Non-Holonomic Gradient Play: Stability and Deception in Games on Manifolds

Mahmoud Abdelgalil, Miroslav Krstic, and Jorge I. Poveda

***Abstract*— We study equilibrium seeking for multi-agent systems whose joint state evolves on a manifold and whose agents are constrained to act through state-dependent control vector fields. Under a natural bundle-splitting assumption, we show that the averaged interaction dynamics admit an intrinsic characterization as a sum of projected Riemannian gradients, a form of *non-holonomic gradient play*, and we analyze the structural and asymptotic stability of equilibria via an intrinsic linearization that reveals the role of the Riemannian connection and the non-holonomy of the actuation frame. Our results subsume existing gradient play characterizations in Euclidean and finite-dimensional manifold settings as special cases, while accommodating non-holonomic control vector fields. We then introduce a geometric theory of *deception* for this setting, wherein a malicious agent, with access to privileged information about others' learning rules, exploits such information to shift the equilibrium of the interaction. We show that deception acts by tilting the Riemannian geometry of the manifold and, consequently, distorting the projected gradients followed by victim agents. We prove that exponential stability of the deception-free equilibrium is robustly preserved over an open set of deception parameters, and we derive an explicit first-order formula for the equilibrium shift induced by the deceiver. Analytical examples are provided to illustrate the presented theory.**

***Index Terms*— Multi-agent learning, games on manifolds, equilibrium seeking, deception.**

I. Introduction

STRATEGIC interactions between autonomous decision-making agents are often constrained by geometry, with examples spanning many practical engineering problems, e.g., rigid body attitude coordination [1], [2], distributed resource allocation [3], [4], and synchronization of coupled oscillators [5]. These settings have two distinguishing structural features. First, the joint state of the interaction evolves on a manifold rather than on a vector space. Second, each agent can influence a restricted set of directions of the joint state, and those admissible directions typically vary with the state itself. That is, each agent acts through a family of state-dependent control vector fields. If, in addition, the agents pursue different payoffs, the natural model of their interaction is *gradient play* wherein each agent seeks to maximize its own payoff along the vector fields it controls. Therefore, game-theoretic notions of equilibrium are typically invoked to characterize the behavior of the interaction.

The earliest study of gradient play in games with continuous strategy spaces seems to be [6], [7], where differential characterizations of equilibria and of their local stability are given for games whose strategy spaces are finite-dimensional manifolds. However, the learning dynamics considered therein are expressed via partial derivatives of the payoffs with respect to a seemingly arbitrary choice of coordinates. Since gradient vector fields on a manifold are defined only relative to a Riemannian metric, and no such metric is introduced in [6], [7], neither the dynamics nor the ensuing stability conditions are coordinate independent. More significantly, the formulation in [6], [7] presupposes that each agent actuates the coordinates of its own strategy space directly and independently. In particular, the framework in [6], [7] does not accommodate agents that act through a restricted family of vector fields, let alone a non-commuting one, even if the family collectively spans the tangent bundle of the manifold. To the best of our knowledge, an intrinsic theory of gradient play in games whose joint state evolves on a manifold and whose agents are constrained to act through state-dependent, possibly non-commuting, vector fields, does not exist in the literature. In parallel, there has been an active research thread on model-free equilibrium seeking [3], [8]–[11] which dispenses with explicit gradient information altogether and relies on zeroth-order information only, i.e., payoff measurements, to approximate gradient play. However, it has remained confined to Euclidean strategy spaces with trivial actuation structures.

The first contribution of the present manuscript is precisely an intrinsic theory of gradient play under such constraints. Specifically, under a natural bundle-splitting assumption on the agents' actuation directions, we construct a Riemannian metric on the state space with respect to which the interaction dynamics amounts to a sum of *projected Riemannian gradients*, one per agent, with each projection being onto the distribution spanned by the corresponding agent's control vector fields. We refer to the resulting dynamics as *non-holonomic gradient play*, and to the corresponding vector field as the *non-holonomic pseudo-gradient*. This representation yields an equilibrium characterization that is coordinate independent. In particular, equilibria are determined solely by the payoffs and the geometry of the actuation directions. We then characterize the structural, as well as the asymptotic, stability of equilibria via an intrinsic linearization of the non-holonomic

M. Abdelgalil (maabdelg@buffalo.edu) is with the Department of Mechanical and Aerospace Engineering, University at Buffalo, State University of New York, Buffalo, NY, 14260, USA. Miroslav Krstic (mkrstic@ucsd.edu) is with the Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA 92093, USA. J. I. Poveda (jipoveda@ucsd.edu) is with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA.

pseudo-gradient, which decomposes into two types of terms: Riemannian Hessian terms, and *Connection* terms. The latter type encodes the non-holonomy of the actuation frame and has no counterpart in the coordinate-based analysis of [6], [7]. Sufficient stability conditions of the block Gershgorin type are derived, and two special cases are worked out in detail. First, when each agent's distribution is involutive, we show that the diagonal blocks of the linearization reduce to intrinsic Hessians of restricted payoffs. When, in addition, the agents' vector fields mutually commute, the state space acquires a local product structure and the characterizations of [6], [7] are recovered as the flat special case. The identical interest game [12], wherein the dynamics reduces to an intrinsic gradient flow, is likewise subsumed.

The second contribution concerns *deception* under asymmetric information, and it is here that the generality of the first contribution proves instrumental. Specifically, in contested and adversarial environments, the assumption that all agents are equally informed is frequently violated. Indeed, an agent may be able to measure, in real time, the inputs applied by its competitors, or may posses privileged knowledge of their learning rules. Recent work [13] showed that, on Euclidean product strategy spaces with additive actuation structures, such asymmetry enables a malicious agent with access to privileged information to relocate the equilibrium of the interaction by manipulating only its own inputs, without tampering with other agents' measurements or actuation. The mechanism of [13], however, is tailored to the additive structure of $\mathbb{R}^n$ and does not extend to the geometric setting considered here. Within the present framework, we show that deception admits an intrinsic geometric description: by embedding the victim's exploration signals into its own inputs, a deceiver *tilts* the victim's effective actuation bundle along its own, and, thereby, distorts the Riemannian projections of the victim in the non-holonomic gradient. As a result, the tilted bundles generically fail to be involutive even when the nominal ones are, and, therefore, analyzing deception necessitates the general theory we develop herein. In particular, we prove that exponentially stable equilibria persist, together with their stability, over an open set of deception parameters, and we derive an explicit first-order formula for the equilibrium shift as a function of those parameters.

As will be discussed in the sequel, the interaction dynamics studied in this manuscript arises as the Lie-bracket approximation [14]–[16] of a control-affine system driven by payoff-only feedback laws. In this sense, the theory developed herein describes the behavior of model-free equilibrium seeking along the lines of [3], [8]–[11], but for games on manifolds and with actuation constraints. Nevertheless, our results are stated for the target dynamics directly and are independent of the particular mechanism that generates them. In particular, any learning algorithm whose target behavior takes the form of non-holonomic gradient play, whether obtained through extremum seeking, stochastic approximation, or otherwise, falls within the scope of our formulation.

The remainder of the manuscript is organized as follows. Section II collects notation and geometric preliminaries. Section III presents the interaction model, its averaged dynamics, and the information-asymmetry model. Section IV develops the intrinsic theory of non-holonomic gradient play: the constructed metric, the equilibrium characterization, the intrinsic linearization, and the involutive and commuting special cases, illustrated throughout by two running examples on the sphere and on the simplex. Section V develops the theory of deception: the tilting mechanism, the persistence of stable equilibria, and the first-order sensitivity of the equilibrium map. Section VI collects the proofs of the main results, and Section VII concludes with directions for future work.

II. Notation

Throughout, we use standard differential geometric notation consistent with, e.g., [17]. Let $\mathcal{M}$ be a smooth n-dimensional manifold, $T\mathcal{M}$ be its tangent bundle, $T^*\mathcal{M}$ its cotangent bundle. We use $\Gamma(T\mathcal{M}\otimes\cdots\otimes T\mathcal{M}\otimes T^*\mathcal{M}\otimes\cdots\otimes T^*\mathcal{M})$ to denote the space of smooth sections of the bundle $T\mathcal{M}\otimes\cdots\otimes T\mathcal{M}\otimes T^*\mathcal{M}\otimes\cdots\otimes T^*\mathcal{M}$, so that the space of smooth vector fields is $\Gamma(T\mathcal{M})$, the space of smooth differential 1-forms is $\Gamma(T^*\mathcal{M})$, and so on. For brevity, we also use $\mathfrak{X}(\mathcal{M})$ to denote the space of smooth vector fields on $\mathcal{M}$, and $\Omega(\mathcal{M})$ to denote the space of smooth differential 1-forms on $\mathcal{M}$. The symbol $\otimes$ denotes *tensor product* so that, if $f_1, f_2 \in \mathfrak{X}(\mathcal{M})$ and $\omega_1, \omega_2 \in \Omega(\mathcal{M})$, then $f_1 \otimes f_2 \in \Gamma(T\mathcal{M}\otimes T\mathcal{M})$, $\omega_1\otimes\omega_2 \in \Gamma(T^*\mathcal{M}\otimes T^*\mathcal{M})$, $\omega_i \otimes f_i \in \Gamma(T^*\mathcal{M}\otimes T\mathcal{M})$, and so on. For $f\in\mathfrak{X}(\mathcal{M})$ and J a smooth function, the Lie derivative of J along f is

$$L_f J(x) := dJ(x)\big[f(x)\big],$$

where $dJ\in\Omega(\mathcal{M})$ denotes the differential of J. A Riemannian metric on $\mathcal{M}$, which is a symmetric positive definite section of $\Gamma(T^*\mathcal{M}\otimes T^*\mathcal{M})$, is denoted by g. Given a smooth function J, the Riemannian gradient with respect to g, denoted by $\mathbf{grad}(J)$, is the unique vector field satisfying

$$g(\mathbf{grad}(J), X) = dJ[X], \quad \forall X\in\mathfrak{X}(\mathcal{M}).$$

The Riemannian *Hessian* of J is the unique symmetric section of $\Gamma(T^*\mathcal{M}\otimes T^*\mathcal{M})$ that satisfies

$$\mathbf{Hess}(J)[X,Y] := g(\nabla_X \mathbf{grad}(J), Y),$$

where $\nabla : \mathfrak{X}(\mathcal{M})\times\mathfrak{X}(\mathcal{M})\to\mathfrak{X}(\mathcal{M})$ is the Levi-Civita connection associated with the metric g. The fact that the Riemannian Hessian is symmetric, i.e., $\mathbf{Hess}(J)[X,Y] = \mathbf{Hess}(J)[Y,X]$, follows from the torsion-free nature of the Levi-Civita connection. If $\mathcal{E}\subset T\mathcal{M}$ is a smooth sub-bundle and $P_{\mathcal{E}} : T\mathcal{M}\to\mathcal{E}$ is the g-orthogonal projection, then we define the *projected gradient*

$$\mathbf{grad}_{\mathcal{E}}(J) := P_{\mathcal{E}}\,\mathbf{grad}(J).$$

Finally, we use $\cdot|_x$ to denote the evaluation operation, i.e., if $f\in\mathfrak{X}(\mathcal{M})$, then $f|_x = f(x)\in T_x\mathcal{M}$, and so on.

III. Problem Formulation

We model the state-space of the strategic environment as a smooth n-dimensional manifold $\mathcal{M}$. The i^{th}-agent can only influence the state of the environment $x\in\mathcal{M}$ through a collection of vector fields $\{f_{i,j}\}_{j=1}^{d_i}\subset\mathfrak{X}(\mathcal{M})$ by manipulating

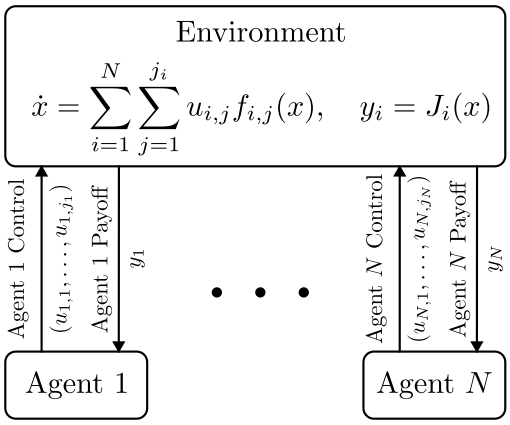


Fig. 1. Block diagram of the strategic interaction between N independent decision-making agents coupled only through the environment.

its input vector $(u_{i,1}, \ldots, u_{i,d_i}) \in \mathbb{R}^{d_i}$. Therefore, in the presence of N independent decision-making agents, and under a control-affine structure, the evolution of the environment state is governed by the control-affine system

$$\dot{x} = \sum_{i=1}^{N} \sum_{j=1}^{d_i} u_{i,j} f_{i,j}(x). \tag{1}$$

In the non-cooperative setting, each agent is interested in their individual payoff which may, in general, be different from other agents' payoffs. We model this situation by a collection of sufficiently regular functions $\{J_i\}_{i=1}^{N}$. A block diagram description of the model is shown in Figure 1.

When the decision-making agents have access to measurements of their own payoff only, an agent can design their control inputs so as to drive the state of the environment x in the directions that maximize their respective individual payoff. The natural class of algorithms in this case is the framework of Lie-Bracket Approximation Extremum Seeking (LieBrA-ES) [14]. Under this design choice, the inputs of the i^{th}-agent may be taken as the feedback law

$$\begin{aligned} u_{i,j}(x,t) = &\sqrt{4\pi\kappa_i\omega_{i,j}\omega}\cos(2\pi\omega_{i,j}\omega t)F^1_{i,j}(J_i(x)) \\ &- \sqrt{4\pi\kappa_i\omega_{i,j}\omega}\sin(2\pi\omega_{i,j}\omega t)F^2_{i,j}(J_i(x)), \end{aligned} \tag{2}$$

where $\omega \in \mathbb{R}_{>0}$, the constants $\cup_{i=1}^{N}\{\omega_{i,j}\}_{j=1}^{d_i} \subset \mathbb{Q}_{>0}$ are any collection of positive rational numbers such that $\omega_{i,j} = \omega_{i',j'}$ if and only if $i = i'$ and $j = j'$, the constants $\kappa_i > 0$ are positive gains, and $F^k_{i,j} : \mathbb{R} \to \mathbb{R}$ is any collection of sufficiently regular functions that satisfy

$$F^2_{i,j}(y) = -F^1_{i,j}(y) \int F^1_{i,j}(y)^{-2}\mathrm{d}y. \tag{3}$$

An example of such functions is

$$F^1_{i,j}(y) := \cos(y), \qquad F^2_{i,j}(y) := -\sin(y), \tag{4}$$

which was introduced in [18], see also [15], [19]. Regardless of which choice is implemented, the behavior of the control-affine system (1) under the output feedback law (2) is known [14], [16] to approximate the behavior of an associated *Lie Bracket Approximation* (LieBrA) system

$$\dot{x} = \sum_{i=1}^{N} \kappa_i \sum_{j=1}^{d_i} L_{f_{i,j}}(J_i)(x) f_{i,j}(x), \tag{5}$$

in the limit $\omega \gg 1$. In particular, the stability properties of the system (5) are inherited by the closed-loop system (1)-(2) in a *practical* sense [14], [16], [20]. Therefore, it is of great interest to characterize the asymptotic behavior of system (5), a goal we pursue in Section IV.

