2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Riemannian polarization of multi-agent gradient flows

La Mi, Jorge Gonçalves, and Johan Markdahl

***Abstract*— Stable polarization of multi-agent systems has been shown to exist over $\mathbb{R}^n$ and highly symmetric nonlinear spaces, especially the n-sphere $\mathcal{S}^n$. Toward a more generalized setting without assuming linearity or symmetry, our previous work established the same type of emergent behavior over general hypersurfaces, subsuming the n-sphere case. In this paper, we discuss our ongoing work of extending our previous hypersurface results to study the stability of polarized equilibria of multi-agent gradient flows evolving on general Riemannian manifolds. The aim is to provide sufficient conditions in terms of the manifold geometry. Special attention is paid to two nonlinear manifolds of interest, the Stiefel manifold and the Grassmannian. While the polarization of the former share similar traits to that of the n-sphere, the latter is shown to have distinct polarization behaviors.**

I. Introduction

The main appeal of studying multi-agent synchronization is the bottom-up philosophy: simple computation and local communication give rise to emergent global phenomena. Given the simple agent assumption, we typically refrain from incorporating too much nonlinearity in the agent dynamics that increases complexity. This practice is in line with the broader philosophy, and additionally makes theoretical analysis tractable. Therefore, the most studied multi-agent synchronization models have agent state-space residing in simple manifolds, such as the linear diffusive coupling model in the Euclidean space $\mathbb{R}^n$ and the Kuramoto model on the circle $\mathcal{S}^1$.

Synchronization over other manifolds exist, often with specialized use cases and are predominantly focused on homogeneous spaces. Spheres $\mathcal{S}^n$ for generalized higher dimensional Kuramoto models [1], Lie groups for cooperative motion control [2] and quantum synchronization [3], Stiefel manifolds and Grassmannians [4] for matrix optimization algorithms. However, as computational power increases, it makes sense to increase the agent complexity for more modeling power, see e.g., the increasing sophistication of nonlinear opinion dynamics [5]. With the trending of customized solutions such as personalized medicine [6], rapid phenomenological modeling is required. Extending multi-agent systems to arbitrary manifolds may be one way to answer the call.

Previous works of synchronization over general manifolds mainly focused on consensus [7, 8]. Our contribution to this topic is to shed light on another type of emergent synchronization behavior over Riemannian manifolds, namely, polarization. The agent dynamics are governed by gradient flows constrained on Riemannian manifolds, which are simple enough to facilitate theoretical guarantee, while maintaining the flexibility to accommodate nonlinearity by shaping the agent state-space. There are two modes of interactions between agents: attraction and repulsion. These features have been previously incorporated in the Kuramoto model to study polarized oscillations [9] and equilibria [10]. Such modeling choices are inspired by real world phenomena, such as activation and inhibition of metabolites, cooperative and antagonistic social interactions. This study shows that the interplay between agent interaction modes and the manifold geometry is a rich source of stable polarization behavior.

The sufficient conditions for local stability in this paper are extended from our previous work on hypersurfaces [11]. Compared to [11], the extension to general Riemannian manifold removes the restrictive modeling assumption that the hypersurface dimension is lower than that of the ambience space by only one. This is advantageous for practice as meaningful data representation methods extract low dimensional manifolds from extremely high dimensional spaces. Another improvement is in the case of purely attractive dynamics, where we state cleaner Lyapunov stability conditions. The revised formulation better reflects the mechanism contributing to polarization, eliminating spurious artifacts of hypersurfaces. The performance of these conditions are illustrated on manifolds with special structures, which in some cases allow us to significantly increase the basin of attraction to almost the entire space.

II. Setup

A. Geometric features on the Riemannian manifold

Consider a closed Riemannian manifold $\mathcal{M}$ embedded in the Euclidean space $\mathbb{R}^{n\times p}$, which can always be done for sufficiently large n and p by the Whitney embedding theorem. The metric tensor is the Euclidean inner product on the tangent space $T_x\mathcal{M}$ at x. The normal space is denoted $N_x\mathcal{M}$, whose elements are normalized $\nu(x)'\nu(x) = I_p$ for convenience.

To define the geometric features on the manifold, we introduce the height function. The height function $h_{\nu(x)}: \mathcal{M} \to \mathbb{R}$ of a point $y \in \mathcal{M}$ with respect to a chosen point $x \in \mathcal{M}$ and a chosen normal vector $\nu(x) \in N_x\mathcal{M}$ is

$$h_{\nu(x)}(y) := \langle \nu(x), y \rangle.$$

Definition 1 (Dimple): If for some $x \in \mathcal{M}$ and some $\nu(x) \in N_x\mathcal{M}$, $y = x$ is a strict local minimizer of $h_{\nu(x)}(y)$ in a neighborhood $\mathcal{I}_x = \{y \in \mathcal{M} \mid \|y - x\| < \epsilon\}$, then $\mathcal{I}_x$ is referred to as a dimple with respect to $\nu(x)$, written $\mathcal{D}_{\nu(x)}$ and x the bottom of the dimple. Similarly, if for some $x \in \mathcal{M}$, $y = x$ is a strict local maximizer of $h_{\nu(x)}(y)$ in $\mathcal{I}_x$, then $\mathcal{I}_x$ is referred to as a pimple with respect to $\nu(x)$, written $\mathcal{P}_{\nu(x)}$ and x the bottom of the pimple.

