

# Stochastic Real-Time Deception in Nash Equilibrium Seeking for Games with Quadratic Payoffs

**Michael Tang**

MYT001@UCSD.EDU

*Department of Electrical and Computer Engineering, University of California, San Diego*

**Miroslav Krstic**

MKRSTIC@UCSD.EDU

*Department of Mechanical and Aerospace Engineering, University of California, San Diego*

**Jorge Poveda**

POVEDA@UCSD.EDU

*Department of Electrical and Computer Engineering, University of California, San Diego\**

**Editors:** N. Ozay, L. Balzano, D. Panagou, A. Abate

## Abstract

In multi-agent autonomous systems, deception is a fundamental concept which characterizes the exploitation of unbalanced information to mislead victims into choosing oblivious actions. This effectively alters the system’s long term behavior, leading to outcomes that may be beneficial to the deceiver but detrimental to victim. We study this phenomenon for a class of model-free Nash equilibrium seeking (NES) where players implement independent stochastic exploration signals to learn the pseudogradient flow. In particular, we show that deceptive players who obtain real-time measurements of other players’ stochastic perturbation can incorporate this information into their own NES action update, consequentially steering the overall dynamics to a new operating point that could potentially improve the payoffs of the deceptive players. We consider games with quadratic payoff functions, as this restriction allows us to derive a more explicit formulation of the capabilities of the deceptive players. By leveraging results on multi-input stochastic averaging for dynamical systems, we establish local exponential (in probability) convergence for the proposed deceptive NES dynamics. To illustrate our results, we apply them to a two player quadratic game.

**Keywords:** Deception, Nash equilibrium seeking, Multi-agent systems

## 1. Introduction

### 1.1. Motivation

The study of data-driven multi-agent systems is becoming increasingly relevant, especially in engineering domains such as smart grids, power systems, robotics and machine learning. Game theory offers a powerful framework for analyzing the interactions between agents in cooperative and competitive settings. However, traditional game-theoretic methods mostly assume that all agents share the same amount of information for the purpose of decision-making. When agents gain access to *privileged information*, traditional game-theoretic concepts, such as the Nash equilibrium Nash (1951), may lose applicability. Such asymmetry of information can be exploited to shape the system’s long-term behavior in a way that favors certain agents while disadvantaging others. In adversarial environments, some agents may leverage privileged information to mislead their opponents into adopting suboptimal policies. This phenomenon, commonly referred to as *deception in games*, has received significant research interest partly due to the growing need for safe, resilient

---

\* This work was supported in part by NSF grants ECCS CAREER 2305756, CMMI 2228791, and AFOSR YIP: FA9550-22-1-0211

decision-making systems in game-theoretic and data-driven settings. Deception has been studied in many domains, such as cybersecurity [Pawlick et al. \(2021\)](#), robotics [Huang and Zhu \(2022\)](#) and biological systems [Smith \(1987\)](#). While deception can serve as a means by which one achieves their goal in a non-competitive environment, it has also been thoroughly studied in competitive settings, such as signaling games [Kouzehgar and Badamchizadeh \(2019\)](#), repeated security games [Nguyen et al. \(2019\)](#), etc. Despite the wide variety of applications, a common theme is the asymmetry of information. For an agent to have the capacity to deceive, they must be given access to information that the other agents are unaware of. Considerable research has also been done to identify when to optimally deceive [Wagner and Arkin \(2011\)](#), how to detect deceptive agents [Zhu et al. \(2021\)](#), and strategies to counter deception [Santos and Li \(2010\)](#).

In this work, we are interested in studying deception within the framework of stochastic model-free Nash equilibrium seeking (NES) algorithms for non-cooperative games. In a non-cooperative game, each player seeks to minimize their cost function, which may depend on the actions of other players. The solution to this optimization problem, known as the *Nash equilibrium (NE)*, is not straightforward to compute. Given this challenge, various NES algorithms have been designed, including distributed [Ye and Hu \(2017\)](#), semi-decentralized [Belgioioso and Grammatico \(2017\)](#) and hybrid [Xu et al. \(2023\)](#) algorithms. However, in many cases, players do not know the mathematical model of their cost function or the structure of the game. In this case, players must implement adaptive seeking dynamics that incorporate exploration and exploitation strategies using real-time measurements of their cost function values. Such algorithms have been studied in, e.g., [Frihauf et al. \(2012\)](#); [Poveda and Quijano \(2015\)](#); [Stanković et al. \(2012\)](#), using extremum-seeking ideas. For this class of algorithms, it was recently shown in [Tang et al. \(2024\)](#) that a player who obtains knowledge of another player’s exploration policy can incorporate this information into their own action update, effectively driving the NES dynamics to a point that might be more beneficial for the deceptive player than the nominal NE. Moreover, conditions that guarantee stability of the closed-loop deceptive NES dynamics were derived. While the results raised important concerns regarding the functionality of real-time adaptive learning algorithms in the presence of adversarial behavior, one important limitation is that they rely on the players using deterministic periodic excitation signals to learn the “deceptive” Nash equilibrium. However, many practical systems, such as biological systems and source-seeking systems, rely on unpredictable and random probing to learn optimal operating points. To the best of our knowledge, the study of stochastic deception in Nash equilibrium-seeking problems has remained mostly unexplored.