The model presented thus far, i.e. the closed-loop system (1)-(2), is formulated under the assumption of *information symmetry*, i.e., that none of the decision-making agents has access to *privileged* information concerning the feedback laws implemented by other agents. As shown recently in [13], when this assumption is violated, i.e., in the presence of *information asymmetry*, agents with access to privileged information can manipulate the asymptotic behavior of the strategic interaction dynamics (1) through careful manipulation of their own control inputs only. In particular, such *deceptive* agents can *bias* the steady-state behavior of the system (1) without ever interfering with the control actions of the *oblivious agents* or their measurements. The context in which this *deception* phenomenon was studied in [13] is that of Nash-Seeking in repeated static non-cooperative games on flat spaces. Herein, we will study a natural generalization of deception in adaptive multi-agent decision-making to the setting captured by (1).

For the sake of simplicity, we focus on the case of two agents and introduce the following model (see Remark 6 for the extension to $N > 2$ agents). We assume that the 1^{st}-agent, i.e., the deceptive agent, can measure the control input (2) of the 2^{nd}-agent, i.e., the *oblivious* agent. We also assume that the deceptive agent, utilizing this privileged information, adjusts its feedback law to

$$\tilde{u}_{1,j}(x,t) := u_{1,j}(x,t) + \sum_{k=1}^{d_2} \gamma^j_k u_{2,k}(x,t), \tag{6}$$

where $\{\gamma^j_k\}$ is a collection of tunable inputs controlled by the deceptive agent. Therefore, the closed-loop system defined by (1) and (6) can be re-written as

$$\dot{x} = \sum_{i=1}^{2} \sum_{j=1}^{d_i} u_{i,j}(x,t)\tilde{f}_{i,j}(x), \tag{7a}$$

where the vector fields $\tilde{f}_{i,j}$ are given by

$$\tilde{f}_{1,j} = f_{1,j}, \quad \tilde{f}_{2,j} := f_{2,j} + \sum_{k=1}^{d_1} \gamma^k_j f_{1,k}, \tag{7b}$$

and, consequently, the associated Lie-Bracket system is

$$\dot{x} = \sum_{i=1}^{2} \kappa_i \sum_{j=1}^{d_i} L_{\tilde{f}_{i,j}}(J_i)(x)\tilde{f}_{i,j}(x). \tag{8}$$

Notably, the system (8) shares the same structure as (5) but differs crucially since the vector fields $\tilde{f}_{2,j}$ are *tilted* from their original value by the deception parameters γ^k_j according to the formula (7b). In particular, we expect that the equilibria of (5)

and (8), when they exist, will, generically, coincide only for the case $\gamma_j^k \equiv 0$. The behavior of (8) is studied in Section V.

We conclude the current section with a discussion on the assumption of asymmetric information adopted herein. To that end, it is worth mentioning that the particular form of information asymmetry adopted in [13] is *tailored* for the classical Extremum-Seeking based Nash-Seeking on $\mathbb{R}^n$. Consequently, the same assumption cannot be used as is to study deception in the context of the model (1) since it relies on the *additive structure* of Euclidean spaces. Moreover, the type of interactions studied in [13] necessitates that the vector fields $f_{i,j}$ are constant and completely decoupled, i.e., that $\mathcal{M} = \mathbb{R}^{j_1} \times \cdots \times \mathbb{R}^{j_N} \simeq \mathbb{R}^{j_1+\cdots j_N}$, and $f_{i,j}(x) = \partial_{(i-1)d_i+j}$. Herein, we relax both of these assumptions as evidenced by the general model (1), and, as a result, we require a stronger notion of information asymmetry to obtain a suitable generalization of the deception mechanism in [13]. Possible relaxations of this assumption will be pursued in future work.

IV. Equilibria and Their Stability

In this section, we focus our attention on system (5), repeated below for convenience,

$$\dot{x} = \sum_{i=1}^{N} \kappa_i \sum_{j=1}^{d_i} L_{f_{i,j}}(J_i)(x) f_{i,j}(x). \tag{5}$$

Characterizing the asymptotic behavior of (5) necessitates identifying its equilibria as the first step. However, without further assumptions on the vector fields $f_{i,j}$, finding the equilibria of (5) requires solving the general system of equations

$$\sum_{i=1}^{N} \kappa_i \sum_{j=1}^{d_i} L_{f_{i,j}}(J_i)(x) f_{i,j}(x) = 0. \tag{9}$$

An issue with (9) is the fact that, in general, the equilibria of (9) will depend on the choice of the gains κ_i, a choice that is arbitrary and independent of the functions J_i. Nevertheless, there are natural assumptions that allow us to characterize those equilibria that satisfy (9) for any choice of the gains κ_i. To that end, we impose the following assumption.

Assumption 1: There exists an open sub-manifold $\mathcal{U} \subset \mathcal{M}$ such that, for all $x \in \mathcal{U}$, we have

$$T_x\mathcal{U} = \mathcal{E}_1(x) \oplus \cdots \oplus \mathcal{E}_N(x), \tag{10}$$

where $\mathcal{E}_i$ is a *regular* distribution on $\mathcal{U}$ defined by the span of the i^{th}-family of vector fields $\{f_{i,j}\}_{j=1}^{d_i}$, i.e.,

$$\mathcal{E}_i(x) := \operatorname{span}_{\mathbb{R}}\{f_{i,j}(x)\}_{j=1}^{d_i}.$$

□

In words, Assumption 1 asks that the family of vector fields $\cup_{i=1}^{N}\{f_{i,j}\}_{j=1}^{d_i}$ spans the tangent space of the manifold $\mathcal{M}$ for all points in an open set $\mathcal{U}$ and that the intersection of the spans of any two distinct families $\{f_{i,j}\}_{j=1}^{d_i}$ and $\{f_{i',j}\}_{j=1}^{d_{i'}}$ is trivial whenever $i \neq i'$. In the context of the multi-agent strategic interaction modeled by (1), Assumption 1 essentially requires that the control vector fields influenced by the i^{th}-agent are linearly independent from the span of the control vector fields influenced by every other agent at every environment state $x \in \mathcal{U}$ and that, collectively, the family $\cup_{i=1}^{N}\{f_{i,j}\}_{j=1}^{d_i}$ spans the entire tangent bundle $T\mathcal{U}$. We now illustrate this assumption with some examples. In these examples, the distributions $\mathcal{E}_i$ are all rank 1, i.e., $d_i = 1$ for all $i \in \{1, \ldots, N\}$. To simplify notation, we use $e_i := f_{i,1}$ to get rid of superfluous indices.

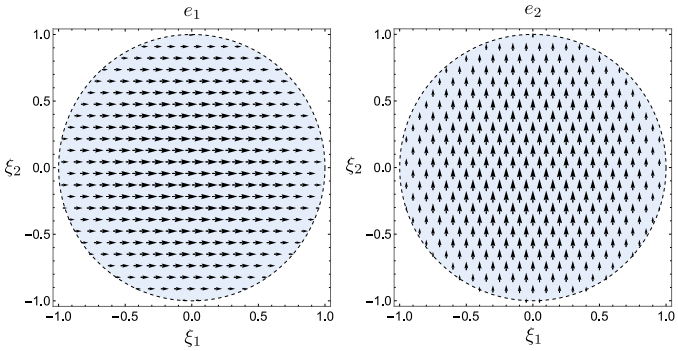


Fig. 2. An illustration of the vector fields $\boldsymbol{e_i}$ for Example 1 in the $\boldsymbol{\xi}$-coordinates. The arrows' sizes represent the magnitude of the vector field.

Example 1: We consider a two-player, i.e. $N = 2$, decision-making problem on the sphere $\mathcal{M} := \mathbb{S}^2 = \{x \in \mathbb{R}^3 \mid x^\top x = 1\} \subset \mathbb{R}^3$, which is a 2-dimensional Euclidean sub-manifold. Let the vector fields e_i be defined in the ambient coordinates $x = (x_1, x_2, x_3)$ by

$$e_i(x) := x_3\partial_i - x_i\partial_3,$$

which can be shown to be tangent to $\mathbb{S}^2$. To verify Assumption 1 on a suitable submanifold $\mathcal{U} \subset \mathcal{M}$, we define the function $\Delta(x) := \langle x, e_1(x) \times e_2(x)\rangle_{\mathbb{R}^3}$, where $\times$ is the cross product of vectors and $\langle\cdot,\cdot\rangle_{\mathbb{R}^3}$ is the inner product between vectors in $\mathbb{R}^3$. Direct computation gives that

$$\Delta(x) = x_3(x_1^2 + x_2^2 + x_3^2) = x_3.$$

Therefore, since the cross product of a pair of vectors in $\mathbb{R}^3$ vanishes if and only if they are linearly dependent, it follows that the vector fields e_i are linearly independent and span the tangent space of $\mathbb{S}^2$ everywhere except at the degenerate set $\{x \in \mathbb{S}^2 \mid x_3 = 0\}$. In particular, by restricting our analysis to the open sub-manifold

$$\mathcal{U} := \{x \in \mathbb{S}^2 \mid x_3 > 0\},$$

we are guaranteed that, if $\mathcal{E}_i(x) := \operatorname{span}\{e_i(x)\}$, then Assumption 1 holds on $\mathcal{U}$. Moreover, the submanifold $\mathcal{U}$ is diffeomorphic to the open disk $\mathbb{D}^2 := \{\xi = (\xi_1, \xi_2) \in \mathbb{R}^2 \mid \xi_1^2 + \xi_2^2 < 1\}$ via the diffeomorphism $(\xi_1, \xi_2) = \varphi(x) = (x_1, x_2)$, which will be convenient for visualizations. For instance, the vector fields e_i are depicted in Figure 2 using the ξ-coordinates. It is also worth mentioning that the vector fields e_i do not commute. Indeed, direct computation gives

$$[e_1, e_2](x) = x_2\partial_1 - x_1\partial_2,$$

which, generically, is non-vanishing on $\mathcal{U}$. □

Example 2: Consider a two-player, i.e., $N = 2$, resource allocation problem on the 2-simplex $\mathcal{M} = \mathcal{S}^2 := \{x \in \mathbb{R}^3 \mid x_i \geq 0, \sum_{i=1}^{3} x_i = 1\}$, which is a 2-dimensional

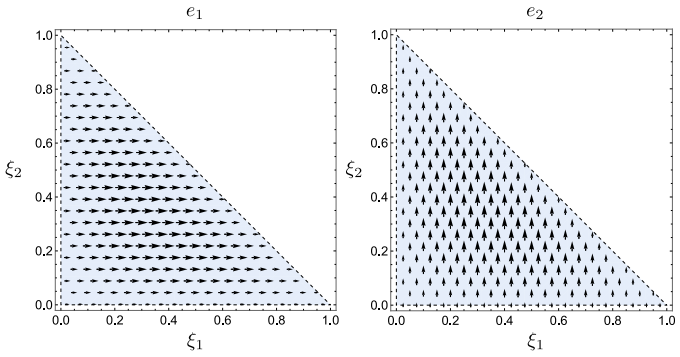


Fig. 3. An illustration of the vector fields e_i for Example 2 in the ξ-coordinates. The arrows' sizes represent the magnitude of the vector field.

Euclidean submanifold. Let $N = 2$ and consider the vector fields e_i defined in the ambient coordinates as follows:

$$e_i(x) := x_1x_2x_3(\partial_i - \partial_3).$$

Direct computation shows that the vector fields e_i are tangent to $\mathcal{M}$ and span the tangent space everywhere except at the degenerate set $\{x \in \mathcal{M} \mid x_1x_2x_3 = 0\}$, which corresponds to the topological boundary of $\mathcal{M}$, denoted by $\partial\mathcal{M}$. Hence, by restricting our analysis to the open submanifold $\mathcal{U} := \mathcal{M}\backslash\partial\mathcal{M}$, we are guaranteed that, with $\mathcal{E}_i = \text{span}\{e_i\}$, then Assumption 1 holds on $\mathcal{U}$. Moreover, similar to the previous example, the submanifold $\mathcal{U}$ is diffeomorphic to an open subset of the plane via the diffeomorphism $(\xi_1, \xi_2) = \varphi(x) = (x_1, x_2)$, which will be convenient for visualizations. Indeed, the vector fields e_i are depicted in Figure 3. We also point out that the vector fields e_i do not commute. In fact, direct computation gives

$$[e_1, e_2](x) = x_1x_2x_3(x_1(x_3 - x_2)\partial_1 + x_2(x_1 - x_3)\partial_2 + x_3(x_2 - x_1)\partial_3),$$

which, generically, is non-vanishing on $\mathcal{U}$.

□

We will be mainly interested in the behavior of the dynamical system (5) within the open sub-manifold $\mathcal{U}$. As such, Assumption 1 allows us to endow system (5) with a natural geometric interpretation: the flow of (5) is a Riemannian pseudo-gradient flow on $\mathcal{M}$.

Theorem 1: Let Assumption 1 be satisfied. Then, there exists a Riemannian metric g on the sub-manifold $\mathcal{U} \subset \mathcal{M}$ such that system (5) takes the form

$$\dot{x} = \sum_{i=1}^{N} \kappa_i \mathbf{grad}_{\mathcal{E}_i}(J_i)(x), \tag{11}$$

where $\mathbf{grad}_{\mathcal{E}_i}(J_i)$ is the vector field obtained through orthogonal projection of $\mathbf{grad}(J_i)$ with respect to g onto $\mathcal{E}_i$.