The same $\mathcal{I}_x$ may be considered as a dimple or a pimple with respect to different choices of normal vectors, e.g., $\nu_1(x) \in N_x\mathcal{M}$ or $\nu_2(x) = -\nu_1(x) \in N_x\mathcal{M}$.

Remark 2: Generally, Riemannian manifolds are non-flat; otherwise it would just be the usual Euclidean (sub)spaces. So the existence of dimples/pimples are to be expected.

B. Multi-agent gradient flows

Evolving on the Riemannian manifold are N homogeneous agents forming an undirected, connected, and weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$. The adjacency matrix $A = [a_{ij}]$ is symmetric and has non-negative entries. The vertices $\mathcal{V}$ are divided into two groups $\mathcal{V}_\mathrm{u} = \{1, 2, \dots M\}$ and $\mathcal{V}_\mathrm{l} = \{M+1, \dots N\}$ for $1 \le M < N$. The edge set $\mathcal{E}$ is partitioned into intragroup and intergroup sets $\mathcal{E}_+ = \{\{i,j\} \in \mathcal{E} \,|\, i,j \in \mathcal{V}_\mathrm{u}$ or $i,j \in \mathcal{V}_\mathrm{l}\}$ and $\mathcal{E}_- = \{\{i,j\} \in \mathcal{E} \,|\, i \in \mathcal{V}_\mathrm{l}, j \in \mathcal{V}_\mathrm{u}\}$.

Such a partition is introduced to enforce different coupling rules over edges in $\mathcal{E}_+$ and $\mathcal{E}_-$. The couplings are positive over all edges in $\mathcal{E}_+$, whereas those over $\mathcal{E}_-$ can be either all positive or all negative. This "edge coloring" equivalently generates a *structurally balanced* graph [12] if we allow the graph to be signed such that edge weights on elements in $\mathcal{E}_+$ are all positive, and edge weights on elements in $\mathcal{E}_-$ are either all positive or all negative. Properties of structurally balanced networks allow us to make stronger stability claims in certain homogeneous manifolds.

Let $\chi := (x_i)_{i=1}^N$ denote the collection of agent positions. The dynamics of each agent is governed by the gradient flow of a disagreement function $V : \mathcal{M}^N \to \mathbb{R}$

$$\dot{x}_i = -\operatorname{grad} V(\chi) = -\Pi_i \nabla_i V(\chi), \tag{1}$$

where grad is the intrinsic gradient on the tangent space, which can be calculated using the second equality in (1) once the manifold embedding is defined. In the second equality of (1), Π_i is the orthogonal projection on the tangent space $T_i\mathcal{M}^N$, ∇_i is the extrinsic gradient in the ambient space with respect to x_i, and it is understood that ∇_i is applied to the extension of V in $\mathbb{R}^{n\times p}$.

The two interaction rules correspond to two quadratic disagreement functions based on the Euclidean distance. The one for attractive-repulsive interactions is

$$V_-(\chi) = \tfrac{1}{2} \sum_{\{i,j\}\in\mathcal{E}_+} a_{ij}\|x_j - x_i\|^2 - \tfrac{1}{2} \sum_{\{i,j\}\in\mathcal{E}_-} a_{ij}\|x_j - x_i\|^2, \tag{2}$$

whereas the one for purely attractive interactions is

$$V_+(\chi) = \frac{1}{2} \sum_{\{i,j\}\in\mathcal{E}} a_{ij}\|x_j - x_i\|^2. \tag{3}$$

The gradient flows corresponding to the two interaction rules are then

$$\begin{aligned} \dot{x}_i &= \Pi_i \left(\sum_{j\in\mathcal{V}_\mathrm{u}} a_{ij}(x_j - x_i) - \sum_{j\in\mathcal{V}_\mathrm{l}} a_{ij}(x_j - x_i) \right), \quad i \in \mathcal{V}_\mathrm{u} \\ \dot{x}_i &= \Pi_i \left(\sum_{j\in\mathcal{V}_\mathrm{l}} a_{ij}(x_j - x_i) - \sum_{j\in\mathcal{V}_\mathrm{u}} a_{ij}(x_j - x_i) \right), \quad i \in \mathcal{V}_\mathrm{l} \end{aligned} \tag{4}$$

and

$$\dot{x}_i = \Pi_i \sum_{j\in\mathcal{V}} a_{ij}(x_j - x_i) \quad \forall i \in \mathcal{V} \tag{5}$$

Remark 3: It is readily seen from (4) and (5) that agent i does not need to know the position of all other agents in the system, as $a_{ij} \neq 0$ only for j connected to i, i.e., for j in the neighborhood of i.

C. Gradient flows on Riemannian manifolds

Summarizing the aforementioned ingredients in §II-A and §II-B, we have a multi-agent gradient flow system with repulsive (4) or attractive (5) intergroup interactions evolving on a manifold $\mathcal{M}^N$ with geometric features. We assume that the manifold is outfitted with a pair of dimples and pimples, each containing one of the two groups of agents $\mathcal{V}_\mathrm{u}$ and $\mathcal{V}_\mathrm{l}$. We are interested in possible polarization arising in this setting as a result of the interplay between the graph couplings and the geometry of the underlying nonlinear space.

Definition 4 (Polarization): The agents are said to be polarized if $x_i = x_j$ for all $\{i,j\} \in \mathcal{E}_+$ and $x_i \neq x_j$ for all $\{i,j\} \in \mathcal{E}_-$.