## 1.2. Contributions

In this paper, we introduce the concept of *stochastic deception* in Nash equilibrium seeking problems. In particular, by leveraging averaging results for stochastic differential equations [Liu and Krstic \(2011\)](#), we extend the deterministic framework of [Tang et al. \(2024\)](#) to the case where players use random probing for the purpose of exploration. While the type of deception considered in [Tang et al. \(2024\)](#) only required knowledge of the victim’s sinusoidal frequency (which could be interpreted as a “key” in the learning mechanism), the deception mechanism in this work requires real-time measurements of the victim’s stochastic perturbation. In other words, to achieve stochastic deception, the deceiver must be able to establish a sensing link with the victim’s probing policy. By focusing on quadratic games with a strictly diagonally dominant pseudogradient, we establish suitable stability properties for the average dynamics, resulting in more explicit characterizations on

the influence of deception compared to earlier works. We further generalize the class of deceptive NES algorithms from [Tang et al. \(2024\)](#) by allowing players to select differing gains and amplitudes for their exploration policies. We then establish conditions which can guarantee stochastic exponential stability of the closed-loop stochastic deceptive NES dynamics. While the deterministic setting required players to choose different exploration frequencies, our work merely requires that each player's perturbation process is chosen independently, resulting in significantly greater flexibility for the players' strategies. All results are further illustrated via numerical examples.

## 2. Preliminaries

Given a matrix  $A \in \mathbb{R}^{m \times n}$ , we use  $[A]_{ij}$  and  $[A]_k$  to denote the  $(i, j)$  entry and  $k$ -th row of  $A$ , respectively. Similarly, given a vector  $b \in \mathbb{R}^n$ , we use  $[b]_i$  to denote the  $i$ -th entry of  $b$ . If  $N \in \mathbb{N}$ , we use  $[N]$  to denote the set of positive integers no greater than  $N$ , i.e.  $[N] = \{1, 2, \dots, N\}$ . Given a vector  $v \in \mathbb{R}^n$ , we use  $\text{diag}(v) \in \mathbb{R}^{n \times n}$  to denote the diagonal matrix with the  $i$ -th diagonal entry given by  $[v]_i$ . For  $r > 0$ , we use  $\mathcal{B}_r(v)$  to denote the open ball of radius  $r$  around  $v$ , i.e.  $\mathcal{B}_r(v) = \{p \in \mathbb{R}^n : |p - v| < r\}$ , where  $|\cdot|$  is the Euclidean norm.

## 3. Problem Formulation

Consider an  $N$ -player noncooperative game, where player  $i \in [N]$  implements action  $x_i \in \mathbb{R}$  and aims to unilaterally minimize their cost function  $J_i : \mathbb{R}^N \rightarrow \mathbb{R}$ . We let  $x = [x_1, \dots, x_N]^\top$  denote the vector of all the players' actions, and we use  $x_{-i}$  to denote the vector of all players' actions except for that of player  $i$ . A policy  $x^* \in \mathbb{R}^N$  is said to be a *Nash equilibrium (NE)* of the game if it satisfies

$$J_i(x_i^*, x_{-i}^*) \leq J_i(x_i, x_{-i}^*), \quad \forall x_i \in \mathbb{R}, \quad \forall i \in [N]. \quad (1)$$

In this paper, we are primarily interested in games with quadratic payoffs, i.e., we assume that for each player  $i$ ,  $J_i$  takes the quadratic form

$$J_i(x) = \frac{1}{2} x^\top A_i x + b_i^\top x + c_i \quad (2)$$

where  $A_i \in \mathbb{R}^{N \times N}$  is symmetric,  $b_i \in \mathbb{R}^N$ , and  $c_i \in \mathbb{R}$ . With this in mind, the *pseudogradient* of the game can be defined as  $\mathcal{G}(x) = \mathcal{A}x + \mathcal{B}$ , where  $\mathcal{A} \in \mathbb{R}^{N \times N}$  satisfies  $[\mathcal{A}]_{i:} = [A_i]_{i:}$  and  $\mathcal{B} \in \mathbb{R}^N$  satisfies  $[\mathcal{B}]_i = [b_i]_i$  for  $i \in [N]$ . We make the following assumption on the pseudogradient

**Assumption 1** *The matrix  $\mathcal{A}$  is strictly diagonally dominant and satisfies  $[\mathcal{A}]_{ii} > 0$  for all  $i \in [N]$ .*

For player  $i$  to learn their NE strategy using only measurements of their payoff  $J_i$ , they implement the following stochastic NE seeking algorithm, originally introduced in [Liu and Krstić \(2011\)](#):

$$x_i = u_i + a_i f_i(\eta_i(t)) \quad (3a)$$

$$\dot{u}_i = -\frac{\gamma_i k_i}{a_i} f_i(\eta_i(t)) J_i(x) \quad (3b)$$

where  $a_i, k_i > 0$  are gains and  $f_i$  is a bounded odd function to be chosen by player  $i$  with

$$\gamma_i = \int_{-\infty}^{\infty} f_i^2(s) \frac{1}{\sqrt{\pi} q_i} e^{-\frac{s^2}{q_i^2}} ds \quad (4)$$