□

In view of (11), we refer to the averaged interaction dynamics as *non-holonomic gradient play*: each agent ascends its own payoff, but only along the state-dependent, and generically non-commuting, directions available to it. In the special case $\mathcal{M} = \mathbb{R}^n$ with $f_{i,j} = \partial_{(i,j)}$ in the standard coordinates, system (11) reduces to the classical gradient play dynamics studied in [6], [7].

The proof of Theorem 1 is constructive in the sense that it explicitly defines the Riemannian metric on the sub-manifold $\mathcal{U}$ for which the conclusions of the Theorem hold. Moreover, if the sub-manifold $\mathcal{U}$ contains any equilibria of the system (5), then, thanks to Assumption 1, the condition (9) simplifies to the natural geometric condition

$$\mathbf{grad}_{\mathcal{E}_i}(J_i)\big|_{x^\star} = 0, \qquad i \in \{1, \dots, N\}. \tag{12}$$

Notably, unlike (9), the condition (12) is independent of the gains κ_i, i.e., no agent can change the equilibria of (5) in the sub-manifold $\mathcal{U}$ by amplifying their gain κ_i, which implies that such equilibria are solely determined by the payoffs J_i and the family of vector fields $\cup_{i=1}^{N}\{f_{i,j}\}_{j=1}^{d_i}$, i.e., by the properties of the environment (cf. Figure 1). Therefore, we may take $\kappa_i = 1$ for all i, which is equivalent to absorbing the constants κ_i into the payoff functions, without affecting the equilibria of (5) in the sub-manifold $\mathcal{U}$.

Having characterized the equilibria of the interaction dynamics (5), we now proceed to study their nature. Specifically, we would like to characterize sufficient conditions under which an equilibrium point, i.e., a point satisfying (12), is *isolated* and locally exponentially stable. This is achieved using the *intrinsic Jacobian*.

The intrinsic Jacobian of the vector field defining (5) is the $(1,1)$-tensor $A : \mathfrak{X}(\mathcal{U}) \to \mathfrak{X}(\mathcal{U})$ defined by

$$A[X] := \nabla_X \left(\sum_{i=1}^{N} \mathbf{grad}_{\mathcal{E}_i}(J_i) \right), \tag{13}$$

for any $X \in \mathfrak{X}(\mathcal{U})$, where we utilize the linearity of the Levi-Civita connection. The local properties of an equilibrium point $x^\star \in \mathcal{U}$ satisfying (12) is dictated by the *eigenvalues* of the linear map $A|_{x^\star} : T_{x^\star}\mathcal{U} \ni v \to A|_{x^\star}[v] \in T_{x^\star}\mathcal{U}$, i.e. the restriction of the Jacobian defined in (13) to the tangent space at the point $x^\star$. In particular, we have the following proposition.[1]

Proposition 1: Let Assumption 1 be satisfied for some $\mathcal{U} \subset \mathcal{M}$ and let $x^\star \in \mathcal{U}$ be such that (12) holds. Then, the following statements are true:

1) The point $x^\star$ is an isolated equilibrium for system (5) **if** the map $A|_{x^\star}$ is an automorphism of $T_{x^\star}\mathcal{U}$.
2) The point $x^\star$ is locally exponentially stable for system (5) **if and only if** the map $A|_{x^\star}$ is Hurwitz.

□

Before we proceed to analyze the structure of the Jacobian (13), it is illuminating to consider the case where $J_1 = J_2 = \cdots = J_N = J$, i.e., the *identical interest* game [12] in which all agents share the same payoff. In that case, system (5) is, essentially, an intrinsic gradient flow. In addition, the equilibrium condition (12) and the operator (13) simplify to

$$\mathbf{grad}(J)|_{x^\star} = 0, \qquad A[X] = \nabla_X \mathbf{grad}(J),$$

where, since $\{P_{\mathcal{E}_i}\}_{i=1}^{N}$ are a complete set of projection operators, we used the property that $\sum_{i=1}^{N} P_{\mathcal{E}_i} = I$. Applying

[1]The proof is classical and is omitted.

Proposition 1 to this special case, we recover the standard conditions of optimization in the geometric setting. In particular, item 2) in Proposition 1 is equivalent to the condition that the Hessian is negative definite at $x^\star$.

Given the discussion above, it is tempting to extend the intuition provided by the identical interest case to the general case. However, as will be shown in the sequel, the curvature induced by the geometry of the bundles $\mathcal{E}_i$ contributes in a non-trivial way to the Jacobian (13). Indeed, recalling that $\mathbf{grad}_{\mathcal{E}_i}(J_i) := P_{\mathcal{E}_i}[\mathbf{grad}(J_i)]$, where the $(1,1)$-tensor $P_{\mathcal{E}_i}$ is the orthogonal projection operator with respect to the metric onto the distribution $\mathcal{E}_i$, and utilizing the product rule of the Levi-Civita connection, we see that

$$A = \sum_{i=1}^{N} P_{\mathcal{E}_i}[\nabla \mathbf{grad}(J_i)] + (\nabla P_{\mathcal{E}_i})[\mathbf{grad}(J_i)]. \tag{14}$$

From (14), it is immediately clear that the curvature of the bundles $\mathcal{E}_i$ has a non-negligible contribution to the Jacobian. Indeed, the second term in (14) does not vanish generically at an equilibrium point $x^\star$ since condition (12) may hold at a point $x^\star \in \mathcal{U}$ even if $\mathbf{grad}(J_i)|_{x^\star} \neq 0$. Using the orthogonal splitting (10), we may decompose A into a *block* form by defining the linear maps

$$A_{ij}[\cdot] = \sum_{k=1}^{N} P_{\mathcal{E}_i}\Big[P_{\mathcal{E}_k}[\nabla_{P_{\mathcal{E}_j}[\cdot]}\mathbf{grad}(J_k)] + (\nabla_{P_{\mathcal{E}_j}[\cdot]}P_{\mathcal{E}_k})[\mathbf{grad}(J_k)]\Big].$$

Since $\mathcal{E}_i$ are orthogonal, it follows that $P_{\mathcal{E}_i}P_{\mathcal{E}_k} = \delta_{ik}P_{\mathcal{E}_k}$, and, therefore, the tensor fields A_{ij} simplify to

$$A_{ij}[\cdot] = P_{\mathcal{E}_i}[\nabla_{P_{\mathcal{E}_j}[\cdot]}\mathbf{grad}(J_i)] + C_{ij}[\cdot], \tag{15a}$$

where C_{ij} are the tensor fields defined by

$$C_{ij}[\cdot] := \sum_{k=1}^{N} P_{\mathcal{E}_i}[(\nabla_{P_{\mathcal{E}_j}[\cdot]}P_{\mathcal{E}_k})[\mathbf{grad}(J_k)]]. \tag{15b}$$

From this block form, it is possible to derive a sufficient condition to ensure that the real parts of the eigenvalues of the linear map $A|_{x^\star}$ are strictly negative and, therefore, that $x^\star$ is both isolated and locally exponentially stable. The following Proposition formalizes this sufficient condition:

Proposition 2: Let $x^\star \in \mathcal{U}$ satisfy (12) and let A_{ij} be the blocks defined in (15). For each pair (i,j), define $\|A_{ij}\|_g$ as the operator norm induced by g, i.e.,

$$\|A_{ij}\|_g := \sup_{v \in \mathcal{E}_j(x^\star)\setminus\{0\}} \frac{\|A_{ij}v\|_g}{\|v\|_g},$$

and, for a linear map M on $\mathcal{E}_i(x^\star)$, let $\sigma^g_{\min}(M) := \inf_{v\in\mathcal{E}_i(x^\star)\setminus\{0\}} \|Mv\|_g/\|v\|_g$. Suppose that there exist positive constants $\chi_i > 0$ $(i = 1,\dots,N)$ such that, for each i,

$$\inf_{\Re(s)\geq 0} \sigma^g_{\min}(A_{ii}|_{x^\star} - sI_i) \;>\; \sum_{j\neq i} \|A_{ij}|_{x^\star}\|_g \,\frac{\chi_j}{\chi_i}, \tag{16}$$

where I_i denotes the identity map on $\mathcal{E}_i(x^\star)$. Then $A|_{x^\star}$ is an automorphism and Hurwitz.

□

In the remainder of this section, we impose the following assumption.

Assumption 2: For each $i \in \{1,\dots,N\}$, the family $\{f_{i,j}\}_{j=1}^{d_i}$ is linearly independent at every $x \in \mathcal{U}$.

□

Under this hypothesis, the Riemannian metric g on $\mathcal{U}$ constructed in the proof of Theorem 1 makes the family

$$\{f_{i,j} : 1 \leq i \leq N,\ 1 \leq j \leq d_i\}$$

a global orthonormal frame on $\mathcal{U}$ under the metric g. Note, however, that this Assumption is not too restrictive and is only made for convenience. Indeed, by restricting the submanifold $\mathcal{U} \subset \mathcal{M}$ if needed, we may always select a set of independent vector fields that span each distribution $\mathcal{E}_i$.

Next, we proceed to define $\{\vartheta^{i,j}\}$ as the co-frame dual to

$$\{f_{i,j} : 1 \leq i \leq N,\ 1 \leq j \leq d_i\},$$

i.e., the unique set of differential 1-forms on $\mathcal{U}$ that satisfy

$$\vartheta^{i,j}[f_{k,\ell}] = \delta_{ik}\,\delta_{j\ell}. \tag{17}$$

With this definition, the metric and the orthogonal projection operators onto the distributions $\mathcal{E}_i$ take the form

$$g(v,w) = \sum_{i,j} \vartheta^{i,j}[v]\,\vartheta^{i,j}[w], \quad P_{\mathcal{E}_i} = \sum_{j=1}^{d_i} \vartheta^{i,j} \otimes f_{i,j}. \tag{18}$$

In addition, for any smooth function J,

$$\mathbf{grad}(J) = \sum_{i=1}^{N} \mathbf{grad}_{\mathcal{E}_i}(J), \quad \mathbf{grad}_{\mathcal{E}_i}(J) = \sum_{j=1}^{d_i} dJ[f_{i,j}]\,f_{i,j}. \tag{19}$$

We then have the following Theorem.

Theorem 2: With respect to the frame $\{f_{i,j}\}$, the linear map $A|_{x^\star}$ admits the block matrix representation

$$A|_{x^\star} = \big[\, A_{ij}|_{x^\star} \,\big]_{i,j=1}^{N}, \qquad A_{ij}|_{x^\star} \in \mathbb{R}^{d_i \times d_j},$$

where the entries of the matrices $A_{ij}|_{x^\star}$ are given by

$$\begin{aligned}\big(A_{ij}|_{x^\star}\big)_{p,q} &= d\big(dJ_i[f_{i,p}]\big)[f_{j,q}]\Big|_{x^\star} \\ &= \mathbf{Hess}(J_i)\big[f_{i,p}, f_{j,q}\big]\Big|_{x^\star} + dJ_i\big[\nabla_{f_{j,q}} f_{i,p}\big]\Big|_{x^\star},\end{aligned} \tag{20}$$

for $p \in \{1,\dots,d_i\}$ and $q \in \{1,\dots,d_j\}$, and where $\mathbf{Hess}(J_i)$ denotes the Hessian of J_i with respect to g.

□

Remark 1: Formula (20) splits each block of the Jacobian into a Hessian term and a *connection* term $dJ_i[\nabla_{f_{j,q}} f_{i,p}]$. The connection term encodes the non-holonomy of the actuation frame $\{f_{i,j}\}$ and, in general, it does not vanish at $x^\star$, since (12) does not imply $\mathbf{grad}(J_i)|_{x^\star} = 0$. Nevertheless, there are two instructive cases with some simplifications. If the inter-block connection coefficients vanish, i.e., $\nabla_{f_{j,q}} f_{i,p} \equiv 0$ whenever $i \neq j$, then the off-diagonal blocks are pure Hessian blocks,

$$\big(A_{ij}|_{x^\star}\big)_{p,q} = \mathbf{Hess}(J_i)[f_{i,p}, f_{j,q}]\big|_{x^\star}, \qquad i \neq j,$$

while the diagonal blocks retain the term $dJ_i[\nabla_{f_{i,q}} f_{i,p}]|_{x^\star}$, which reflects the intra-block geometry of the frame. Conversely, if the intra-block connection coefficients vanish, i.e., $\nabla_{f_{i,q}} f_{i,p} \equiv 0$ for all i, then the diagonal blocks are pure Hessian blocks, while the off-diagonal blocks retain the term $dJ_i[\nabla_{f_{j,q}} f_{i,p}]|_{x^\star}$ for $i \neq j$, which reflects the inter-block geometry of the frame. Finally, if the frame is holonomic, i.e., $[f_{i,p}, f_{j,q}] \equiv 0$ for all pairs, then the frame is a coordinate frame in some chart and the metric constructed in Theorem 1 is flat on that chart. In particular, all connection terms vanish identically, and, in this case, the blocks (20) reduce to matrices of mixed partial derivatives in this chart, i.e.,

$$\left(A_{ij}|_{x^\star}\right)_{p,q} = \partial^2_{(j,q),(i,p)} J_i\big|_{x^\star},$$

which recovers the characterization in [6]. In general, however, the terms induced by the non-holonomy of the frame $\{f_{i,j}\}$ must be taken into account in studying (5).

□

Before we move on from the current section, we provide examples that illustrate the concepts introduced thus far.

Example 3 (Example 1 cont'd): With the vector fields $\{e_i\}$ defined in Example 1, we define the co-vector fields $\{\vartheta^i\}$ as

$$\vartheta^i(x)[v] := (-1)^{3-i}\Delta(x)^{-1}\,\langle v, e_{3-i}(x) \times x\rangle_{\mathbb{R}^3},$$

for all $(x, v) \in T\mathcal{U}$. Then, it can be verified through direct computation that $\vartheta^i[e_j] = \delta_{ij}$, i.e., $\{\vartheta^i\}$ is the dual co-frame to the frame $\{e_i\}$. We now define a metric g on the submanifold $\mathcal{U}$ by (18) for any pair of vector fields v, w.