It follows that a polarized configuration must be of the following form

$$\begin{aligned} \chi^* := \{\chi \in \mathcal{M}^N \,| x_i &= x_\mathrm{u} \in \mathcal{M},\ i \in \mathcal{V}_\mathrm{u}, \\ x_i &= x_\mathrm{l} \in \mathcal{M},\ i \in \mathcal{V}_\mathrm{l},\ \ x_\mathrm{u} \neq x_\mathrm{l}\}. \end{aligned} \tag{6}$$

We focus on polarization as an equilibrium (set) and its stability properties, as equilibria are the only possible attractor for gradient flows on manifolds [13, App. C.12]. For polarization to occur, a necessary condition is that there exists $\nu_\mathrm{u} \in N_{x_\mathrm{u}}\mathcal{M}$ and $\nu_\mathrm{l} \in N_{x_\mathrm{l}}\mathcal{M}$ such that ν_u and ν_l are aligned, as implied by the following result.

Proposition 5: If the gradient flow system (5) or (4) converges to a polarized equilibrium, then $x_\mathrm{u} - x_\mathrm{l} \in N_{x_\mathrm{u}}\mathcal{M} \cap N_{x_\mathrm{l}}\mathcal{M}$.

Proof: Consider the system (5). At a polarized equilibrium (6), the aggregate attraction $u_i := \sum_{j\in\mathcal{V}} a_{ij}(x_j - x_i)$ for agent $i \in \mathcal{V}_\mathrm{u}$ is equal to $\sum_{j\in\mathcal{V}_\mathrm{l}} a_{ij}(x_\mathrm{l} - x_\mathrm{u})$. It is nonzero since $x_\mathrm{l} \neq x_\mathrm{u}$ and belongs to $\ker \Pi_\mathrm{u}$ by (5). Likewise, $0 \neq u_i = \sum_{j\in\mathcal{V}_\mathrm{u}} a_{ij}(x_\mathrm{u} - x_\mathrm{l}) \in \ker \Pi_\mathrm{l}$ for agent $i \in \mathcal{V}_\mathrm{l}$. The kernel space of a projection operator on $T_x\mathcal{M}$ is equal to the normal space $N_x\mathcal{M}$, whereby the conclusion follows. The reasoning for system (4) is identical. ■

In certain geometrical conditions, this necessary condition also becomes sufficient for the stability of the polarized equilibrium, to be shown in §III-A.

Before we proceed, we collect a few previous results and associated definitions that will pave the way for the development of our main results. The definitions of concepts such as stability and a local minimizer of a real function are well known. Here, we clarify their meanings when a set rather than a point is in question, which is perhaps less standard.

Consider two sets in the Euclidean space $\mathcal{Y}, \mathcal{Z} \subset \mathbb{R}^{n\times p}$ when using the Hausdorff distance to define stability

$$d_{\mathrm{H}}(\mathcal{Y}, \mathcal{Z}) := \max\{\sup_{y\in\mathcal{Y}} \inf_{z\in\mathcal{Z}} \|y - z\|, \sup_{z\in\mathcal{Z}} \inf_{y\in\mathcal{Y}} \|y - z\|\}.$$

Definition 6 (Stability): A set of equilibria $\mathcal{S}$ is Lyapunov stable if, for each $\epsilon > 0$, there is $\delta = \delta(\epsilon)$ such that $d_{\mathrm{H}}(x, \mathcal{S})|_{t=0} < \delta$ implies $d_{\mathrm{H}}(x, \mathcal{S})|_t < \epsilon$ for all $t \geq 0$; is asymptotically stable if it is stable and δ can be chosen such that $d_{\mathrm{H}}(x, \mathcal{S})|_{t=0} < \delta$ implies $\lim_{t\to\infty} d_{\mathrm{H}}(x, \mathcal{S}) = 0$.

Definition 7 (Local minimizer): A set $\mathcal{S} \subset \mathcal{M}$ is said to be a *local minimizer* of a real function $f : \mathcal{M} \to \mathbb{R}$ from a metric space $(\mathcal{M}, d_{\mathrm{H}})$ if for some $\epsilon > 0$, there is an open neighborhood $\mathcal{N}(\mathcal{S}) = \{x \in \mathcal{M} \mid d_{\mathrm{H}}(x, \mathcal{S}) < \epsilon\}$ such that $f|_{\mathcal{S}} \leq f(x)$ for all $x \in \mathcal{N}(\mathcal{S})$. Moreover, if the inequality is strict for all $x \in \mathcal{N}(\mathcal{S})\backslash\mathcal{S}$, then $\mathcal{S}$ is said to be a *strict local minimizer*.

Definition 8 (Isolated critical): A set $\mathcal{S} \subset \mathcal{M}$ of critical points of a real function $f : \mathcal{M} \to \mathbb{R}$ from a metric space $(\mathcal{M}, d_{\mathrm{H}})$ is said to be *isolated critical* if for some $\epsilon > 0$, there is an open neighborhood $\mathcal{N}(\mathcal{S}) = \{x \in \mathcal{M} \mid d_{\mathrm{H}}(x, \mathcal{S}) < \epsilon\}$ such that $\mathcal{N}(\mathcal{S})\backslash\mathcal{S}$ is void of critical points.

Proposition 9 (Manifold Lyapunov theorem): Let $\mathcal{M}$ be a closed manifold and take any $V : \mathcal{M} \to \mathbb{R}$ that is C^2. Let $\mathcal{S}$ be a compact set of local minimizers of V. If $\mathcal{S}$ is a strict local minimizer, then $\mathcal{S}$ is a Lyapunov stable equilibrium set of $\dot{x} = -\operatorname{grad} V$. If $\mathcal{S}$ is also isolated critical, then it is asymptotically stable.

