de tipo Dirichlet: de  $\alpha_1(D_1, t) = \theta_1(t)$  a  $\alpha_1(0, t) = \Theta_1(t)$ , com  $D_1 = 1$  para o jogador  $J1$  e de  $\alpha_2(D_2, t) = \theta_2(t)$  a  $\alpha_2(0, t) = \Theta_2(t)$ , com  $D_2 = 3$  para o jogador  $J2$ .