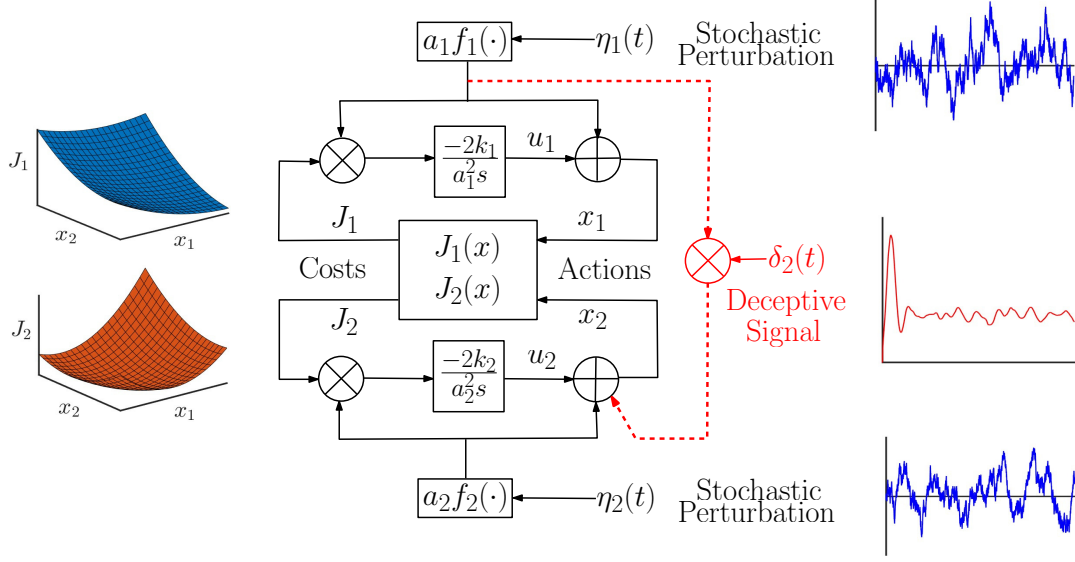


Figure 1: Block diagram depicting the deception mechanism for stochastic NES in a two player game where player 2 deceives player 1.

where  $q_i > 0$ . The signals  $\eta_i(t), i \in [N]$  are selected to be independent time homogeneous continuous Markov ergodic processes chosen by player  $i$ , such as the Ornstein-Uhlenbeck (OU) process:

$$\vartheta_i d\eta_i(t) = -\eta_i(t)dt + \sqrt{\vartheta_i q_i} dW_i(t) \quad (5)$$

where  $\vartheta_i > 0$  and  $W_i(t), i \in [N]$  are independent 1-dimensional standard Brownian motion on a complete probability space  $(\Omega, \mathcal{F}, P)$  with sample space  $\Omega$ ,  $\sigma$ -field  $\mathcal{F}$ , and probability measure  $P$ . We express  $\vartheta_i$  in the form  $\vartheta_i = \bar{\vartheta} \bar{\vartheta}_i$  for some  $\bar{\vartheta} > 0$ , as this form is convenient for our stability analysis. Without loss of generality, we assume  $\bar{\vartheta}_1 = 1$ . Now we introduce the concept of *stochastic deception* within the context of (3):

**Definition 1** A player  $i$  is said to be stochastically deceptive towards a set of players  $\mathcal{D}_i \subset [N] \setminus \{i\}$  if its actions are updated according to the following dynamics:

$$x_i(t) = u_i(t) + a_i f_i(\eta_i(t)) + \delta_i \sum_{k \in \mathcal{D}_i} a_k f_k(\eta_k(t)) \quad (6a)$$

$$\dot{u}_i = -\frac{\gamma_i k_i}{a_i} f_i(\eta_i(t)) J_i(x) \quad (6b)$$

where  $\delta_i(t)$  is a tunable deterministic deceptive gain that satisfies  $\sup_{t \geq 0} |\delta_i(t)| > 0$ .

For a player to be stochastically deceptive to player  $d$ , they must obtain knowledge of player  $d$ 's amplitude  $a_d$ , their function  $f_d$ , and real-time measurements of player  $d$ 's stochastic exploration signal  $\eta_d(t)$ . While this may seem more restrictive than what is considered in Tang et al. (2024), which only required knowledge of exploration frequencies  $\omega_d$ , we note that our setting is in fact more lenient. Indeed, the NES dynamics in Tang et al. (2024) assumed  $a_k = a$  for all  $k \in [N]$  and that all players use a deterministic sinusoidal exploration policy, i.e  $f_k(t) = \sin(t)$  and  $\eta_k(t) = \omega_k t$

for all  $k \in [N]$ . Hence, a player who is deceptive to player  $d$  in Tang et al. (2024) can compute  $a_d f_d(\eta_d(t))$  in real-time, which is what we require in (6a). Figure 1 depicts this phenomenon for a two player game where player 2 is deceptive to player 1.

In this work, we assume there are  $n$  deceptive players, and the set of deceptive players is given by  $\mathcal{D} = \{d_1, \dots, d_n\}$ . Moreover, we let  $\mathcal{V} = \{j \in [N] : \exists i \in [N] \text{ s.t } j \in \mathcal{D}_i\}$  denote the set of “victims”, i.e the players who are being deceived. We say player  $i$  is deceiving  $n_i \in [N - 1]$  players in  $\mathcal{D}_i$  if they implement (6) and update  $\delta_i$  according to the following dynamics:

$$\dot{\delta}_i = \varepsilon \varepsilon_i (J_i(x) - J_i^{\text{ref}}). \quad (7)$$

where  $J_i^{\text{ref}}$  is the desired payoff of player  $i$  and  $\varepsilon, \varepsilon_i > 0$  are small parameters. Now, we can state the full model-free stochastic dynamics for the deceptive game:

$$x_i = \begin{cases} u_i + a_i f_i(\eta_i(t)) + \delta_i \sum_{j \in \mathcal{D}_i} a_j f_j(\eta_j(t)) & \text{if } i \in \mathcal{D} \\ u_i + a_i f_i(\eta_i(t)) & \text{else} \end{cases} \quad (8a)$$

$$\dot{u}_i = -\frac{\gamma_i k_i}{a_i} J_i(x) f_i(\eta_i(t)), \quad (8b)$$

$$\dot{\delta}_i = \varepsilon \varepsilon_i (J_i(x) - J_i^{\text{ref}}), \quad i \in \mathcal{D} \quad (8c)$$