III. Main results

Though the geometric features defined in Definition 1 can be considered ubiquitous as mentioned in Remark 2, not every pair admits the existence of an equilibrium for a given gradient flow, let alone a stable one. In this section, we present and discuss sufficient conditions for the existence of polarized equilibria and their stability properties. They arise in different combinations of attractive/repulsive interactions with dimple/pimple geometric features.

A. Dimple pairs with attractive intergroup couplings

Assumption 10: For all $i \in \mathcal{V}_{\mathrm{u}}$, $x_i \in \mathcal{D}_{\nu_{\mathrm{u}}(x_{\mathrm{u}})}$; for all $i \in \mathcal{V}_{\mathrm{l}}$, $x_i \in \mathcal{D}_{\nu_{\mathrm{l}}(x_{\mathrm{l}})}$.

Proposition 11: For system (5) under Assumption 10, the polarized configuration χ^* defined in (6) is a strict local minimizer of V_+ if

1) $\langle x_{\mathrm{u}} - x_{\mathrm{l}}, \nu_{\mathrm{u}}\rangle = \|x_{\mathrm{u}} - x_{\mathrm{l}}\|$ and
2) $\nu_{\mathrm{u}} = -\nu_{\mathrm{l}}$.

Proof: We examine the disagreement terms $\|x_j - x_i\|^2$ over $\mathcal{E}_-$ and $\mathcal{E}_+$ in (3) separately. For $\{i, j\} \in \mathcal{E}_-$, suppose without loss of generality that $j \in \mathcal{V}_{\mathrm{u}}$ and $i \in \mathcal{V}_{\mathrm{l}}$, then

$$\begin{aligned} p\|x_j - x_i\|^2 \geq \langle x_j - x_i, \nu_{\mathrm{u}}\rangle^2 &= \big(h_{\nu_{\mathrm{u}}}(x_j) - h_{\nu_{\mathrm{u}}}(x_i)\big)^2 \\ &\geq (h_{\nu_{\mathrm{u}}}(x_{\mathrm{u}}) - h_{\nu_{\mathrm{u}}}(x_{\mathrm{l}}))^2, \end{aligned}$$

The first lower bound due to Cauchy-Schwartz (mind that $\|\nu_{\mathrm{u}}\|^2 = p$) is achieved only when $x_j - x_i$ is parallel to ν_{u}. To show the last inequality, note that

$$h_{\nu_{\mathrm{u}}}(x_j) \geq h_{\nu_{\mathrm{u}}}(x_{\mathrm{u}}), \quad \forall j \in \mathcal{V}_{\mathrm{u}} \tag{7a}$$
$$h_{\nu_{\mathrm{l}}}(x_i) \geq h_{\nu_{\mathrm{l}}}(x_{\mathrm{l}}), \quad \forall i \in \mathcal{V}_{\mathrm{l}} \tag{7b}$$

as per Assumption 10. Equation (7b) leads to

$$-h_{\nu_{\mathrm{u}}}(x_i) \geq -h_{\nu_{\mathrm{u}}}(x_{\mathrm{l}}), \quad \forall i \in \mathcal{V}_{\mathrm{l}} \tag{8}$$

due to condition 2. Combining (7a) and (8) to obtain $h_{\nu_{\mathrm{u}}}(x_j) - h_{\nu_{\mathrm{u}}}(x_i) \geq h_{\nu_{\mathrm{u}}}(x_{\mathrm{u}}) - h_{\nu_{\mathrm{u}}}(x_{\mathrm{l}}) > 0$, where the second inequality is a result of condition 1. Therefore, squaring both sides does not change the direction of the inequality. The second lower bound is achieved only when $x_j = x_{\mathrm{u}}$ and $x_i = x_{\mathrm{l}}$.

For $\{i, j\} \in \mathcal{E}_+$, the lower bound is trivially zero and is achieved only when $x_i = x_j$. In summary, the only configuration to simultaneously achieve all the lower bounds above is χ^* in (6). ■

Applying the first part of Prop. 9 to Prop. 11 thus establishes when the necessary condition in Prop. 5 becomes sufficient for stable polarization in the Lyapunov sense. We next provide a local asymptotic stability result by imposing a further requirement on the manifold to satisfy the second condition of Prop. 9.

Theorem 12: For system (5) under Assumption 10, if the conditions in Prop. 11 are satisfied, and the manifold is analytic, then χ^* defined in (6) is an asymptotically stable polarized equilibrium.

The proof hinges on a variant [14, Sec. 9] of the Łojasiewicz inequality valid on analytic Riemannian manifolds. The proof steps for the corresponding hypersurface result [11, Thm. 14] can be applied in the Riemannian manifold case; a repetition is unnecessary.

B. Pimple pairs with repulsive intergroup couplings

Assumption 13: For all $i \in \mathcal{V}_{\mathrm{u}}$, $x_i \in \mathcal{P}_{\nu_{\mathrm{u}}(x_{\mathrm{u}})}$; for all $i \in \mathcal{V}_{\mathrm{l}}$, $x_i \in \mathcal{P}_{\nu_{\mathrm{l}}(x_{\mathrm{l}})}$.