Next, we introduce the payoff functions that characterize the decision-making problem:

$$J_i(x) = 2^{\frac{1}{2}} + (x_3 + x_{3-i}),$$

which are depicted in Figure 4. Direct computation gives

$$dJ_i[e_i](x) = -x_i. \tag{21}$$

Therefore, the only equilibrium point in $\mathcal{U}$ is given by $x^\star = (0, 0, 1)$. To verify the structural, as well as the asymptotic, stability properties of $x^\star$, we need to compute the intrinsic Jacobian A defined in (13). To that end, we invoke Theorem 2 and utilize the simplified expression obtained in (20). Substituting from (21) shows that $d(dJ_i[e_i])[e_j]|_{x^\star} = -\delta_{ij}$, i.e., the matrix representation of $A|_{x^\star}$ with respect to the frame $\{e_1, e_2\}$ is given by:

$$A|_{x^\star} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

which is clearly an automorphism of $T_{x^\star}\mathcal{U}$ and also Hurwitz. It follows from Proposition 1 that $x^\star$ is a non-degenerate isolated locally exponentially stable equilibrium. □

Example 4 (Example 2 cont'd): Next, we define the co-vector fields ϑ^i by

$$\vartheta^i(x)[v] := \frac{v_i}{x_1 x_2 x_3},$$

for any tangent vector $v = (v_1, v_2, v_3) \in T_x\mathcal{U}$. Then, it can be verified through direct computation that $\vartheta^i[e_j] = \delta_{ij}$, i.e., the co-frame $\{\vartheta^i\}$ is dual to the frame $\{e_i\}$. We now define a metric g on the submanifold $\mathcal{U}$ by (18), which, automatically, makes the frame $\{e_i\}$ orthonormal with respect to g.

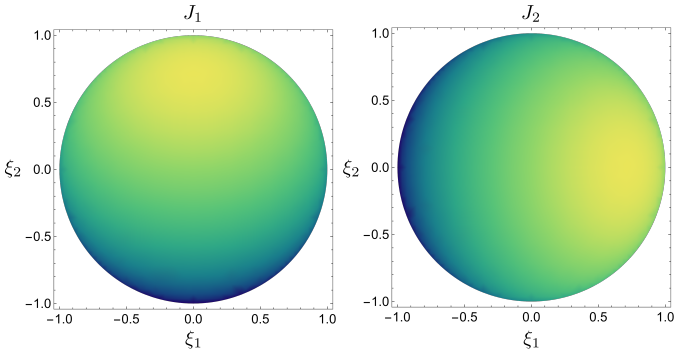


Fig. 4. An illustration of the payoffs $\boldsymbol{J_i}$ for Example 1 in the $\boldsymbol{\xi}$-coordinates. The red dots represent the maximizer of the corresponding function.

The payoff functions of the agents are:

$$J_i(x) := 1 - \frac{1}{2}((x_{3-i} - \alpha)^2 + (x_i - 1 + 2\alpha)^2 + (x_3 - \alpha)^2),$$

for any $\alpha \in \mathbb{R}$, which are depicted in Figure 5 for $\alpha = 4/9$. Direct computation gives that

$$dJ_i[e_i](x) = -x_1 x_2 x_3 (x_i - x_3 + 3\alpha - 1), \tag{22}$$

which can be shown to imply that the only equilibrium in $\mathcal{U}$ is the point

$$x^\star = \frac{1}{3}(2 - 3\alpha, 2 - 3\alpha, 6\alpha - 1),$$

provided that $\alpha \in (1/6, 2/3)$. To verify the structural, as well as the asymptotic, stability properties of $x^\star$, we compute the intrinsic Jacobian A by invoking Theorem 2 and using (20). Indeed, substituting from (22) into (20), we obtain:

$$\begin{aligned}
d(dJ_1[e_1])[e_1] &= x_1 x_2^2 x_3 \big(x_1\,(3\alpha - 4x_3 - 1) \\
&\quad + x_3\,(-3\alpha + x_3 + 1) + x_1^2\big), \\
d(dJ_1[e_1])[e_2] &= x_1^2 x_2 x_3 \big(x_2\,(3\alpha + x_1 - 2x_3 - 1) \\
&\quad + x_3\,(-3\alpha - x_1 + x_3 + 1)\big), \\
d(dJ_1[e_2])[e_1] &= x_1 x_2^2 x_3 \big(x_1\,(3\alpha + x_2 - 2x_3 - 1) \\
&\quad + x_3\,(-3\alpha - x_2 + x_3 + 1)\big), \\
d(dJ_1[e_2])[e_2] &= x_1^2 x_2 x_3 \big(x_2\,(3\alpha - 4x_3 - 1) \\
&\quad + x_3\,(-3\alpha + x_3 + 1) + x_2^2\big),
\end{aligned}$$

so that, at the point $x^\star$, we have that the matrix representation of the linear map $A|_{x^\star}$ is

$$A|_{x^\star} = \frac{1}{729}(1 - 6\alpha)^2(2 - 3\alpha)^4 \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix},$$

which is clearly an automorphism of $T_{x^\star}\mathcal{U}$ and Hurwitz. It follows from Proposition 1 that $x^\star$ is a non-degenerate isolated locally exponentially stable equilibrium. □

To end this section, we consider additional assumptions on the family of vector fields and investigate the resulting simplifications on the geometric structures introduced earlier. We begin with the case in which each agent's control distribution is involutive.

Assumption 3: Suppose that, for each $i \in \{1,\dots,N\}$, $[f_{i,p}, f_{i,q}] \in \operatorname{span}\{f_{i,r}\}_{r=1}^{d_i}$, for all $p,q \in \{1,\dots,d_i\}$. □

In words, Assumption 3 asks that each agent cannot directly interfere in the directions influenced by other agents by exciting Lie Brackets of its own vector fields, which is a natural Assumption. Note that Assumption 3 holds trivially in Examples 1 and 2, since each distribution therein is of rank one. Under Assumptions 1 and 3, each distribution $\mathcal{E}_i$ defines a submanifold $\mathcal{U}_i$ such that $T_x\mathcal{U}_i = \mathcal{E}_i(x)$. Note, however, that this is not enough to conclude that $\mathcal{U}$ admits a product structure $\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_N$. Indeed, as in Examples 1 and 2, each distribution $\mathcal{E}_i$ can be 1-dimensional, which is trivially involutive, and yet the manifold $\mathcal{U}$ does not admit a product structure that respects the splitting in Assumption 1 due to the fact that vector fields belonging to different distributions may not commute. Nevertheless, Assumption 3 allows us to simplify the expressions (20) as follows.

First, by Frobenius Theorem [17, Theorem 19.12], there exists an embedded submanifold $\mathcal{U}_i \subset \mathcal{U}$ such that $T\mathcal{U}_i = \mathcal{E}_i$, for all i. That is, $\mathcal{U}_i \subset \mathcal{U}$ is a *weakly embedded* submanifold [17, Theorem 19.17] of the Riemannian manifold $\mathcal{U}$ and, when equipped with the ambient metric, is isometrically weakly embedded. Consequently, restrictions of smooth functions to $\mathcal{U}_i$ are also smooth functions. Introducing the notation $J_i^{\mathcal{U}_i}$ to denote the restriction of the smooth function J_i to the submanifold $\mathcal{U}_i$, as well as g^i, ∇^i, to denote the Riemannian metric on $\mathcal{U}_i$ induced by the ambient metric g, and its associated Levi-Civita connection, respectively, we have the following Proposition.

Proposition 3: Suppose that Assumptions 1-3 hold, and that $x^\star \in \mathcal{U}$ satisfies (12). Then, with respect to the frame $\{f_{i,j}\}$, the diagonal blocks in (20) are given by

$$(A_{ii}|_{x^\star})_{p,q} = \mathbf{Hess}^i(J_i^{\mathcal{U}_i})[f_{i,p}, f_{i,q}]|_{x^\star}, \tag{23}$$

where $\mathbf{Hess}^i$ is the intrinsic Hessian of $J_i^{\mathcal{U}_i}$, i.e. the restriction of the function J_i to the submanifold $\mathcal{U}_i$, computed with respect to the induced metric g^i. □

Remark 2: It is worth mentioning that a major simplification provided by Proposition 3 is the fact that the matrix $A_{ii}|_{x^\star}$ is now symmetric, since the Hessian is a symmetric operator. We also remark that Assumption 3 will no longer hold, in general, if deception is involved. Indeed, as shown in Proposition 5, the effect of deception in the presence of information asymmetry is to tilt the distributions $\mathcal{E}_i$ controlled by the oblivious agents along the distributions controlled by the deceiver. Therefore, as soon as deception is involved, the new distributions $\tilde{\mathcal{E}}_i$ may no longer be involutive in the general case. This is a major motivation behind considering the general analysis in Section IV, which allows us to study the stability of deception in Section V. □

Next, we consider the following Assumption.

Assumption 4: For all $i,j \in \{1,\dots,N\}$ with $i \neq j$,

$$[f_{i,p}, f_{j,q}] = 0,$$

for all $p \in \{1,\dots,d_i\}$ and all $q \in \{1,\dots,d_j\}$.

In words, Assumption 4 requires that the vector fields controlled by each agent commute with those controlled by every other agent. Note that Assumption 4 fails in both Examples 1 and 2, in view of the non-vanishing commutators computed therein, which illustrates that the product-structure setting considered below is genuinely more restrictive. Under Assumptions 1-4, Frobenius Theorem can be applied to show the existence of a local diffeomorphism $\phi : U \to \mathcal{U}_1 \times \cdots \times \mathcal{U}_N$, where each $\mathcal{U}_i$ is a smooth d_i-dimensional Riemannian manifold and $\phi_* \mathcal{E}_i = T\mathcal{U}_i$. In particular, the Riemannian metric g from Theorem 1 takes a block-diagonal form $g = g^1 \oplus \cdots \oplus g^N$ in these coordinates.

The following proposition establishes the key geometric consequence of the combined assumptions: the Levi-Civita connection of g has no inter-agent components.

Proposition 4: Let Assumptions 1-4 be satisfied. Then,

$$\nabla_{f_{j,q}} f_{i,p} = 0 \quad \text{on } U, \quad \text{for all } i \neq j. \tag{24}$$

□

As a consequence of Proposition 4, the intrinsic Jacobian from Theorem 2 simplifies completely.

Corollary 1: Let the hypotheses of Proposition 4 hold and let $x^\star \in \mathcal{U}$ satisfy (12). Then, the blocks of the intrinsic Jacobian $A|_{x^\star}$ are given by

$$\left(A_{ij}\big|_{x^\star}\right)_{p,q} = \begin{cases} \mathbf{Hess}^i(J_i^{\mathcal{U}_i})[f_{i,p},\, f_{i,q}]\,\big|_{x^\star}, & i = j, \\ \mathbf{Hess}(J_i)[f_{i,p},\, f_{j,q}]\,\big|_{x^\star}, & i \neq j. \end{cases} \tag{25}$$

□

Proof: For diagonal blocks, (25) follows from Proposition 3, which requires only Assumptions 1-3. For off-diagonal blocks with $i \neq j$, substituting (24) into the second equality of (20) immediately gives

$$(A_{ij}|_{x^\star})_{p,q} = \mathbf{Hess}(J_i)[f_{i,p}, f_{j,q}]|_{x^\star}.$$

■

Remark 3: In the special case where each $\mathcal{E}_i$ is spanned by coordinate vector fields in some chart, i.e., $f_{i,j} = \partial_{(i,j)}$, Assumptions 2-4 are automatically satisfied, and system (5) takes the form

$$\dot{x} = \sum_{i=1}^{N} \sum_{j=1}^{d_i} \partial_{(i,j)} J_i(x) \partial_{(i,j)},$$

in coordinates, i.e., the interaction is governed by the vector field whose components are partial derivatives of the payoff functions in local coordinates. Moreover, we compute that

$$\mathbf{Hess}^i(J_i)[f_{i,p}, f_{i,q}] = \partial^2_{(i,p),(i,q)} J_i,$$
$$\mathbf{Hess}(J_i)[f_{i,p}, f_{j,q}] = \partial^2_{(j,q),(i,p)} J_i.$$

In this case, (25) recovers the coordinate dependent formulas in [6], [7] as a special case of Theorem 2. Note, however, that even if Assumptions 2-4 hold, the results in [6], [7] still do not apply due to the fact that, for each i, the commutators $[f_{i,p}, f_{i,q}]$ need not vanish in general. In particular, the expression (25) accounts for the intrinsic curvature of the submanifolds $\mathcal{U}_i$ which, in general, does not vanish. □

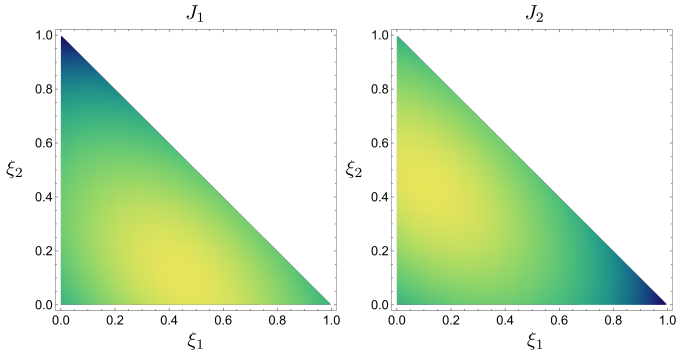


Fig. 5. An illustration of payoffs $\boldsymbol{J_i}$ for Example 2 in the $\boldsymbol{\xi}$-coordinates. The red dots represent the maximizer of the corresponding function.

V. Information Asymmetry and Deception

As discussed in Section III, when the assumption of information symmetry is violated, an agent with access to privileged information can bias the steady-state behavior of (1) by manipulating only its own control inputs, without interfering with the actions or measurements of the oblivious agents. In this section, we analyze the interaction dynamics induced by the deceptive feedback law (6), i.e., system (8), repeated below for convenience,

$$\dot{x} = \sum_{i=1}^{2} \kappa_i \sum_{j=1}^{d_i} L_{\tilde{f}_{i,j}}(J_i)(x)\tilde{f}_{i,j}(x), \tag{8}$$

where, as in (7b), the vector fields $\tilde{f}_{i,j}$ are given by

$$\tilde{f}_{1,j} = f_{1,j}, \quad \tilde{f}_{2,j} := f_{2,j} + \sum_{k=1}^{d_1} \gamma_j^k f_{1,k}.$$

Similar to Section IV, we impose Assumption 1 on the vector fields $f_{i,j}$ and, for simplicity, we also impose Assumption 2. Then, we have the following proposition.