Let  $\delta = [\delta_{d_1}, \dots, \delta_{d_n}]^\top$  and let  $\mathcal{K}_j$  represent the set of players who are deceptive to player  $j$ , i.e.,  $\mathcal{K}_j = \{i \in \mathcal{D} : j \in \mathcal{D}_i\}$ . We use  $\zeta \in \mathbb{R}^{N+n}$  to denote the full state vector, i.e  $\zeta = [u^\top \ \delta^\top]^\top$ . We define the perturbed pseudogradient parameters as  $\bar{\mathcal{A}}(\delta) \in \mathbb{R}^{N \times N}$  and  $\bar{\mathcal{B}}(\delta) \in \mathbb{R}^N$ , which have entries satisfying:

$$[\bar{\mathcal{A}}(\delta)]_{ii} = [\mathcal{A}]_{ii} + \sum_{k \in \mathcal{K}_i} \delta_k [A_i]_{ki}, \quad [\bar{\mathcal{B}}(\delta)]_i = [\mathcal{B}]_i + \sum_{k \in \mathcal{K}_i} \delta_k [b_i]_k \quad (9)$$

The intuition behind (9) will become clear when the averaged dynamics of (8) are computed in Section 4. We can then define the following *stability preserving set*

$$\Delta := \{\delta \in \mathbb{R}^n : -\text{diag}([k_1, \dots, k_N]^\top) \bar{\mathcal{A}}(\delta) \text{ is Hurwitz}\}. \quad (10)$$

By leveraging Assumption 1, we obtain an estimate of  $\Delta$  that is somewhat conservative but an improvement of the results from Tang et al. (2024).

**Lemma 2** *Pick  $r \in (0, \infty)$  satisfying*

$$r = \min_{i \in \mathcal{V}} \min \left( \frac{[\mathcal{A}]_{ii} - \sum_{j \neq i}^N |[\mathcal{A}]_{ij}|}{\sum_{k \in \mathcal{K}_i} \sum_{j=1}^N |[A_i]_{kj}|}, \frac{[\mathcal{A}]_{ii}}{\sum_{k \in \mathcal{K}_i} |[A_i]_{ki}|} \right) \quad (11)$$

where we interpret  $\frac{p}{q} := \infty$  if  $p > 0$  and  $q = 0$ . Then  $\mathcal{B}_r(\mathbf{0}_n) \subset \Delta$ , where  $\mathbf{0}_n \in \mathbb{R}^n$  is the  $n$ -dimensional zero vector.

**Proof** Fix  $i \in \mathcal{V}$ , and pick  $\delta$  satisfying  $|\delta| < \frac{[\mathcal{A}]_{ii}}{\sum_{k \in \mathcal{K}_i} |[A_i]_{ki}|}$ . Then,

$$[\bar{\mathcal{A}}(\delta)]_{ii} = [\mathcal{A}]_{ii} + \sum_{k \in \mathcal{K}_i} \delta_k [A_i]_{ki} \geq [\mathcal{A}]_{ii} - |\delta| \sum_{k \in \mathcal{K}_i} |[A_i]_{ki}| > 0 \quad (12)$$

Now let  $\delta$  also satisfy  $|\delta| < \frac{[\mathcal{A}]_{ii} - \sum_{j \neq i}^N |[\mathcal{A}]_{ij}|}{\sum_{k \in \mathcal{K}_i} \sum_{j=1}^N |[\mathcal{A}]_{kj}|}$ . Then, for  $\bar{\mathcal{A}}(\delta)$  to be strictly diagonally dominant, we require

$$[\mathcal{A}]_{ii} + \sum_{k \in \mathcal{K}_i} \delta_k [A_i]_{ki} > \sum_{j \neq i}^N \left| [\mathcal{A}]_{ij} + \sum_{k \in \mathcal{K}_i} \delta_k [A_i]_{kj} \right| \quad (13)$$

For (13) to hold, it suffices to have

$$[\mathcal{A}]_{ii} - |\delta| \sum_{k \in \mathcal{K}_i} |[A_i]_{ki}| > \sum_{j \neq i}^N \left( |[\mathcal{A}]_{ij}| + |\delta| \sum_{k \in \mathcal{K}_i} |[A_i]_{kj}| \right) \quad (14)$$

which can be guaranteed whenever

$$[\mathcal{A}]_{ii} - \sum_{j \neq i}^N |[\mathcal{A}]_{ij}| > |\delta| \sum_{k \in \mathcal{K}_i} \sum_{j=1}^N |[A_i]_{kj}| \quad (15)$$

Taking the minimum over all  $i \in \mathcal{V}$ , it follows that  $\bar{\mathcal{A}}(\delta)$  is strictly diagonally dominant for  $|\delta| < r$ , and  $[\bar{\mathcal{A}}(\delta)]_{ii} > 0$  for all  $i \in [N]$ . By the Gershgorin Circle Theorem, we conclude that  $\delta \in \Delta$ . ■

This estimate is a significant improvement compared to the results of Tang et al. (2024), which only guaranteed that  $\Delta$  contains a neighborhood of the origin. By further exploiting the diagonal dominance assumption, we are able to derive an explicit lower bound on the “size” of this neighborhood. To characterize when the deceptive players are able to achieve their desired payoffs via deception while preserving stability, we introduce the notion of *attainability* as was presented in Tang et al. (2024):

**Definition 3** A vector  $J^{ref} = [J_{d_1}^{ref}, \dots, J_{d_n}^{ref}]^\top$  is said to be attainable if there exists  $\delta^* \in \Delta$  such that:

1.  $J_{d_k}(-\bar{\mathcal{A}}(\delta^*)^{-1}\bar{\mathcal{B}}(\delta^*)) = J_{d_k}^{ref}, \quad \forall k \in [n]$ .
2. The matrix  $\Lambda(\delta^*) \in \mathbb{R}^{n \times n}$  with  $[\Lambda(\delta^*)]_{jk} = \nabla_j \xi_k(\delta^*)$  is Hurwitz, where  $\xi_k : \mathbb{R}^n \rightarrow \mathbb{R}$  is given by  $\xi_k(\delta) := \varepsilon_{d_k} J_{d_k}(-\bar{\mathcal{A}}(\delta)^{-1}\bar{\mathcal{B}}(\delta))$ .