Although the setting in this subsection mirrors that in §III-A, reusing the same conditions in Prop. 11 would not work, because here we need to upperbound rather than lowerbound the disagreement terms over $\mathcal{E}_-$. To state the alternative condition, we denote by $\mathcal{B}_{r_o}(x_o)$ the *closed* ball centered at $x_o = \frac{1}{2}(x_{\mathrm{u}}+x_{\mathrm{l}})$ with radius $r_o = \frac{1}{2}\|x_{\mathrm{u}}-x_{\mathrm{l}}\|$. The equilibrium we are interested in is the following set

$$\begin{aligned} \mathcal{C}_{\mathrm{lem}} := \{\chi \in \mathcal{M}^N \mid &x_i = x,\ i \in \mathcal{V}_{\mathrm{u}}, \\ &x_i = y,\ i \in \mathcal{V}_{\mathrm{l}},\ (x, y) \in Y\}, \end{aligned} \tag{9}$$

where $Y := \{(x, y) \in \mathcal{P}_{\nu_{\mathrm{u}}(x_{\mathrm{u}})} \times \mathcal{P}_{\nu_{\mathrm{l}}(x_{\mathrm{l}})} \mid \|x - y\| = 2r_o > 0\}$. Of course, each element of (9) conforms to the form in (6).

Proposition 14: For system (4) under Assumption 13, the polarized configuration $\mathcal{C}_{\mathrm{lem}}$ defined in (9) is a strict local minimizer of V_- if $\mathcal{P}_{\nu_{\mathrm{u}}(x_{\mathrm{u}})}$ and $\mathcal{P}_{\nu_{\mathrm{l}}(x_{\mathrm{l}})}$ are entirely contained in $\mathcal{B}_{r_o}(x_o)$ with $r_o > 0$.

Proof: We shall minimize V_-, which is the sum of disagreement terms $\|x_j - x_i\|^2$ among agents. Under Assumption 13, the bounds on the disagreement terms differ for agents belonging to the same or opposing groups. For all

$\{i,j\} \in \mathcal{E}_-$, assume without loss of generality that $i \in \mathcal{V}_\mathrm{u}$ and $j \in \mathcal{V}_\mathrm{l}$. The term $\|x_j - x_i\|^2$ in (2) is upper bounded by

$$\|x_j - x_i\| \le 2r_o = \|x_\mathrm{u} - x_\mathrm{l}\|.$$

The inequality is because both pimples are entirely contained in $\mathcal{B}_{r_o}(x_o)$. The upper bound is achieved when $x_i = x$ and $x_j = y$ for every pair $(x, y) \in \mathcal{P}_{\nu_\mathrm{u}(x_\mathrm{u})} \times \mathcal{P}_{\nu_\mathrm{l}(x_\mathrm{l})}$ such that $\|x - y\| = 2r_o$, an example of which is $x = x_\mathrm{u}$ and $y = x_\mathrm{l}$.

For all $\{i,j\} \in \mathcal{E}_+$,

$$\|x_j - x_i\|^2 \ge 0,$$

where the equality is achieved when $x_i = x_j$, of which a special case is $x_i = x_j = x_\mathrm{u}$ for $i, j \in \mathcal{V}_\mathrm{u}$ and $x_i = x_j = x_\mathrm{l}$ for $i, j \in \mathcal{V}_\mathrm{l}$. Therefore,

$$V_-(\chi) \ge -\tfrac{1}{2}\sum_{\{i,j\}\in\mathcal{E}_-} a_{ij}\|x_j - x_i\|^2 \ge -2r_o^2 \sum_{\{i,j\}\in\mathcal{E}_-} a_{ij}.$$

The minimum is achieved only when $x_i = x, i \in \mathcal{V}_\mathrm{u}$ and $x_i = y, i \in \mathcal{V}_\mathrm{l}$ for every pair of $(x, y) \in \mathcal{P}_{\nu_\mathrm{u}(x_\mathrm{u})} \times \mathcal{P}_{\nu_\mathrm{l}(x_\mathrm{l})}$ such that $\|x - y\| = 2r_o$. ■

Again, Lyapunov stability can be concluded for $\mathcal{C}_\mathrm{lem}$ by applying Prop. 9 to Prop. 14.

Theorem 15: For system (4) under Assumption 13, if the conditions given in Prop. 14 is satisfied, and in addition, the manifold is analytic, then $\mathcal{C}_\mathrm{lem}$ defined in (9) is an asymptotically stable set of polarized equilibria.

The proof of Theorem 15 is more involved, as it concerns the asymptotic stability of an equilibrium set, rather than an equilibrium point as the singleton set χ^* in Theorem 12. The first part of the proof is to show that pointwise, there are no equilibrium points other than those in $\mathcal{C}_\mathrm{lem}$ in the neighborhood of each $\chi \in \mathcal{C}_\mathrm{lem}$. That is, if another equilibrium set $\mathcal{Q}$ exists on the manifold such that $\mathcal{Q} \cap \mathcal{C}_\mathrm{lem} = \emptyset$, each $\chi \in \mathcal{C}_\mathrm{lem}$ is isolated from $\mathcal{Q}$. This is again done by invoking the Łojasiewicz inequality as in the proof of Theorem 12. In the second part, we show that $\mathcal{C}_\mathrm{lem}$ is isolated critical by proving no sequence in $\mathcal{Q}$ can approach $\mathcal{C}_\mathrm{lem}$ arbitrarily close using the Bolzano-Weierstrass theorem. For details, we refer the readers to the proof of [11, Thm. 18].

Since we can freely choose any normal vector in the normal space, making the opposite choices for ν_u and ν_l brings us the following corollaries.