Proposition 5: Let $\{\tilde{f}_{i,j}\}$ be given by (7b) for any choice of deception parameters $\{\gamma_k^j\}$. Then, Assumption 1 is satisfied on $\mathcal{U}$ with $\tilde{\mathcal{E}}_i = \operatorname{span}\{\tilde{f}_{i,j}\}_{j=1}^{d_i}$. In particular, the family $\{\tilde{f}_{i,j}\}$ remains pointwise linearly independent on $\mathcal{U}$ and

$$T_x\mathcal{U} = \tilde{\mathcal{E}}_1(x) \oplus \tilde{\mathcal{E}}_2(x) \qquad \text{for all } x \in \mathcal{U}.$$

□

Remark 4: Although $\tilde{\mathcal{E}}_1 = \mathcal{E}_1$ as subbundles of $T\mathcal{U}$, since the deceiver does not alter their own directions, the victim's new bundle $\tilde{\mathcal{E}}_2$ is generically distinct from $\mathcal{E}_2$ since, from (7b), the vector field $\tilde{f}_{2,j}$ acquires components along the deceiver's directions $f_{1,k}$ with weights γ_j^k. This is the geometric mechanism of deception in our setting: the deceiver *rotates the victim's effective actuation directions* into its own bundle without modifying its own reach. As a consequence, projected gradients on $\tilde{\mathcal{E}}_2$ are biased relative to those on $\mathcal{E}_2$, and the equilibrium of the averaged dynamics is generically shifted. Moreover, the proposition holds for *all* choices of deception parameters, including arbitrarily large ones, and uniformly in $x \in \mathcal{U}$. This is a direct consequence of the block-triangular structure of Γ in (50), which decouples the invertibility analysis from the magnitude of the tilt block γ. □

As a direct consequence of Theorem 1 and Proposition 5, we have the following Corollary.

Corollary 2: There exists a Riemannian metric $\tilde{g}$ on the open sub-manifold $\mathcal{U} \subset \mathcal{M}$ such that (8) takes the form

$$\dot{x} = \sum_{i=1}^{2} \kappa_i \, \mathbf{grad}_{\tilde{\mathcal{E}}_i}(J_i)(x), \tag{26}$$

where $\mathbf{grad}_{\tilde{\mathcal{E}}_i}(J_i)$ is the vector field obtained through orthogonal projection of $\mathbf{grad}(J_i)$ with respect to $\tilde{g}$ onto $\tilde{\mathcal{E}}_i$. □

This result allows us to, once again, take $\kappa_i = 1$ and look for equilibria of system (8) using the condition

$$\mathbf{grad}_{\tilde{\mathcal{E}}_i}(J_i)(x^\star) = 0, \quad i \in \{1, 2\}. \tag{27}$$

Notably, since $\tilde{\mathcal{E}}_1 = \mathcal{E}_1$, this condition is unchanged from the analysis in Section IV for $i = 1$. However, unlike Section IV, the bundle $\tilde{\mathcal{E}}_2$ is tilted along the bundle $\mathcal{E}_1$. In particular, utilizing the assumption that the vector fields $\{f_{i,j}\}$, and, consequently the vector fields $\{\tilde{f}_{i,j}\}$, are linearly independent, as well as (7b), the new equilibrium conditions take the form

$$0 = dJ_1[f_{1,j}](x), \qquad j \in \{1, \dots, d_1\},$$
$$0 = dJ_2[f_{2,j}](x) + \sum_{k=1}^{d_1} \gamma_j^k dJ_2[f_{1,k}](x), \quad j \in \{1, \dots, d_2\},$$

which implies that an equilibrium point of system (8) will, in general, depend on the matrix of deception parameters $\gamma \in \mathbb{R}^{d_1 \times d_2}$. Nevertheless, we have the following Theorem.

Theorem 3: Suppose that $x^\star \in \mathcal{U}$ is a locally exponentially stable equilibrium point for the interaction dynamics (5). Then, there exists an open neighborhood $\Omega \subset \mathbb{R}^{d_1 \times d_2}$ with $0 \in \Omega$ and a smooth map $X^\star : \Omega \to \mathcal{U}$ such that $X^\star(0) = x^\star$ and, for all $\gamma \in \Omega$, the point $X^\star(\gamma)$ is a locally exponentially stable equilibrium point for the interaction dynamics (8). □

Corollary 3: Under the hypotheses of Theorem 3, and under the identification of $T_{x^\star}\mathcal{U}$ with $\mathbb{R}^n$ provided by the frame $\{f_{i,j}(x^\star)\}$, the map $X^\star$ satisfies

$$\left.\frac{\partial X^\star}{\partial \gamma_j^k}\right|_{\gamma=0} = -\, dJ_2[f_{1,k}](x^\star) \left(A|_{x^\star}\right)^{-1} \left[f_{2,j}(x^\star)\right], \tag{28}$$

for all $j \in \{1, \dots, d_2\}$ and $k \in \{1, \dots, d_1\}$. □

Remark 5: Formula (28) makes the mechanism of Proposition 5 quantitative: to first order, the deceiver shifts the equilibrium along the directions $(A|_{x^\star})^{-1}[f_{2,j}(x^\star)]$, weighted by the sensitivities $dJ_2[f_{1,k}](x^\star)$ of the victim's payoff along the deceiver's directions. In particular, if $dJ_2[f_{1,k}](x^\star) = 0$ for all k, i.e., if the victim's payoff is stationary along the deceiver's directions at $x^\star$, then deception is ineffective to first order, regardless of the choice of γ. Formula (28) thus provides a natural starting point for the constructive design of deception parameters, a direction we pursue in future work. □

Remark 6: The restriction to two agents is made for clarity of exposition, not out of necessity. Consider N agents and

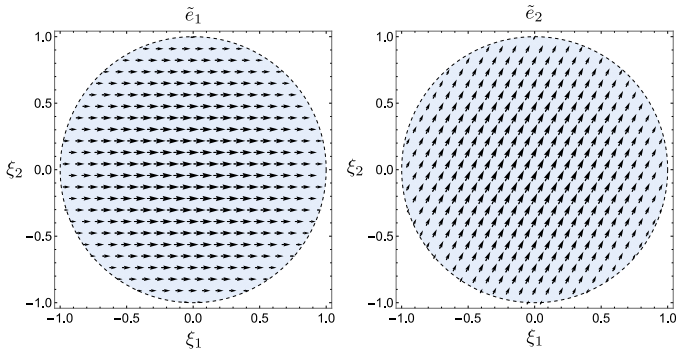


Fig. 6. An illustration of the tilted vector fields $\tilde{\boldsymbol{e}}_{\boldsymbol{i}}$ for Example 5 in the $\boldsymbol{\xi}$-coordinates. The arrows' sizes represent the magnitude of the vector field.

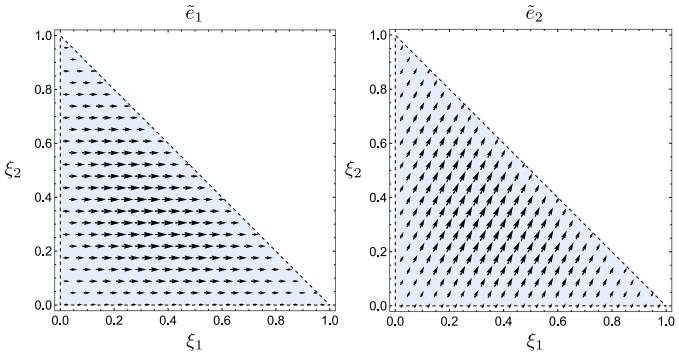


Fig. 7. An illustration of the tilted vector fields $\tilde{\boldsymbol{e}}_{\boldsymbol{i}}$ for Example 6 in the $\boldsymbol{\xi}$-coordinates. The arrows' sizes represent the magnitude of the vector field.

suppose that the information-asymmetry structure is *acyclic*, i.e., there exists an ordering of the agents such that agent i can measure, and embed into its own inputs, the inputs of agent ℓ only if $i < \ell$. Then, the matrix Γ representing the tilt, cf. (50), is strictly block upper-triangular in this ordering, hence nilpotent, and $\det(I + \Gamma) = 1$ irrespective of the deception parameters. Consequently, Proposition 5, Theorem 3, and Corollary 3 extend verbatim to this setting. The genuinely distinct case is that of a *cyclic* information structure, e.g., two agents simultaneously attempting to deceive one another, in which case $I + \Gamma$ may become singular for certain parameter choices and the tilted family may fail to span the tangent space; the characterization of equilibria in this regime is an open problem that we leave for future work. □

We conclude this section by revisiting the examples from Section IV when deception is involved.

Example 5 (Example 1 cont'd): We consider the case in which the 1^{st}-agent implements a deceptive strategy towards the 2^{nd}-agent. In this case, $\{e_i\}$ are replaced by

$$\tilde{e}_1(x) := e_1(x), \qquad \tilde{e}_2(x) = e_2(x) + \gamma e_1(x),$$

where $\gamma \in \mathbb{R}$ is the deception parameter. The new vector fields are depicted in Figure 6 for $\gamma = \frac{1}{2}$. As can be observed from the figure, the vector field $\tilde{e}_2$ is tilted compared to Figure 2. Repeating the same computations as in Example 1, we find that the new frame $\{\tilde{e}_i\}$, with $\tilde{e}_1 = e_1$ and $\tilde{e}_2$ as above, still spans the tangent space of the submanifold $\mathcal{U}$. In particular, we have that

$$dJ_1[e_1](x) = -x_1, \qquad dJ_2[\tilde{e}_2] = -x_2 - \gamma(x_1 - x_3),$$

so that the only equilibrium in $\mathcal{U}$ is given, as a function of the deception parameter γ, by

$$X^\star(\gamma) = \left(0, \frac{\gamma}{\sqrt{1+\gamma^2}}, \frac{1}{\sqrt{1+\gamma^2}}\right). \tag{29}$$

Next, we proceed to verify the structural, as well as the asymptotic, stability properties of the point $X^\star(\gamma)$. Thanks to Theorem 3, we already know that there exists an open neighborhood $\Omega \subset \mathbb{R}$ with $0 \in \Omega$ such that, for all $\gamma \in \Omega$, $X^\star(\gamma)$ is locally exponentially stable. Nevertheless, because we have an explicit expression of the equilibrium point, we can obtain stronger results. To do so, we employ the frame $\{\tilde{e}_1, \tilde{e}_2\}$, which is orthonormal with respect to the tilted metric $\tilde{g}$ of Corollary 2. In particular, we compute that

$$\begin{aligned}
d(dJ_1[\tilde{e}_1])[\tilde{e}_1] &= -x_3,\\
d(dJ_1[\tilde{e}_1])[\tilde{e}_2] &= -\gamma x_3,\\
d(dJ_2[\tilde{e}_2])[\tilde{e}_1] &= -\gamma(x_1 + x_3),\\
d(dJ_2[\tilde{e}_2])[\tilde{e}_2] &= -\gamma(\gamma x_1 + x_2) - (1+\gamma^2)x_3.
\end{aligned}$$

Hence, substituting from (29), we see that the matrix representation of $A|_{X^\star(\gamma)}$ with respect to $\{\tilde{e}_i\}$ is

$$A|_{X^\star(\gamma)} = \frac{1}{\sqrt{1+\gamma^2}} \begin{bmatrix} -1 & -\gamma \\ -\gamma & -(1+2\gamma^2) \end{bmatrix}.$$

To verify that $A|_{X^\star(\gamma)}$ is an automorphism of $T_{X^\star(\gamma)}\mathcal{U}$ and Hurwitz, we compute its trace and determinant as:

$$\begin{aligned}
\operatorname{trace}(A|_{X^\star(\gamma)}) &= -2\sqrt{1+\gamma^2}\\
\det(A|_{X^\star(\gamma)}) &= 1,
\end{aligned}$$

which, since it is a 2×2 matrix, imply that its eigenvalues are strictly negative, and, therefore, $X^\star(\gamma)$ is a non-degenerate isolated exponentially stable equilibrium for all $\gamma \in \mathbb{R}$. In conclusion, deception preserves the stability of the new equilibrium point (29), which is controlled by the deceiver via the parameter γ. We also note that differentiating (29) at $\gamma = 0$ gives $\frac{d}{d\gamma} X^\star(0) = (0, 1, 0) = e_2(x^\star)$, in agreement with Corollary 3, since $dJ_2[e_1](x^\star) = 1$ and $A|_{x^\star} = -I$.

□

Example 6 (Example 2 cont'd): We consider the case in which the 1^{st}-agent implements a deceptive strategy towards the 2^{nd}-agent in the setting of Example 2. In this case, the vector fields $\{e_i\}$ are replaced by

$$\tilde{e}_1(x) := e_1(x), \qquad \tilde{e}_2(x) := e_2(x) + \gamma\, e_1(x),$$

where $\gamma \in \mathbb{R}$ is the deception parameter. The new vector fields are depicted in Figure 7 for $\gamma = \frac{1}{2}$. As can be observed from the figure, the vector field $\tilde{e}_2$ is tilted compared to Figure 3. We compute that

$$dJ_2[e_1](x) = x_1 x_2 x_3\,(x_3 - x_1),$$

and, consequently,

$$dJ_1[\tilde{e}_1](x) = x_1 x_2 x_3\,(-x_1 + x_3 + 1 - 3\alpha),$$

$$dJ_2[\tilde{e}_2](x) = x_1x_2x_3\big[(-x_2+x_3+1-3\alpha)+\gamma(x_3-x_1)\big],$$

where we utilized (22). On $\mathcal{U}$ we have $x_1x_2x_3 > 0$, so the equilibrium conditions reduce to the two linear equations

$$\begin{aligned} -x_1+x_3+1-3\alpha &= 0,\\ (-x_2+x_3+1-3\alpha)+\gamma(x_3-x_1) &= 0,\end{aligned}$$

together with the simplex constraint $x_1+x_2+x_3=1$. Introducing $\beta := 1-3\alpha$, the first equation gives $x_1 = x_3+\beta$, so that $x_3-x_1 = -\beta$. Substituting into the second and simplifying yields $x_2 = x_3+\beta(1-\gamma)$. Therefore, the simplex constraint determines

$$x_3 = \frac{(6\alpha-1)+\gamma(1-3\alpha)}{3}.$$

From the above, we see that the unique equilibrium in $\mathcal{U}$, when it exists as a function of the deception parameter γ, is

$$X^\star(\gamma) = \frac{1}{3}\begin{pmatrix}(2-3\alpha)+\gamma(1-3\alpha)\\ (2-3\alpha)-2\gamma(1-3\alpha)\\ (6\alpha-1)+\gamma(1-3\alpha)\end{pmatrix}. \tag{30}$$

In particular, $X^\star(0)$ recovers the nominal equilibrium from Example 2, and the γ-dependent shift

$$X^\star(\gamma)-X^\star(0) = \frac{\gamma(1-3\alpha)}{3}\,(1,\,-2,\,1)$$

is exactly linear in γ. For $X^\star(\gamma)$ to remain in the interior $\mathcal{U}$, we require $\gamma\in\Omega$, where $\Omega\subset\mathbb{R}$ is the open interval

$$\Omega = \{\gamma\in\mathbb{R} \;:\; X_i^\star(\gamma)>0 \text{ for } i=1,2,3\}. \tag{31}$$

The three positivity constraints in (31) read

$$\begin{aligned}(2-3\alpha)+\gamma(1-3\alpha)&>0,\\ (2-3\alpha)-2\gamma(1-3\alpha)&>0,\\ (6\alpha-1)+\gamma(1-3\alpha)&>0.\end{aligned}$$

For $\alpha\in(1/6,1/3)$, these give that

$$\Omega = \left(\frac{1-6\alpha}{1-3\alpha},\,\frac{2-3\alpha}{2(1-3\alpha)}\right),$$

which contains 0 and is bounded on both sides. For $\alpha=1/3$, the map $\gamma\mapsto X^\star(\gamma)$ is constant ($\beta=0$), and deception has no effect on the equilibrium. For $\alpha\in(1/3,2/3)$ (so $1-3\alpha<0$), the inequalities flip direction and

$$\Omega = \left(\frac{2-3\alpha}{2(1-3\alpha)},\,\frac{1-6\alpha}{1-3\alpha}\right).$$

In both non-degenerate regimes, Ω is an explicit, bounded, open interval containing 0, giving a quantitative counterpart to the qualitative neighborhood guaranteed by Theorem 3.