We let  $\mathcal{J} \subset \mathbb{R}^n$  denote the set of all attainable vectors  $J^{ref} = [J_{d_1}^{ref}, \dots, J_{d_n}^{ref}]^\top$ .

We can now state the main result of this paper, which establishes stochastic stability of the NES dynamics (8).

**Theorem 4** Consider the stochastic NES dynamics (8) where the costs  $J_i$  are of the quadratic form (2),  $J^{ref} \in \mathcal{J}$ , the pseudogradient matrix  $\mathcal{A}$  satisfies Assumption 1, the noise  $\eta_i(t)$  are generated according to (5) and  $f_i(\cdot)$  are bounded odd functions for  $i \in [N]$ . Then for each  $\tilde{\varepsilon} > 0$  there exists  $a^*$  such that for  $a_1, \dots, a_N$  satisfying  $\max_i a_i < a^*$ , there exists  $\varepsilon^*$  such that for all  $\varepsilon \in (0, \varepsilon^*)$  there exists  $R > 0, C > 0, M > 0$  and  $\zeta^* = \begin{bmatrix} u^* \\ \delta^* \end{bmatrix} \in \mathbb{R}^{N+n}$  satisfying the following:

1.  $|J_{d_k}(u^*) - J_{d_k}^{ref}| < \tilde{\varepsilon}$  for all  $k \in [n]$

2. For any  $\tilde{r} > 0$  and any initial condition  $\zeta_0 \in \mathbb{R}^{N+n}$  with  $|\zeta_0 - \zeta^*| < R$ , the solution of (8) satisfies

$$\lim_{\vartheta \rightarrow 0} \inf \{t \geq 0 : |\zeta(t) - \zeta^*| > C|\zeta_0 - \zeta^*|e^{-Mt} + \tilde{r}\} = \infty \quad a.s. \quad (16)$$

3. There exists  $\varepsilon_0 > 0$  and a function  $T : (0, \varepsilon_0) \rightarrow \mathbb{N}$  such that

$$\lim_{\vartheta \rightarrow 0} P(|\zeta(t) - \zeta^*| \leq C|\zeta_0 - \zeta^*|e^{-Mt} + \tilde{r} \quad \forall t \in [0, T(\vartheta)]) = 1 \quad \text{with} \quad \lim_{\vartheta \rightarrow 0} T(\vartheta) = \infty. \quad (17)$$

This result establishes weak stochastic stability of  $\zeta^*$  for (8) under random perturbations generated by (5). Moreover,  $\zeta^*$  can be made such that the costs of the deceptive players evaluated at the “ $u$ ” component (which itself can be made arbitrarily close to the players’ actions  $x$ ) are arbitrarily close to their desired reference values. The proof is presented in the next section.

#### 4. Analysis of the Deceptive NES Dynamics and Proof of Convergence

To prove Theorem 4, we compute the “average” dynamics of (8) and apply Theorem A.3 from Liu and Krstić (2011). First, we let  $\eta(t) \in \mathbb{R}^N$  denote the vector of the stochastic components, i.e  $\eta(t) = x - u$  where  $x$  and  $u$  are given by (8a) and (8b). We can also define  $\chi_i(t) = \eta_i(\vartheta_i t)$  and  $B_i(t) = \frac{1}{\sqrt{\vartheta_i}} W_i(\vartheta_i t)$ . Then, by (5) we have

$$d\chi_i(t) = -\chi_i(t)dt + q_i dB_i(t), \quad (18)$$

where  $[B_1(t), \dots, B_N(t)]^\top$  is an  $N$ -dimensional Brownian motion on the probability space  $(\Omega, \mathcal{F}, P)$ . Following the notation in Liu and Krstić (2011), we introduce the change of variable

$$Z_1(t) = \chi_1(t), \quad Z_2(t) = \chi_2(t/\bar{\vartheta}_2), \dots, Z_N(t) = \chi_N(t/\bar{\vartheta}_N), \quad (19)$$

which will allow us to rewrite the  $\dot{u}_i$  system (8b) in the following way:

$$\dot{u}_i = -\frac{\gamma_i k_i}{a_i} f_i(Z_i(t/\vartheta)) J_i(u + \eta(t)) \quad (20)$$

where

$$[\eta(t)]_k = \begin{cases} a_k f_k(Z_k(t/\vartheta)) + \delta_k \sum_{j \in \mathcal{D}_k} a_j f_j(Z_j(t/\vartheta)) & \text{if } k \in \mathcal{D} \\ a_k f_k(Z_k(t/\vartheta)) & \text{else} \end{cases} \quad (21)$$

Since  $(\chi_i(t), t \geq 0)$  is ergodic with invariant distribution  $\mu_i(dz_i) = \frac{1}{q_i \sqrt{\pi}} e^{-z_i^2/q_i^2} dz_i$ , it follows from Lemma A.2 in Liu and Krstić (2011) that  $[Z_1(t), \dots, Z_N(t)]^\top$  is ergodic with invariant distribution  $\mu_1 \times \dots \times \mu_N$ . Now we define state  $z_i = Z_i(t/\vartheta)$  and we let  $\bar{\eta}(z) \in \mathbb{R}^N$  be given by (21) but with  $z_i$  in place of  $Z_i(t/\vartheta)$ , i.e

$$[\bar{\eta}(z)]_k = \begin{cases} a_k f_k(z_k) + \delta_k \sum_{j \in \mathcal{D}_k} a_j f_j(z_j) & \text{if } k \in \mathcal{D} \\ a_k f_k(z_k) & \text{else} \end{cases} \quad (22)$$