Corollary 16: For system (4) under Assumption 10, the polarized configuration $\mathcal{C}_\mathrm{lem}$ defined in (9) is a strict local minimizer of V_- if $\mathcal{D}_{\nu_\mathrm{u}(x_\mathrm{u})}$ and $\mathcal{D}_{\nu_\mathrm{l}(x_\mathrm{l})}$ are entirely contained in $\mathcal{B}_{r_o}(x_o)$ with $r_o > 0$, and is additionally asymptotically stable if $\mathcal{M}$ is analytic.

Corollary 17: For system (5) under Assumption 13, the polarized configuration χ^* defined in (6) is a strict local minimizer of V_+ if $\langle x_\mathrm{u} - x_\mathrm{l}, \nu_\mathrm{u}\rangle = -\|x_\mathrm{u} - x_\mathrm{l}\|$ and $\nu_\mathrm{u} = -\nu_\mathrm{l}$, and is additionally asymptotically stable if $\mathcal{M}$ is analytic.

IV. EXAMPLES

A. The Stiefel manifold

The compact real Stiefel manifold $\mathrm{St}(n, p)$ is the set of p-frames in n-dimensional Euclidean space $\mathbb{R}^n$ [15]. It can be embedded in $\mathbb{R}^{n\times p}$ as an analytic matrix manifold

$$\mathrm{St}(n, p) = \{U \in \mathbb{R}^{n\times p} \mid U'U = I_p\}.$$

The inner product and the norm are the trace operator and the Frobenius norm, respectively. Two useful projections when dealing with the Stiefel manifold are sym: $\mathbb{R}^{p\times p} \to \mathrm{SO}(p)^\perp : X \mapsto (X + X')/2$ and skew: $\mathbb{R}^{p\times p} \to \mathrm{SO}(p) : X \mapsto (X - X')/2$.

Every point on the Stiefel manifold can be seen as a pimple bottom. To see this, start with the normal space at a point $U \in \mathrm{St}(n, p)$

$$N_U \mathrm{St}(n, p) = \{U \,\mathrm{sym}\, U'X \mid X \in \mathbb{R}^{n\times p}\}.$$

The normal vector corresponding to the choice of $X = U$ is U. The height function of U with respect to U_0 and the particular choice of normal vector $U_0 \in N_{U_0}\mathrm{St}(n, p)$ is then

$$\begin{aligned} h_{U_0}(U) &= \langle U_0, U\rangle \\ &= \mathrm{tr}\, U'U_0 < \mathrm{tr}\, U_0'U_0 = p = h_{U_0}(U_0), \quad \forall U \ne U_0. \end{aligned}$$

Thus, we have shown the existence of an X, e.g., $X = U_0$ such that $U = U_0$ is a strict local maximizer of $h_{\nu(U_0)}(U)$ on $\mathrm{St}(n, p)$, as required by the pimple definition in Def. 1. In fact, this choice of the normal coincides with the (unique) normal of the $\mathcal{S}^{n\times p-1}$ at a point $\mathrm{vec}\, U_0$, which is known to be everywhere pimple [11, §3.3] and of which $\mathrm{St}(n, p)$ is a subset.

On $\mathrm{St}(n, p)$, choose U_u and U_l such that $U_\mathrm{u} = -U_\mathrm{l}$ to define the closed ball $\mathcal{B}_{r_o}(U_o)$, where $U_o = \frac{1}{2}(U_\mathrm{u} + U_\mathrm{l}) = 0$ and $r_o = \frac{1}{2}\|U_\mathrm{u} - U_\mathrm{l}\| = \sqrt{p}$. A polarized configuration on the Stiefel manifold is

$$\begin{aligned} \mathcal{C}_\mathrm{St} := \{\chi \in \mathrm{St}(n, p)^N \mid\, & U_i = U_\mathrm{u},\ \forall i \in \mathcal{V}_\mathrm{u}, \\ & U_i = U_\mathrm{l},\ \forall i \in \mathcal{V}_\mathrm{l},\ U_\mathrm{u} = -U_\mathrm{l}\}. \end{aligned} \quad (10)$$

It can be verified that this configuration is an equilibrium of the Stiefel version of (4), where $\Pi_i(X) = U_i \,\mathrm{skew}\, U_i'X + (I - U_iU_i')X$.

It can be verified that every point on $\mathrm{St}(n, p)$ is on the boundary of the ball $\partial\mathcal{B}_{\sqrt{p}}(0)$, and therefore Prop. 14 applies. In this respect, $\mathrm{St}(n, p)$ resembles the sphere of radius $\sqrt{p}$. Moreover, for certain (n, p) pairs, the equilibrium set $\mathcal{C}_\mathrm{St}$ is almost global asymptotically stable (AGAS), by which we mean

Definition 18 (almost global asymptotic stability): A set of equilibria $\mathcal{D} \subset \mathcal{M}^N$ is almost globally attractive if for all initial conditions except a measure-zero set, it holds that $\lim_{t\to\infty} \chi(t) \in \mathcal{D}$. Moreover, if $\mathcal{D}$ is stable, $\mathcal{D}$ is almost globally asymptotically stable.

The measure-zero set is with respect to the Lebesgue measure.

Proposition 19: For $p \le \frac{2}{3}n - 1$, the polarized configuration $\mathcal{C}_\mathrm{St}$ is an AGAS equilibrium set of the gradient descent flow (4) on $\mathrm{St}(n, p)$.

The proof exploits the symmetry of $\mathrm{St}(n, p)$ and employs a coordinate transform that turns the polarization problem into a consensus one that was already solved. The coordinate transform is akin to the gauge transformation noted in [12, Lem. 1] that brings a structurally balanced signed graph to a nonnegative one.