To verify local exponential stability of $X^\star(\gamma)$, we compute the intrinsic Jacobian $A|_{X^\star(\gamma)}$ with respect to the frame $\{\tilde{e}_1,\tilde{e}_2\}$, which is orthonormal with respect to the tilted metric $\tilde{g}$, via Theorem 2.

Observe that each directional derivative above factors as $\chi\cdot\rho$, where $\chi := x_1x_2x_3$ and ρ is a polynomial of degree 1 in x. At $X^\star(\gamma)$, the linear factor ρ vanishes by construction of the equilibrium, while $\chi|_{X^\star(\gamma)}>0$. Hence, using $\tilde{e}_1=\chi(\partial_1-\partial_3)$ and $\tilde{e}_2=\chi(\gamma\partial_1+\partial_2-(1+\gamma)\partial_3)$,

$$d(dJ_i[\tilde{e}_i])[\tilde{e}_j]\big|_{X^\star(\gamma)} = \chi\cdot\tilde{e}_j[\rho_i]\big|_{X^\star(\gamma)},$$

where $\rho_1 = -x_1+x_3+\beta$ and $\rho_2 = (-x_2+x_3+\beta)+\gamma(x_3-x_1)$. Direct computation gives

$$\begin{aligned} d(dJ_1[\tilde{e}_1])[\tilde{e}_1] &= -2\,\chi^2,\\ d(dJ_1[\tilde{e}_1])[\tilde{e}_2] &= -(1+2\gamma)\,\chi^2,\\ d(dJ_2[\tilde{e}_2])[\tilde{e}_1] &= -(1+2\gamma)\,\chi^2,\\ d(dJ_2[\tilde{e}_2])[\tilde{e}_2] &= -2(1+\gamma+\gamma^2)\,\chi^2.\end{aligned}$$

Hence, substituting from (30), the matrix representation of $A|_{X^\star(\gamma)}$ with respect to $\{\tilde{e}_i\}$ is

$$A|_{X^\star(\gamma)} = \chi^2\big|_{X^\star(\gamma)}\begin{bmatrix}-2 & -(1+2\gamma)\\ -(1+2\gamma) & -2(1+\gamma+\gamma^2)\end{bmatrix}. \tag{32}$$

In particular, at $\gamma=0$, the scalar $\chi^2\big|_{x^\star} = \frac{1}{729}(2-3\alpha)^4(6\alpha-1)^2$ and the matrix (32) reduces to the expression in Example 4. Next, we verify that $A|_{X^\star(\gamma)}$ is Hurwitz by computing its trace and determinant:

$$\begin{aligned}\operatorname{tr}\big(A|_{X^\star(\gamma)}\big) &= -2\,\chi^2\big|_{X^\star(\gamma)}\cdot(2+\gamma+\gamma^2),\\ \det\big(A|_{X^\star(\gamma)}\big) &= \chi^4\big|_{X^\star(\gamma)}\cdot\Big[4(1+\gamma+\gamma^2)-(1+2\gamma)^2\Big]\\ &= 3\,\chi^4\big|_{X^\star(\gamma)}.\end{aligned}$$

Since $2+\gamma+\gamma^2 = (\gamma+1/2)^2+7/4>0$ for all $\gamma\in\mathbb{R}$, the trace is strictly negative for every $\gamma\in\Omega$. The determinant equals $3\,\chi^4|_{X^\star(\gamma)}>0$ for every $\gamma\in\Omega$. A real 2×2 matrix with strictly negative trace and strictly positive determinant is Hurwitz, and therefore $X^\star(\gamma)$ is a non-degenerate isolated locally exponentially stable equilibrium for system (8), for every $\gamma\in\Omega$. In conclusion, as in Example 5, deception preserves the stability of the system on the entire admissible parameter interval Ω, with the equilibrium smoothly controlled by γ according to (30) with $\gamma\in\Omega$. □

VI. Proofs

In this section, we collect all proofs of the main results presented in the manuscript.

A. Proof of Theorem 1

Under Assumption 1, and for each $x\in\mathcal{U}$, we have the (pointwise) direct-sum decomposition

$$T_x\mathcal{U} = \mathcal{E}_1(x)\oplus\cdots\oplus\mathcal{E}_N(x),$$

where, thanks to Assumption 1, each $\mathcal{E}_i$ is a regular distribution on $\mathcal{U}$ generated by the family $\{f_{i,j}\}_{j=1}^{d_i}$. Consequently, using annihilators yields the canonical dual decomposition

$$T_x^\star\mathcal{U} = \mathcal{E}_1(x)^\circ\oplus\cdots\oplus\mathcal{E}_N(x)^\circ, \tag{33}$$

$$\mathcal{E}_i(x)^\circ := \operatorname{Ann}\Big(\bigoplus_{k\neq i}\mathcal{E}_k(x)\Big)\subset T_x^\star\mathcal{U}. \tag{34}$$

Let $\pi_i(x) : T_x^\star\mathcal{U} \to \mathcal{E}_i(x)^\circ$ denote the corresponding linear projections. For each $i \in \{1, \dots, N\}$, define a linear operator

$$P_i(x)[\omega] := \sum_{j=1}^{d_i} \big(\pi_i(x)\omega\big)\big(f_{i,j}(x)\big)\, f_{i,j}(x), \tag{35}$$

for all $\omega \in T_x^\star\mathcal{U}$. By construction, $P_i(x)[\omega] \in \mathcal{E}_i(x)$ for all $\omega \in T_x^\star\mathcal{U}$. Moreover, since $dJ_i(x)\big(f_{i,j}(x)\big) = L_{f_{i,j}}(J_i)(x)$, system (5) can be rewritten as

$$\dot{x} = \sum_{i=1}^{N} \kappa_i P_i(x)\big[dJ_i(x)\big]. \tag{36}$$

We now construct a metric g on $\mathcal{U}$ via its co-metric. For each $x \in \mathcal{U}$, we define a symmetric bilinear form g_x^{-1} on $T_x^\star\mathcal{U}$ by declaring it block-diagonal with respect to the splitting (33) and, for $\alpha, \beta \in T_x^\star\mathcal{U}$ with unique decompositions $\alpha = \sum_{i=1}^N \alpha_i$ and $\beta = \sum_{i=1}^N \beta_i$ such that $\alpha_i, \beta_i \in \mathcal{E}_i(x)^\circ$, defining

$$\begin{aligned} g_x^{-1}(\alpha, \beta) &:= \sum_{i=1}^{N} g_{i,x}^{-1}(\alpha_i, \beta_i), \\ g_{i,x}^{-1}(\alpha_i, \beta_i) &:= \sum_{j=1}^{d_i} \alpha_i\big(f_{i,j}(x)\big)\, \beta_i\big(f_{i,j}(x)\big). \end{aligned} \tag{37}$$

Since $\{f_{i,j}(x)\}_{j=1}^{d_i}$ spans $\mathcal{E}_i(x)$, the quadratic form $\alpha_i \mapsto g_{i,x}^{-1}(\alpha_i, \alpha_i) = \sum_{j=1}^{d_i} \alpha_i(f_{i,j}(x))^2$ is positive definite on $\mathcal{E}_i(x)^\circ$. Hence, g_x^{-1} is positive definite on $T_x^\star\mathcal{U}$ for every $x \in \mathcal{U}$. Smoothness follows from the smoothness of the vector fields $f_{i,j}$ and of the projections $\pi_i(x)$ for regular distributions. Therefore, the tensor g^{-1} defines a Riemannian co-metric on $\mathcal{U}$, and, consequently, a unique Riemannian metric g.

Next, let $g_x^\sharp : T_x^\star\mathcal{U} \to T_x\mathcal{U}$ denote the tangent-cotangent isomorphism induced by g, i.e.,

$$\alpha(v) = g_x\big(g_x^\sharp\alpha, v\big) \qquad \forall\, \alpha \in T_x^\star\mathcal{U},\ v \in T_x\mathcal{U}.$$

By construction of (37), we have the block formula

$$g_x^\sharp\alpha = \sum_{i=1}^{N} P_i(x)[\alpha_i], \tag{38}$$

where $\alpha = \sum_{i=1}^N \alpha_i$, $\alpha_i \in \mathcal{E}_i(x)^\circ$ is the unique decomposition of α with respect to the splitting (33). For each i, the Riemannian gradient of the function J_i with respect to the metric g is defined by $\mathbf{grad}(J_i)(x) := g_x^\sharp\,(dJ_i(x))$. Utilizing the splitting (33) once more, we have the unique decomposition $dJ_i(x) = \sum_{k=1}^N (dJ_i(x))_k$ with $(dJ_i(x))_k \in \mathcal{E}_k(x)^\circ$, which, in combination with (38), implies that

$$\mathbf{grad}(J_i)(x) = \sum_{k=1}^{N} P_k(x)\big[(dJ_i(x))_k\big].$$

Moreover, since g_x^{-1} is block-diagonal with respect to the splitting (33), it follows that the decomposition (10) is g_x-orthogonal, and, therefore, the g_x-orthogonal projection $P_{\mathcal{E}_i}(x) : T_x\mathcal{U} \to \mathcal{E}_i(x)$ is precisely the extraction of the i-th component. Hence

$$\mathbf{grad}_{\mathcal{E}_i}(J_i)(x) := P_{\mathcal{E}_i}[\mathbf{grad}(J_i)](x) = P_i(x)\big[(dJ_i(x))_i\big].$$

Finally, since $f_{i,j}(x) \in \mathcal{E}_i(x)$ we have that

$$(dJ_i(x))_i(f_{i,j}(x)) = dJ_i(x)(f_{i,j}(x)),$$

and, therefore, it follows that

$$\mathbf{grad}_{\mathcal{E}_i}(J_i)(x) = \sum_{j=1}^{d_i} L_{f_{i,j}}(J_i)(x)\, f_{i,j}(x).$$

Combining this identity with (36) proves that system (5) is of the form (11), which concludes the proof.

B. Proof of Proposition 2

Let $X := \mathrm{diag}(\chi_1 I_1, \dots, \chi_N I_N)$, where I_i is the identity on $\mathcal{E}_i(x^\star)$, and define the operator $\tilde{A} := X^{-1}AX$, which is similar to A, i.e., has the same spectrum. Direct computation shows that its blocks with respect to the splitting (10) are $\tilde{A}_{ii} = A_{ii}$, and $\tilde{A}_{ij} = \frac{\chi_j}{\chi_i}A_{ij}$. Therefore,

$$\tilde{R}_i := \sum_{j\neq i} \|\tilde{A}_{ij}\|_g = \sum_{j\neq i} \|A_{ij}\|_g \frac{\chi_j}{\chi_i},$$

with respect to $\|\cdot\|_g$. By the block Gershgorin circle theorem [21], each $\lambda \in \sigma(\tilde{A})$ satisfies, for some i,

$$\|(\tilde{A}_{ii} - \lambda I_i)^{-1}\|_g^{-1} \leq \tilde{R}_i. \tag{39}$$

Since $\tilde{A}_{ii} = A_{ii}$ and $\|M^{-1}\|_g^{-1} = \sigma_{\min}^g(M)$, we obtain that $\sigma_{\min}^g(A_{ii} - \lambda I_i) \leq \tilde{R}_i$. To arrive at a contradiction, suppose that, for some $\lambda \in \sigma(\tilde{A})$, we have that $\Re(\lambda) \geq 0$. Then, we would have that

$$\sigma_{\min}^g(A_{ii} - \lambda I_i) \geq \inf_{\Re(s)\geq 0} \sigma_{\min}^g(A_{ii} - sI_i) > \tilde{R}_i$$

from (16), which contradicts (39). Hence, $\sigma(\tilde{A}) \subset \{\Re(z) < 0\}$ and, because $\sigma(A) = \sigma(\tilde{A})$, A is Hurwitz.

C. Proof of Theorem 2

Recall from (15) that

$$A_{ij}[\,\cdot\,] = P_{\mathcal{E}_i}\Big[\nabla_{P_{\mathcal{E}_j}[\,\cdot\,]}\,\mathbf{grad}(J_i)\Big] + C_{ij}[\,\cdot\,], \tag{40}$$

$$C_{ij}[\,\cdot\,] = \sum_{k=1}^{N} P_{\mathcal{E}_i}\Big[\big(\nabla_{P_{\mathcal{E}_j}[\,\cdot\,]}P_{\mathcal{E}_k}\big)[\mathbf{grad}(J_k)]\Big]. \tag{41}$$

We evaluate $A_{ij}[v]$ for an arbitrary vector field v and then extract its components in the orthonormal frame $\{f_{i,j}\}$.