Then the system (20) becomes

$$\dot{u}_i = -\frac{\gamma_i k_i}{a_i} f_i(z_i) \left( J_i(u) + \bar{\eta}(z)^\top \nabla J_i(u) + \frac{1}{2} \bar{\eta}(z)^\top A_i \bar{\eta}(z) \right) \quad (23)$$

We can then use A.4 in [Liu and Krstić \(2011\)](#) to compute the average system of (23). Given a state  $u$ , we use  $\tilde{u}$  to denote the corresponding state in the average system. Then, by repeatedly applying Fubini's theorem, we obtain

$$\dot{\tilde{u}}_i = -\frac{\gamma_i k_i}{a_i} \int_{\mathbb{R}^N} f_i(z_i) \left( J_i(\tilde{u}) + \bar{\eta}(z)^\top \nabla J_i(\tilde{u}) + \frac{1}{2} \bar{\eta}(z)^\top A_i \bar{\eta}(z) \right) \mu_1(dz_1) \times \dots \times \mu_N(dz_N) \quad (24)$$

$$= -\frac{\gamma_i k_i}{a_i} \int_{\mathbb{R}^N} f_i(z_i) \bar{\eta}(z)^\top \nabla J_i(\tilde{u}) \left( \prod_{l=1}^N \frac{1}{q_l \sqrt{\pi}} e^{-\frac{z_l^2}{q_l^2}} \right) dz_1 \dots dz_N \quad (25)$$

$$= -\frac{\gamma_i k_i}{a_i} \int_{\mathbb{R}^N} \left( a_i f_i^2(z_i) \nabla_i J_i(\tilde{u}) + \sum_{j \in \mathcal{K}_i} \tilde{\delta}_j a_i f_i^2(z_i) \nabla_j J_i(\tilde{u}) \right) \left( \prod_{l=1}^N \frac{1}{q_l \sqrt{\pi}} e^{-\frac{z_l^2}{q_l^2}} \right) dz_1 \dots dz_N \quad (26)$$

$$= -k_i \left( \nabla_i J_i(\tilde{u}) + \sum_{j \in \mathcal{K}_i} \tilde{\delta}_j \nabla_j J_i(\tilde{u}) \right) \quad (27)$$

$$= -k_i [\bar{\mathcal{A}}(\tilde{\delta}) \tilde{u} + \bar{\mathcal{B}}(\tilde{\delta})]_i \quad (28)$$

For the  $\delta$  subsystem, we have for  $i \in \mathcal{D}$ :

$$\dot{\tilde{\delta}}_i = \varepsilon \varepsilon_i \int_{\mathbb{R}^N} (J_i(\tilde{u} + \bar{\eta}(z)) - J_i^{\text{ref}}) \mu_1(dz_1) \times \dots \times \mu_N(dz_N) \quad (29)$$

$$= \varepsilon \varepsilon_i \int_{\mathbb{R}^N} \left( J_i(\tilde{u}) + \bar{\eta}(z)^\top \nabla J_i(\tilde{u}) + \frac{1}{2} \bar{\eta}(z)^\top A_i \bar{\eta}(z) - J_i^{\text{ref}} \right) \left( \prod_{l=1}^N \frac{1}{q_l \sqrt{\pi}} e^{-\frac{z_l^2}{q_l^2}} \right) dz_1 \dots dz_N \quad (30)$$

$$= \varepsilon \varepsilon_i \left( J_i(\tilde{u}) - J_i^{\text{ref}} + \int_{\mathbb{R}^N} \frac{1}{2} \bar{\eta}(z)^\top A_i \bar{\eta}(z) \left( \prod_{l=1}^N \frac{1}{q_l \sqrt{\pi}} e^{-\frac{z_l^2}{q_l^2}} \right) dz_1 \dots dz_N \right) \quad (31)$$

$$= \varepsilon \varepsilon_i \left( J_i(\tilde{u}) - J_i^{\text{ref}} + \mathcal{P}_i \left( \begin{bmatrix} a \\ \tilde{\delta} \end{bmatrix} \right) \right) \quad (32)$$

where  $a = [a_1, \dots, a_N]$  and  $\mathcal{P}_i : \mathbb{R}^{N+n} \rightarrow \mathbb{R}$  is a quadratic function satisfying  $\mathcal{P}_i \left( \begin{bmatrix} \mathbf{0}_N \\ \delta \end{bmatrix} \right) = 0$ . Consider the  $\tilde{\zeta}$  system with  $a = \mathbf{0}_N$ . Since  $J^{\text{ref}} \in \mathcal{J}$ , we can apply a standard singular perturbation argument (as done in [Tang et al. \(2024\)](#)) to conclude that there exists some  $\varepsilon^*$  such that for all  $\varepsilon \in (0, \varepsilon^*)$ , this system has an exponentially stable equilibrium point of the form  $\zeta^* = \begin{bmatrix} -\bar{\mathcal{A}}(\delta^*)^{-1} \bar{\mathcal{B}}(\delta^*) \\ \delta^* \end{bmatrix}$  for some  $\delta^* \in \Delta$ . By the implicit function theorem, there exists some neighborhood  $U$  of  $\mathbf{0}_N$  and  $\mathcal{C}^1$