Proof: Apply a coordinate transformation $V_i = U_i$ for all $i \in \mathcal{V}_u$ and $V_i = -U_i$ for all $i \in \mathcal{V}_l$. The gradient descent flow (4) on $\mathrm{St}(n, p)$ becomes a consensus seeking system

$$\dot{U}_i = U_i \,\mathrm{skew}(U_i' \sum_j a_{ij} U_j) + (I - U_i U_i') \sum_j a_{ij} U_j, \quad \forall i \in \mathcal{V}. \quad (11)$$

This system was considered in [16, Thm. 4], which guarantees the AGAS property of the consensus set

$$\mathcal{C} = \{\chi \in \mathrm{St}(n, p)^N \mid U_i = U_j \ \forall i, j \in \mathcal{E}\}.$$

The consensus set $\mathcal{C}$ is mapped to the polarization set $\mathcal{C}_{\mathrm{St}}$ by reversing the bijective coordinate transform. Therefore, applying [16, Thm. 4] to (11) and then reversing the coordinate transformation, we obtain the desired conclusion. ■

The AGAS property of polarization on the n-sphere for $n > 1$ is shown in [11, Thm. 24]. Thus, the Stiefel manifold shares similarities with the sphere in terms of polarization properties. However, the next example cautions that not all manifolds that are a subset of the sphere exhibit comparable polarization behaviors.

B. The Grassmannian manifold

A point on the Grassmannian $\mathrm{Gr}(n, p)$ is a p-dimensional linear subspace $\mathcal{U}$ in $\mathbb{R}^n$ [15]. Since $\mathrm{Gr}(n, n-p)$ is isomorphic to $\mathrm{Gr}(n, p)$, we identify orthogonally complementary subspaces and assume without loss of generality [17, §2.2] that $p \leq n/2$. Points on the Grassmann manifold can be represented in various ways. We prefer a unique representation for unique identification of each agent in the multi-agent gradient flow systems. We therefore choose to represent each Grassmannian point as an orthogonal projector $P \in \mathbb{R}^{n\times n}$ onto $\mathcal{U}$. It can be uniquely represented by $P = UU'$, where the columns of $U \in \mathbb{R}^{n\times p}$ form an orthonormal basis of $\mathcal{U}$ [18]. The projector representation of the Grassmannian can thus be viewed as

$$\mathrm{Gr}(n, p) := \{P \in \mathbb{R}^{n\times n} \mid P^2 = P, \mathrm{rank}(P) = p, P = P'\}. \quad (12)$$

The inner product of P and Q is $\mathrm{tr}\, Q'P$, and thus we have the Frobenius norm $\|P\| = \sqrt{\mathrm{tr}\, P'P}$. For all $P \in \mathrm{Gr}(n, p)$, we have $\mathrm{tr}\, P'P = \mathrm{tr}\, UU' = \mathrm{tr}\, U'U = \mathrm{tr}\, I_p = p$, implying that $\mathrm{Gr}(n, p)$ is a subset of a sphere of radius $\sqrt{p}$.

Again, we can regard the Grassmanniann as everywhere pimple. The normal space at a point $P \in \mathrm{Gr}(n, p)$ is [19, (2.17)]

$$N_P \mathrm{Gr}(n, p) = \{X - [P, [P, X]] \mid X = X'\},$$

where we use the matrix commutator bracket $[A, B] = AB - BA$. The height function of P with respect to P_0 and a normal vector $\nu(P_0) \in N_{P_0}\mathrm{Gr}(n, p)$ is

$$\begin{aligned} h_{\nu(P_0)}(P) &= \langle X - [P_0, [P_0, X]], P\rangle \\ &= \mathrm{tr}(X - P_0 X + 2P_0 X P_0 - X P_0)P. \end{aligned}$$

Choose $X = P_0$, then the normal vector corresponding to the choice $X = P_0$ is P_0 and

$$h_{\nu(P_0)}(P) = \mathrm{tr}\, P_0 P < \mathrm{tr}\, P_0 = h_{\nu(P_0)}(P_0), \quad \forall P \neq P_0.$$

Thus, we have shown the existence of an X, e.g., $X = P_0$ such that $P = P_0$ is a strict local maximizer of $h_{\nu(P_0)}(P)$ on $\mathrm{Gr}(n, p)$, as required by the pimple definition in Def. 1. This choice of the normal corresponds to the (unique) normal of the $\mathcal{S}^{n^2-1}$ at a point P_0, which is known to be everywhere pimple.

For $\mathrm{Gr}(n, p)$, the disagreement term satisfies $\|P_j - P_i\|^2 = \mathrm{tr}(P_j - P_i)'(P_j - P_i) = 2(p - \mathrm{tr}\, P_j' P_i)$. The Grassmannian version of (4) then reads

$$\begin{aligned} \dot{P}_i &= \sum_{j\in\mathcal{V}_u} a_{ij}(P_i P_j P_i^\perp + P_i^\perp P_j P_i) \\ &\quad - \sum_{j\in\mathcal{V}_l} a_{ij}(P_i P_j P_i^\perp + P_i^\perp P_j P_i), \ \ \forall i \in \mathcal{V}_u \\ \dot{P}_i &= \sum_{j\in\mathcal{V}_l} a_{ij}(P_i P_j P_i^\perp + P_i^\perp P_j P_i) \\ &\quad - \sum_{j\in\mathcal{V}_u} a_{ij}(P_i P_j P_i^\perp + P_i^\perp P_j P_i), \ \ \forall i \in \mathcal{V}_l, \end{aligned} \quad (13)$$

where $P^\perp = I - P$. The distance between two subspaces can also be measured using principal angles [20], denoted θ_i, so that $\mathrm{tr}\, Q'P = \sum_i^p \cos^2\theta_i$. They can be thought of as the angles between the orthonormal bases of P and Q.