From (18) and the product rule, we compute that

$$\nabla P_{\mathcal{E}_k} = \sum_{\ell=1}^{d_k} \Big(\nabla\vartheta^{k,\ell} \otimes f_{k,\ell} + \vartheta^{k,\ell} \otimes \nabla f_{k,\ell}\Big).$$

Since $P_{\mathcal{E}_j}[v] = \sum_{q=1}^{d_j} \vartheta^{j,q}[v]\, f_{j,q}$, we obtain that

$$\begin{aligned} \nabla_{P_{\mathcal{E}_j}[v]}P_{\mathcal{E}_k} = \sum_{q=1}^{d_j} \vartheta^{j,q}[v] \sum_{\ell=1}^{d_k} \Big(&\vartheta^{k,\ell} \otimes \nabla_{f_{j,q}} f_{k,\ell} \\ &+ (\nabla_{f_{j,q}}\vartheta^{k,\ell}) \otimes f_{k,\ell}\Big). \end{aligned} \tag{42}$$

From the definition of the covariant derivative of a 1-form,

$$(\nabla_{f_{j,q}}\vartheta^{k,\ell})[\,\cdot\,] = d\big(\vartheta^{k,\ell}[\,\cdot\,]\big)[f_{j,q}] - \vartheta^{k,\ell}\big[\nabla_{f_{j,q}}[\,\cdot\,]\big].$$

Applying this to $\mathbf{grad}(J_k)$ and utilizing (17) and (19),

$$(\nabla_{f_{j,q}}\vartheta^{k,\ell})[\mathbf{grad}(J_k)] = d\big(dJ_k[f_{k,\ell}]\big)[f_{j,q}] - \vartheta^{k,\ell}\big[\nabla_{f_{j,q}}\mathbf{grad}(J_k)\big].$$

Utilizing (19), the product rule, and (17):

$$\vartheta^{k,\ell}\big[\nabla_{f_{j,q}}\mathbf{grad}(J_k)\big] = d\big(dJ_k[f_{k,\ell}]\big)[f_{j,q}] + \sum_{m,r} dJ_k[f_{m,r}]\,\vartheta^{k,\ell}[\nabla_{f_{j,q}}f_{m,r}].$$

Hence, we obtain that

$$(\nabla_{f_{j,q}}\vartheta^{k,\ell})[\mathbf{grad}(J_k)] = -\sum_{m,r} dJ_k[f_{m,r}]\,\vartheta^{k,\ell}[\nabla_{f_{j,q}}f_{m,r}]. \tag{43}$$

Substituting (42) into (41) and applying $P_{\mathcal{E}_i}$, which extracts components along $\{f_{i,p}\}_{p=1}^{d_i}$ via (18), we obtain:

$$\begin{aligned} &P_{\mathcal{E}_i}\Big[\big(\nabla_{P_{\mathcal{E}_j}[v]}P_{\mathcal{E}_k}\big)[\mathbf{grad}(J_k)]\Big] \\ &\quad= \sum_{q,p}\left(\sum_{\ell} dJ_k[f_{k,\ell}]\,\vartheta^{i,p}[\nabla_{f_{j,q}}f_{k,\ell}]\right)\vartheta^{j,q}[v]f_{i,p} \\ &\qquad+ \sum_{q,\ell}(\nabla_{f_{j,q}}\vartheta^{k,\ell})[\mathbf{grad}(J_k)]\vartheta^{j,q}[v]P_{\mathcal{E}_i}[f_{k,\ell}]. \end{aligned}$$

Using $P_{\mathcal{E}_i}[f_{k,\ell}] = \delta_{ik}\,f_{i,\ell}$ and substituting (43), the second sum contributes only when $k = i$:

$$-\delta_{ik}\sum_{q,\ell}\left(\sum_{m,r} dJ_k[f_{m,r}]\,\vartheta^{k,\ell}[\nabla_{f_{j,q}}f_{m,r}]\right)\vartheta^{j,q}[v]f_{i,\ell}.$$

Summing over k and relabeling $\ell \mapsto p$, we obtain that

$$C_{ij}[v] = \sum_{q,p}\Bigg(\sum_{k,\ell}\big(dJ_k[f_{k,\ell}] - dJ_i[f_{k,\ell}]\big)\,\vartheta^{i,p}[\nabla_{f_{j,q}}f_{k,\ell}]\Bigg)\vartheta^{j,q}[v]f_{i,p}. \tag{44}$$

From (19) and the product rule,

$$\nabla_{f_{j,q}}\mathbf{grad}(J_i) = \sum_{k,\ell}\Big(d\big(dJ_i[f_{k,\ell}]\big)[f_{j,q}]\,f_{k,\ell} + dJ_i[f_{k,\ell}]\,\nabla_{f_{j,q}}f_{k,\ell}\Big).$$

Extracting the (i,p)-component with $\vartheta^{i,p}\circ P_{\mathcal{E}_i}$ and using (17),

$$P_{\mathcal{E}_i}[\nabla_{P_{\mathcal{E}_j}[v]}\mathbf{grad}(J_i)] = \sum_{q,p}\Bigg(d\big(dJ_i[f_{i,p}]\big)[f_{j,q}] + \sum_{k,\ell} dJ_i[f_{k,\ell}]\,\vartheta^{i,p}[\nabla_{f_{j,q}}f_{k,\ell}]\Bigg)\vartheta^{j,q}[v]f_{i,p}. \tag{45}$$

Adding (44) and (45) and canceling identical terms, we get

$$A_{ij}[v] = \sum_{q,p}\Bigg(d\big(dJ_i[f_{i,p}]\big)[f_{j,q}] + \sum_{k,\ell} dJ_k[f_{k,\ell}]\,\vartheta^{i,p}[\nabla_{f_{j,q}}f_{k,\ell}]\Bigg)\vartheta^{j,q}[v]f_{i,p}. \tag{46}$$

By condition (12) and (19), $\mathbf{grad}_{\mathcal{E}_k}(J_k)(x^\star) = 0$ implies

$$dJ_k[f_{k,\ell}]\big|_{x^\star} = 0 \qquad \text{for all } k,\ell.$$

Hence the inner sum in (46) vanishes at $x^\star$, leaving

$$A_{ij}[v]\big|_{x^\star} = \sum_{q,p} d\big(dJ_i[f_{i,p}]\big)[f_{j,q}]\,\vartheta^{j,q}[v]f_{i,p}\big|_{x^\star}.$$

Reading off the matrix entries with respect to the orthonormal frame gives the first equality in (20). Finally, from the defining identity of the Hessian,

$$\mathbf{Hess}(J_i)[X,Y] = d(dJ_i[Y])[X] - dJ_i[\nabla_X Y],$$

applied with $X = f_{j,q}$ and $Y = f_{i,p}$,

$$d\big(dJ_i[f_{i,p}]\big)[f_{j,q}] = \mathbf{Hess}(J_i)[f_{i,p},f_{j,q}] + dJ_i\big[\nabla_{f_{j,q}}f_{i,p}\big],$$

establishing the second equality in (20).

D. Proof of Proposition 3

By the first equality in (20),

$$(A_{ii}|_{x^\star})_{p,q} = d(dJ_i[f_{i,p}])[f_{i,q}]|_{x^\star}.$$

The definition of the Hessian implies that

$$d(dJ_i[f_{i,p}])[f_{i,q}] = \mathbf{Hess}(J_i)[f_{i,q},f_{i,p}] + dJ_i[\nabla_{f_{i,q}}f_{i,p}]. \tag{47}$$

Since $\mathcal{U}_i$ is a weakly embedded submanifold,

$$\nabla_{f_{i,q}}f_{i,p} = \nabla^i_{f_{i,q}}f_{i,p} + \Pi^i(f_{i,q},f_{i,p}),$$

where $\nabla^i_{f_{i,q}}f_{i,p} \in \mathcal{E}_i$, and Π^i is the second fundamental form. The equilibrium condition (12) gives $dJ_i[f_{i,r}]|_{x^\star} = 0$ for all r, so $dJ_i[\nabla^i_{f_{i,q}}f_{i,p}]|_{x^\star} = 0$, and hence $dJ_i[\nabla_{f_{i,q}}f_{i,p}]|_{x^\star} = dJ_i[\Pi^i(f_{i,q},f_{i,p})]|_{x^\star}$. From [22, Proposition 1.2], which is obtained by applying the Gauss equation of Riemannian geometry to the Hessian, we obtain that

$$\mathbf{Hess}(J_i^{\mathcal{U}_i})[f_{i,q},f_{i,p}] = \mathbf{Hess}^i(J_i)[f_{i,q},f_{i,p}] - dJ_i[\Pi^i(f_{i,q},f_{i,p})]. \tag{48}$$

Substituting into (47), the second fundamental form contributions cancel, and we obtain

$$d(dJ_i[f_{i,p}])[f_{i,q}]\big|_{x^\star} = \mathbf{Hess}^i(J_i^{\mathcal{U}_i})[f_{i,p},f_{i,q}]\big|_{x^\star}.$$

E. Proof of Proposition 4

Since $\{f_{i,j}\}$ is a global orthonormal frame on $\mathcal{U}$ with respect to g (see the proof of Theorem 1), we have $g(f_{i,p},f_{k,r}) = \delta_{ik}\delta_{pr}$, which is constant on $\mathcal{U}$. Therefore, all Lie-derivative terms in the Koszul formula [23, Corollary 5.11] vanish identically, and for any $k \in \{1,\dots,N\}$ and $r \in \{1,\dots,d_k\}$, we have that

$$\begin{aligned} 2g(\nabla_{f_{j,q}}f_{i,p},\,f_{k,r}) = &-g(f_{j,q},\,[f_{i,p},f_{k,r}]) \\ &+g(f_{i,p},\,[f_{k,r},f_{j,q}]) \\ &+g(f_{k,r},\,[f_{j,q},f_{i,p}])\,. \end{aligned} \tag{49}$$

Fixing $i \neq j$, the last term vanishes for all k, since, by Assumption 4, $[f_{j,q},f_{i,p}] = 0$. Next, we consider each possible value for k:

- For $k \notin \{i,j\}$, Assumption 4 gives that
$$[f_{i,p}, f_{k,r}] = 0, \quad [f_{k,r}, f_{j,q}] = 0,$$
so the right hand side in (49) vanishes.
- For $k = i$, $[f_{i,p}, f_{k,r}] \in \mathcal{E}_i$ by Assumption 3, so
$$g(f_{j,q}, [f_{i,p}, f_{i,r}]) = 0,$$
since $\mathcal{E}_j \perp_g \mathcal{E}_i$ (by construction), and, by Assumption 4,
$$[f_{i,r}, f_{j,q}] = 0.$$
- For $k = j$, Assumption 4 gives that
$$[f_{i,p}, f_{k,r}] = 0,$$
whereas, by Assumption 3, we have that $[f_{j,r}, f_{j,q}] \in \mathcal{E}_j$, and, therefore,
$$g(f_{i,p}, [f_{j,r}, f_{j,q}]) = 0,$$
since $\mathcal{E}_i \perp_g \mathcal{E}_j$ (by construction).

From the above, it follows that
$$g(\nabla_{f_{j,q}} f_{i,p}, f_{k,r}) = 0,$$
for all k, r, and, since $\{f_{k,r}\}$ is a global frame, (24) follows.

F. Proof of Proposition 5

First, we fix an order of the global frame as
$$(f_{1,1}, \dots, f_{1,d_1}, f_{2,1}, \dots, f_{2,d_2}),$$
and let $n := d_1 + d_2$. From (7b), the linear map sending $f_{i,p} \mapsto \tilde{f}_{i,p}$ has the matrix representation $I + \Gamma^\top$ with respect to this ordered frame, where
$$\Gamma = \begin{bmatrix} 0_{d_1 \times d_1} & \gamma \\ 0_{d_2 \times d_1} & 0_{d_2 \times d_2} \end{bmatrix}, \tag{50}$$
with $\gamma \in \mathbb{R}^{d_1 \times d_2}$ being the matrix of deception parameters γ_j^k. To prove the proposition, it suffices to show that $I + \Gamma^\top$, or, equivalently, $I + \Gamma$, is invertible, since, in that case, $\{\tilde{f}_{i,j}(x)\}$ is a basis of $T_x\mathcal{U}$, the rank-d_i subspaces $\tilde{\mathcal{E}}_i(x)$ are well-defined, and the direct-sum decomposition $T_x\mathcal{U} = \tilde{\mathcal{E}}_1(x) \oplus \tilde{\mathcal{E}}_2(x)$ follows from the basis property. Since the bottom-left block of Γ is zero, the matrix $I + \Gamma$ is block lower-triangular:
$$I + \Gamma = \begin{bmatrix} I_{d_1} & \gamma \\ 0 & I_{d_2} \end{bmatrix},$$
and consequently
$$\det(I + \Gamma) = \det(I_{d_1}) \cdot \det(I_{d_2}) = 1,$$
so $I + \Gamma$ is invertible at every $x \in \mathcal{U}$. Since $\{f_{i,j}(x)\}$ is a basis of $T_x\mathcal{U}$ at every $x \in \mathcal{U}$ by Assumption 1, and $I + \Gamma$ is invertible, the tilted family $\{\tilde{f}_{i,j}(x)\}$ is also a basis of $T_x\mathcal{U}$. In particular, $\{\tilde{f}_{i,j}(x)\}_{j=1}^{d_i}$ are linearly independent and span the rank-d_i subspace $\tilde{\mathcal{E}}_i(x)$, and the decomposition $T_x\mathcal{U} = \tilde{\mathcal{E}}_1(x) \oplus \tilde{\mathcal{E}}_2(x)$ follows from dimension counting.

G. Proof of Theorem 3

The proof proceeds in three steps. First, we recast the equilibrium condition (27) as a smooth map $F : \mathcal{U} \times \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}^n$ whose zero set parametrizes the equilibria of (8) as a function of γ. Second, we apply the implicit function theorem at $(x^\star, 0)$ using the non-degeneracy of the nominal Jacobian $A|_{x^\star}$ to obtain the smooth map $X^\star$. Third, we use continuity of the spectrum of the deceptive Jacobian $\tilde{A}|_{X^\star(\gamma)}$ to transfer local exponential stability from $\gamma = 0$ to a neighborhood Ω.