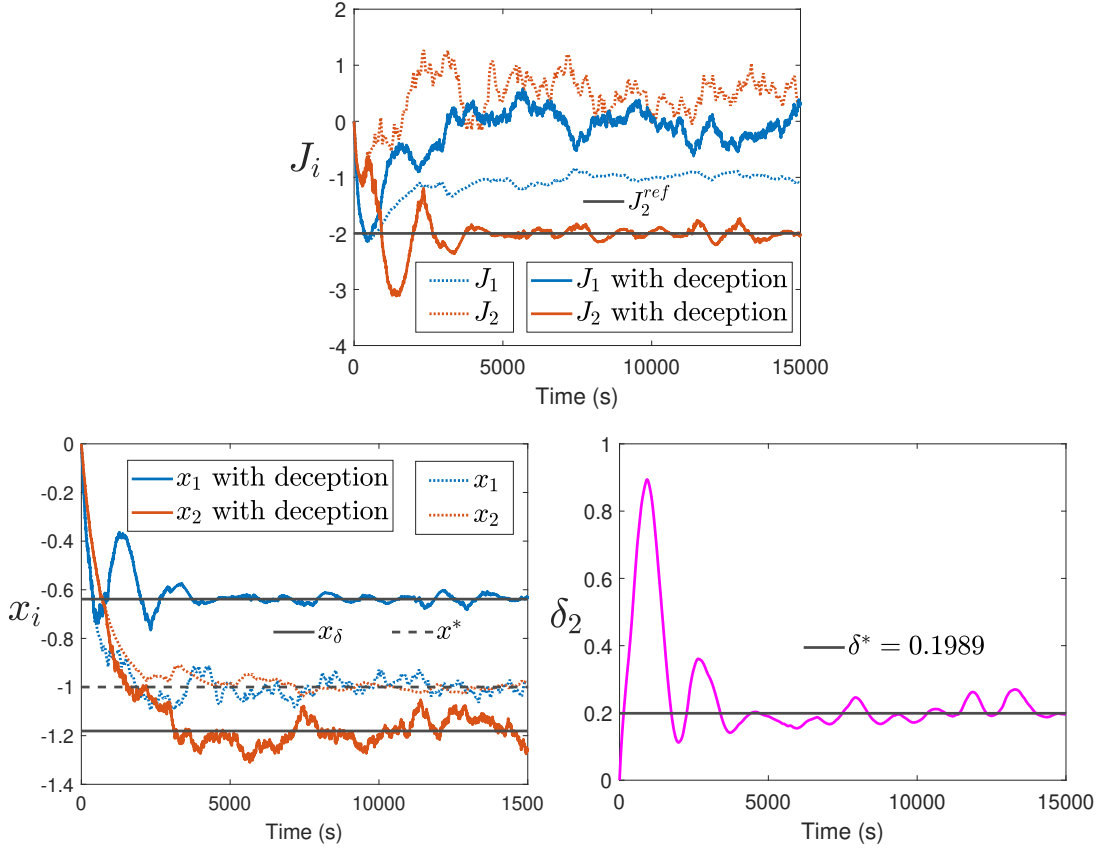


Figure 2: Top: Cost function convergence for the two player game from Section 5 with  $J_2^{ref} = -2$ . Left: The action convergence for the two player game. Right: The corresponding trajectories of  $\delta_1$  when player 2 deceives player 1.

function  $g : U \rightarrow \mathbb{R}^{N+n}$  such that  $g(a)$  is an exponentially stable equilibrium point of the  $\tilde{\zeta}$  system for all  $a \in U$ . Moreover,  $U$  can be chosen such that

$$\left| \mathcal{P}_i \left( \begin{bmatrix} a \\ [g(a)]_{N+1} \\ \vdots \\ [g(a)]_{N+n} \end{bmatrix} \right) \right| < \tilde{\varepsilon} \quad \forall i \in \mathcal{D}, \quad a \in U \quad (33)$$

which establishes item 1) of Theorem 4. We can then define  $a^* = \sup\{r > 0 : \mathcal{B}_r(\mathbf{0}_N) \subset U\}$  and apply Theorem 3 from Liu and Krstić (2011) to complete the proof.

**Remark 5** One key difference between the above proof and the proof of the main stability result in Tang et al. (2024) is that the approach taken in Tang et al. (2024) treats  $\mathcal{P}_i$  as an  $\mathcal{O}(a)$  perturbation and applies standard robustness results for deterministic systems to establish exponential stability in a semi-global practical sense. Since the conditions of Theorem 3 from Liu and Krstić (2011) require exponential stability of an equilibrium point of the averaged system, we leverage the implicit

function theorem to establish the existence of such a point for  $\max_i a_i$  sufficiently small. As the averaged integrator dynamics are still perturbed, we can at best claim that the new equilibrium can be made arbitrarily close to one that achieves each deceptive players' desired payoff.

## 5. Numerical Example

To further illustrate our results, we present a numerical example. Consider a two player quadratic game with costs of the form (2), with parameters

$$A_1 = \begin{bmatrix} 3 & 1 \\ 1 & 5 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \quad c_1 = c_2 = 0. \quad (34)$$

The pseudogradient of this game is given by

$$\mathcal{G}(x) = \mathcal{A}x + \mathcal{B}, \quad \mathcal{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \quad (35)$$

where  $\mathcal{A}$  clearly satisfies Assumption 1. This game has a unique Nash equilibrium at  $x^* = [-1, -1]^\top$  and steady state costs  $J_1(x^*) = -1$ ,  $J_2(x^*) = 0.5$ . For the stochastic NES dynamics (3), the players use  $a_1 = 0.1, k_1 = 0.06, f_1(s) = \text{sat}(s), a_2 = 0.12, k_2 = 0.05, f_2(s) = \sin(s)$  where the saturation function  $\text{sat}(\cdot)$  is given by

$$\text{sat}(s) = \begin{cases} 1 & \text{if } s \geq 1 \\ s & \text{if } s \in (-1, 1) \\ -1 & \text{if } s \leq -1 \end{cases} \quad (36)$$

The stochastic perturbations  $\eta_i(t)$  are generated by (5) with  $\vartheta_1 = 0.0005, q_1 = 0.1, \vartheta_2 = 0.0004, q_2 = 0.07$ . Now we suppose player 2 is deceptive to player 1, so player 2 implements (6) and (7) with  $\varepsilon\varepsilon_2 = 0.001$  and  $J_2^{\text{ref}} = -2$ . With Lemma 2, we obtain that  $\mathcal{B}_{\frac{1}{3}}(0) \subset \Delta$ . It can be verified that  $J_2^{\text{ref}}$  is achievable with  $\delta^* = 0.1989 \in \Delta$ , which results in a “deceptive” Nash equilibrium of  $x_\delta = -\overline{\mathcal{A}}(\delta^*)\overline{\mathcal{B}}(\delta^*) = [-0.6385, -1.1808]^\top$ . We also have  $\nabla\xi_2(\delta^*) = -12\varepsilon_2 < 0$  for  $\varepsilon_2$ , where  $\xi_2$  is given in Definition 3. Hence, the closed loop stochastic deceptive NES dynamics are stable. The trajectories of the players' actions and costs are provided in Figure 2.

## 6. Conclusion

In this paper, we introduced stochastic deception in Nash equilibrium seeking problems for noncooperative games with finitely many players. We leverage results on multi-input stochastic averaging to establish local exponential (in probability) convergence for the deceptive stochastic NES dynamics. We primarily consider quadratic games, as this allows us to consider a more general class of NES dynamics with an improved estimate of the “stability preserving set”  $\Delta$ , but future directions will focus on extending the results to general nonlinear payoffs. Since our deception mechanism requires real-time measurements of the victim's stochastic perturbation, it is also of future interest to study methods a deceptive player could use to obtain such measurements or information in real time.

## References

- Giuseppe Belgioioso and Sergio Grammatico. Semi-decentralized nash equilibrium seeking in aggregative games with separable coupling constraints and non-differentiable cost functions. *IEEE control systems letters*, 1(2):400–405, 2017.
- Paul Frihauf, Miroslav Krstic, and Tamer Basar. Nash equilibrium seeking in noncooperative games. *IEEE Transactions on Automatic Control*, 57(5):1192–1207, 2012. doi: 10.1109/TAC.2011.2173412.
- Linan Huang and Quanyan Zhu. A dynamic game framework for rational and persistent robot deception with an application to deceptive pursuit-evasion. *IEEE Transactions on Automation Science and Engineering*, 19(4):2918–2932, 2022. doi: 10.1109/TASE.2021.3097286.
- Maryam Kouzehgar and Mohammad Ali Badamchizadeh. Fuzzy signaling game of deception between ant-inspired deceptive robots with interactive learning. *Applied Soft Computing*, 75:373–387, 2019.
- S. Liu and M. Krstic. Stochastic Nash equilibrium seeking for games with general nonlinear payoffs. *SIAM J. Control Optim.*, 49(4):1659–1679, 2011.
- Shu-Jun Liu and Miroslav Krstić. Stochastic nash equilibrium seeking for games with general nonlinear payoffs. *SIAM Journal on Control and Optimization*, 49(4):1659–1679, 2011. doi: 10.1137/100811738. URL <https://doi.org/10.1137/100811738>.
- John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951. ISSN 0003486X, 19398980. URL <http://www.jstor.org/stable/1969529>.
- Thanh H Nguyen, Yongzhao Wang, Arunesh Sinha, and Michael P Wellman. Deception in finitely repeated security games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 2133–2140, 2019.
- Jeffrey Pawlick, Quanyan Zhu, et al. *Game theory for cyber deception*. Springer, 2021.
- J. I. Poveda and N. Quijano. Shahshahani gradient like extremum-seeking. *Automatica*, 58:51–59, 2015.
- Eugene Santos and Deqing Li. On deception detection in multiagent systems. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40(2):224–235, 2010. doi: 10.1109/TSMCA.2009.2034862.
- Euclid O Smith. Deception and evolutionary biology. *Cultural Anthropology*, 2(1):50–64, 1987.
- M. Stanković, K. H. Johansson, and D. M. Stipanović. Distributed seeking of Nash equilibria with applications to mobile sensor networks. *IEEE Trans. Autom. Contr.*, 57(4):904–919, 2012.
- Michael Tang, Umar Javed, Xudong Chen, Miroslav Krstic, and Jorge I. Poveda. Deception in nash equilibrium seeking, 2024. URL <https://arxiv.org/abs/2407.05168>.
- Alan R Wagner and Ronald C Arkin. Acting deceptively: Providing robots with the capacity for deception. *International Journal of Social Robotics*, 3:5–26, 2011.

- Wenying Xu, Zidong Wang, Guoqiang Hu, and Jürgen Kurths. Hybrid nash equilibrium seeking under partial-decision information: An adaptive dynamic event-triggered approach. *IEEE Transactions on Automatic Control*, 68(10):5862–5876, 2023. doi: 10.1109/TAC.2022.3226142.
- Maojiao Ye and Guoqiang Hu. Distributed nash equilibrium seeking by a consensus based approach. *IEEE Transactions on Automatic Control*, 62(9):4811–4818, 2017.
- Mu Zhu, Ahmed H. Anwar, Zelin Wan, Jin-Hee Cho, Charles A. Kamhoua, and Munindar P. Singh. A survey of defensive deception: Approaches using game theory and machine learning. *IEEE Communications Surveys and Tutorials*, 23(4):2460–2493, 2021. doi: 10.1109/COMST.2021.3102874.