Let $\{\vartheta^{i,j}\}$ denote the dual co-frame of $\{f_{i,j}\}$ on $\mathcal{U}$, as in (17). The equilibrium condition (27) for system (8) reads
$$\mathbf{grad}_{\tilde{\mathcal{E}}_1}(J_1)(x) = 0, \qquad \mathbf{grad}_{\tilde{\mathcal{E}}_2}(J_2)(x) = 0.$$
Since $\tilde{f}_{1,j} = f_{1,j}$, the first condition is independent of γ and amounts to
$$dJ_1[f_{1,j}](x) = 0, \qquad j = 1, \dots, d_1.$$
Since $\tilde{f}_{2,j} = f_{2,j} + \sum_{k=1}^{d_1} \gamma_j^k f_{1,k}$, the second condition is
$$dJ_2[\tilde{f}_{2,j}](x) = dJ_2[f_{2,j}](x) + \sum_{k=1}^{d_1} \gamma_j^k \, dJ_2[f_{1,k}](x) = 0,$$
for $j = 1, \dots, d_2$. We now introduce the smooth map $F : \mathcal{U} \times \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}^n$ (with $n = d_1 + d_2$) by its n scalar components
$$\begin{aligned} F_{1,j}(x, \gamma) &= dJ_1[f_{1,j}](x), \\ F_{2,j}(x, \gamma) &= dJ_2[f_{2,j}](x) + \sum_{k=1}^{d_1} \gamma_j^k \, dJ_2[f_{1,k}](x). \end{aligned} \tag{51}$$
Then, $x \in \mathcal{U}$ is an equilibrium of (8) with deception parameter γ if and only if $F(x, \gamma) = 0$. By construction, $F(x^\star, 0) = 0$ (the nominal equilibrium conditions (12) at $x^\star$).

The partial derivative $\partial_x F(x^\star, 0) : T_{x^\star}\mathcal{U} \to \mathbb{R}^n$ is the linear map with scalar components, for $v \in T_{x^\star}\mathcal{U}$,
$$\begin{aligned} \partial_x F_{1,j}(x^\star, 0)[v] &= d\big(dJ_1[f_{1,j}]\big)[v]\big|_{x^\star}, \\ \partial_x F_{2,j}(x^\star, 0)[v] &= d\big(dJ_2[f_{2,j}]\big)[v]\big|_{x^\star}. \end{aligned}$$
Expanding $v = \sum_{i',p'} \vartheta^{i',p'}[v] \, f_{i',p'}(x^\star)$ and using the orthonormality of the frame, we obtain that
$$\partial_x F_{i,j}(x^\star, 0)[v] = \sum_{i',p'} d\big(dJ_i[f_{i,j}]\big)[f_{i',p'}]\big|_{x^\star} \, \vartheta^{i',p'}[v].$$
By the first equality in (20) of Theorem 2, the coefficients $d(dJ_i[f_{i,j}])[f_{i',p'}]|_{x^\star}$ are precisely the entries of the matrix representation of the nominal Jacobian $A|_{x^\star}$ with respect to the orthonormal frame $\{f_{i,j}\}$. Therefore, under the identification $\mathbb{R}^n \simeq T_{x^\star}\mathcal{U}$ provided by the frame $\{f_{i,j}(x^\star)\}$, the linear map $\partial_x F(x^\star, 0)$ coincides with the matrix representation of $A|_{x^\star}$:
$$\partial_x F(x^\star, 0) \;\simeq\; A|_{x^\star}. \tag{52}$$
By the hypothesis of local exponential stability of $x^\star$ and the "only if" direction of Proposition 1, $A|_{x^\star}$ is Hurwitz and, in particular, invertible. Hence $\partial_x F(x^\star, 0)$ is a linear isomorphism from $T_{x^\star}\mathcal{U}$ to $\mathbb{R}^n$.

The map F defined in (51) is smooth in both arguments (since $J_1, J_2, f_{i,j}$ are all smooth), $F(x^\star, 0) = 0$, and

$\partial_x F(x^\star, 0)$ is a linear isomorphism. By the implicit function theorem, there exists an open neighborhood $\Omega \subset \mathbb{R}^{d_1 \times d_2}$ of $\gamma = 0$ and a smooth map $X^\star : \Omega \to \mathcal{U}$ such that $X^\star(0) = x^\star$ and $F(X^\star(\gamma), \gamma) = 0$ for all $\gamma \in \Omega$. By construction, $X^\star(\gamma)$ is an equilibrium of (8) with deception parameter γ for every $\gamma \in \Omega$, and uniqueness in a neighborhood of $x^\star$ is part of the statement of the implicit function theorem.

Let $\tilde{A}(\gamma) := \tilde{A}|_{X^\star(\gamma)}$ denote the intrinsic Jacobian of the deceptive dynamics (8), evaluated at the deceptive equilibrium $X^\star(\gamma)$. Since both $X^\star$ and the frame $\{\tilde{f}_{i,j}\} = \{\tilde{f}_{i,j}(\gamma)\}$ depend smoothly on γ, and since the projected gradients $\mathbf{grad}_{\tilde{\mathcal{E}}_i}(J_i)$ depend smoothly on γ through $\tilde{f}_{i,j}$, the Jacobian $\tilde{A}(\gamma)$ is a smooth function of γ on Ω. At $\gamma = 0$, the deceptive frame $\{\tilde{f}_{i,j}\}$ coincides with the nominal frame $\{f_{i,j}\}$, the deceptive metric $\tilde{g}$ coincides with g, the deceptive bundles $\tilde{\mathcal{E}}_i$ coincide with $\mathcal{E}_i$, and $X^\star(0) = x^\star$. Consequently,

$$\tilde{A}(0) = A|_{x^\star},$$

which is Hurwitz by hypothesis. Since the spectrum of a matrix depends continuously on its entries, and smoothness of $\tilde{A}(\gamma)$ in γ implies continuity of its entries, there exists an open neighborhood $\Omega' \subset \Omega$ of $\gamma = 0$ such that $\tilde{A}(\gamma)$ remains Hurwitz for all $\gamma \in \Omega'$. Shrinking Ω to Ω' if necessary, the "if" direction of the second statement of Proposition 1 applied to (8) gives that $X^\star(\gamma)$ is a locally exponentially stable equilibrium of (8) for all $\gamma \in \Omega$.

H. Proof of Corollary 3

Differentiating the identity $F(X^\star(\gamma), \gamma) = 0$ with respect to γ_j^k at $\gamma = 0$ gives

$$\partial_x F(x^\star, 0) \left[\frac{\partial X^\star}{\partial \gamma_j^k}\bigg|_{\gamma=0} \right] + \frac{\partial F}{\partial \gamma_j^k}(x^\star, 0) = 0.$$

From (51), the only component of F that depends on γ_j^k is $F_{2,j}$, and $\partial F_{2,j}/\partial \gamma_j^k(x^\star, 0) = dJ_2[f_{1,k}](x^\star)$. Under the identification (52), the vector $\partial F/\partial \gamma_j^k(x^\star, 0) \in \mathbb{R}^n$ corresponds to $dJ_2[f_{1,k}](x^\star)\, f_{2,j}(x^\star) \in T_{x^\star}\mathcal{U}$, and inverting $\partial_x F(x^\star, 0) \simeq A|_{x^\star}$, which is invertible by assumption, yields (28).

VII. Conclusion

We studied model-free multi-agent decision making on manifolds under payoff-only measurements and control-affine actuation along state-dependent vector fields. Using a Lie-bracket approximation viewpoint, we showed that, under a natural bundle-splitting condition, the interaction dynamics admits an intrinsic geometric representation as a sum of projected Riemannian gradients, i.e., as a non-holonomic gradient play, which yields an equilibrium characterization and an intrinsic local stability test via a connection-based linearization. We then introduced an asymmetric-information model suitable for manifold-constrained interactions and showed how embedding victims' exploration signals into the strategy of a deceiver tilts the victims' effective actuation directions, generically shifting equilibria and enabling the deceiver to control the location of the equilibria, with an explicit first-order characterization of the induced shift. Future work will develop constructive systematic design rules and tuning algorithms for the deception parameters, and will address cyclic information structures in which multiple agents attempt to deceive one another.

Acknowledgement

This manuscript is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Grant No. HR00112530225 (COMPASS program).

References

[1] A. Sarlette, R. Sepulchre, and N. E. Leonard, "Autonomous rigid body attitude synchronization," *Automatica*, vol. 45, no. 2, pp. 572–577, 2009.
[2] W. Ren, "Distributed cooperative attitude synchronization and tracking for multiple rigid bodies," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 2, pp. 383–392, 2009.
[3] J. Poveda and N. Quijano, "Distributed extremum seeking for real-time resource allocation," in *2013 American Control Conference*, pp. 2772–2777, IEEE, 2013.
[4] D. E. Ochoa, J. I. Poveda, C. A. Uribe, and N. Quijano, "Hybrid robust optimal resource allocation with momentum," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 3954–3959, IEEE, 2019.
[5] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag, "Synchronization of coupled oscillators is a game," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 920–935, 2011.
[6] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "On the characterization of local Nash equilibria in continuous games," *IEEE transactions on automatic control*, vol. 61, no. 8, pp. 2301–2307, 2016.
[7] E. Mazumdar, L. J. Ratliff, and S. S. Sastry, "On gradient-based learning in continuous games," *SIAM Journal on Mathematics of Data Science*, vol. 2, no. 1, pp. 103–131, 2020.
[8] P. Frihauf, M. Krstic, and T. Basar, "Nash equilibrium seeking in non-cooperative games," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1192–1207, 2011.
[9] S.-J. Liu and M. Krstić, "Stochastic nash equilibrium seeking for games with general nonlinear payoffs," *SIAM Journal on Control and Optimization*, vol. 49, no. 4, pp. 1659–1679, 2011.
[10] M. S. Stankovic, K. H. Johansson, and D. M. Stipanovic, "Distributed seeking of nash equilibria with applications to mobile sensor networks," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 904–919, 2011.
[11] S. Krilasevic and S. Grammatico, "Learning generalized nash equilibria in multi-agent dynamical systems via extremum seeking control," *Automatica*, vol. 133, p. 109846, 2021.
[12] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *Journal of economic theory*, vol. 68, no. 1, pp. 258–265, 1996.
[13] M. Tang, U. Javed, X. Chen, M. Krstic, and J. I. Poveda, "Deception in nash equilibrium seeking," *IEEE Transactions on Automatic Control*, 2025.
[14] H.-B. Dürr, M. S. Stanković, C. Ebenbauer, and K. H. Johansson, "Lie bracket approximation of extremum seeking systems," *Automatica*, vol. 49, no. 6, pp. 1538–1552, 2013.
[15] V. Grushkovskaya, A. Zuyev, and C. Ebenbauer, "On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties," *Automatica*, vol. 94, pp. 151–160, 2018.
[16] M. Abdelgalil, D. E. Ochoa, and J. I. Poveda, "Multi-time scale control and optimization via averaging and singular perturbation theory: From odes to hybrid dynamical systems," *Annual Reviews in Control*, vol. 56, p. 100926, 2023.
[17] J. M. Lee in *Introduction to Smooth Manifolds*, Springer, 2012.
[18] A. Scheinker, "Model independent beam tuning," in *Proceedings of the 2013 International Particle Accelerator Conference, Shanghai, China*, pp. 12–17, 2013.
[19] A. Scheinker and M. Krstić, "Extremum seeking with bounded update rates," *Systems & Control Letters*, vol. 63, pp. 25–31, 2014.
[20] L. Moreau and D. Aeyels, "Practical stability and stabilization," *IEEE Transactions on Automatic Control*, vol. 45, no. 8, pp. 1554–1558, 2002.
[21] D. G. Feingold and R. S. Varga, "Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem.," 1962.

[22] M. Dajczer and R. Tojeiro, *Submanifold Theory: Beyond an Introduction*. Springer, 2019.
[23] J. M. Lee, *Introduction to Riemannian Manifolds*. Springer, 2018.



Mahmoud Abdelgalil is Assistant Professor in the Mechanical and Aerospace Engineering department at the University at Buffalo, State University of New York, Buffalo, NY. He received his BSc degree in Aerospace and Aeronautical Engineering from Zewail University, Egypt, in 2018, and his MSc and PhD in Mechanical and Aerospace Engineering from the University of California, Irvine, in 2020 and 2023, respectively, followed by a Postdoctoral appointment in the Electrical and Computer Engineering department at the University of California, San Diego. He is a recipient of the Holmes fellowship and the Henry Samueli fellowship at the University of California, Irvine, and the awardee of the 2023 ACC Best Student Paper.



Miroslav Krstic is Distinguished Professor and Senior Associate Vice Chancellor for Research at University of California, San Diego. He is Fellow of IEEE, IFAC, ASME, SIAM, AAAS, AIAA (Assoc. Fellow), IET (UK), and member of the Serbian Academy of Sciences and Arts. Krstic is a recipient of the Bellman Control Heritage Award, Bode Lecture Prize, SIAM Reid Prize, ASME Oldenburger Medal, Nyquist Lecture Prize, Paynter Award, Ragazzini Education Award, IFAC Nonlinear Control Systems Award, IFAC Ruth Curtain Distributed Parameter Systems Award, IFAC Adaptive and Learning Systems Award, Chestnut textbook prize, CSS Distinguished Member Award, the PECASE, NSF Career, and ONR Young Investigator awards, and the Schuck ('96 and '19) and Axelby paper prizes. Krstic is also the inaugural recipient of the A.V. Balakrishnan Award for the Mathematics of Systems. He is Editor-in-Chief of IEEE Transactions on Automatic Control, having previously served as Editor-in-Chief of Systems & Control Letters. Krstic has coauthored 19 books and about 500 journal papers.



Jorge I. Poveda is Associate Professor in the ECE Department at the University of California, San Diego, La Jolla, CA, USA. He received his M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of California, Santa Barbara in 2016 and 2018, respectively, followed by a Postdoctoral position at Harvard University. He was also a Research Intern at the Mitsubishi Electric Research Laboratories in 2016 and 2017. He has received the NSF CRII and CAREER awards, the AFOSR and SHPE Young Investigator awards, the 2023 AACC Donald P. Eckman award, the UCSB-CCDC Outstanding Scholar Fellowship (2013) and Best Ph.D. Thesis award (2020), and the 2023 IEEE Transactions on Control of Network Systems Best Paper award. He serves as Associate Editor for Automatica, IEEE Control Systems Letters, and Nonlinear Analysis: Hybrid Systems.