On $\mathrm{Gr}(n, p)$, choose P_u and P_l such that $P_u' P_l = 0$ to define the closed ball $\mathcal{B}_{r_o}(P_o)$, where we emphasize that $P_o = \frac{1}{2}(P_u + P_l)$ does not belong to $\mathrm{Gr}(n, p)$ but rather to $\mathrm{Gr}(n, 2p)$. A polarized configuration on $\mathrm{Gr}(n, p)$ is

$$\mathcal{C}_{\mathrm{Gr}} := \{\chi \in \mathrm{Gr}(n, p)^N \mid P_i = P_u, \forall i \in \mathcal{V}_u, P_i = P_l, \forall i \in \mathcal{V}_l, \ P_u' P_l = 0\}. \quad (14)$$

It can be checked that this is an equilibrium of (13).

Proposition 20: Every point $P \in \mathrm{Gr}(n, p)$ is on the boundary $\partial\mathcal{B}_{r_o}(P_o)$ for n even and $p = n/2$. Otherwise, for $p < n/2$, there exists $P \in \mathrm{Gr}(n, p) \notin \mathcal{B}_{r_o}(P_o)$.

Proof: The radius of the ball is

$$r_o^2 = (\tfrac{1}{2}\|P_u - P_l\|)^2 = \tfrac{1}{4}\,\mathrm{tr}\,(P_u - P_l)^2 = \tfrac{1}{4}\,\mathrm{tr}(P_u + P_l) = \tfrac{1}{2}p.$$

The distance between any point P and P_o is

$$\begin{aligned} \|P - P_o\|^2 &= \mathrm{tr}(P^2 - 2PP_o + P_o^2) \\ &= \mathrm{tr}\, P - \mathrm{tr}\, P(P_u + P_l) + \tfrac{1}{4}\,\mathrm{tr}\,(P_u + P_l)^2 \\ &= p - \mathrm{tr}\, P(P_u + P_l) + \tfrac{1}{4}\,\mathrm{tr}(P_u + P_l) \\ &= \tfrac{3}{2}p - \mathrm{tr}\, P(P_u + P_l). \end{aligned}$$

For $p = n/2$, P_u and P_l are orthogonal complements of each other in $\mathbb{R}^n$, which implies that $P_u + P_l = I_n$, as $\mathbb{R}^n$ itself is represented by the projection matrix I_n. (To see this, form the matrix U whose columns are n orthonormal bases so that $U'U = I$. But $U'U = UU'$ because U is full rank, so $UU' = I$ is the projector representation of $\mathrm{Gr}(n, n)$.) Thus

$$\|P - P_o\|^2 = \tfrac{3}{2}p - p = \tfrac{1}{2}p = r_o^2.$$

This is proves the first statement.

Otherwise, for $p < n/2$, the term $\operatorname{tr} P(P_u + P_l)$ is upper-bounded by p. To see this, we state the maximization problem as

$$\max \operatorname{tr} P(P_u + P_l) = \max \sum_{i}^{p} \left(\cos^2 \theta_{u,i} + \cos^2 \theta_{l,i}\right)$$
$$\text{subject to } P_u P_l = 0,$$

where $\theta_{u,i}$ and $\theta_{l,i}$ are one of the principal angles between P and P_u and between P and P_l, respectively. Note that there is an ordering in the definition of the principal angles [20], but it is not essential here. This form suggests that the maximization can be done for each i independently, since the choice of each basis of P does not depend on other bases. We can restate the maximization problem as

$$\max \cos^2 \theta_{u,i} + \cos^2 \theta_{l,i}, \quad \forall i$$
$$\text{subject to } \theta_{lu,i} = \pi/2,$$

where $\theta_{lu,i}$ is one of the p principal angle between P_u and P_l and is equal to $\pi/2$ as implied by the constraint $P_u P_l = 0$. Roughly speaking, to maximize the cosine's is to minimize the θ's, and the minimizing solution must be found on the plane defined by the ith bases of P_u and P_l. Any solution on this plane must satisfy $\theta_{u,i} + \theta_{l,i} = \pi/2$, and in turn achieves the maximum as $\max \cos^2 \theta_{u,i} + \cos^2(\pi/2 - \theta_{u,i}) = 1$. Consequently, $\max \operatorname{tr} P(P_u + P_l) = p$ and so for $p < n/2$, $\|P - P_o\|^2 \geq 1/2p$, meaning that there exists $P \in \mathrm{Gr}(n, p)$ outside $\mathcal{B}_{r_o}(P_o)$. ∎

The first part of Prop. 20 admits the application of Prop. 15, as $\mathrm{Gr}(n, p)$ is analytic. Therefore we can conclude that $\mathcal{C}_{\mathrm{Gr}}$ is an asymptotically stable polarized equilibrium. The ball argument in Prop. 14 only applies to Grassmannian polarization with $\mathcal{C}_{\mathrm{Gr}}$ for $p = n/2$, nevertheless, it is still true that $\mathcal{C}_{\mathrm{Gr}}$ is a stable polarized equilibrium of (13) for general (n, p) pairs.

Proposition 21: For system (13), the global minimum of V_- is $-p \sum_{\{i,j\}\in\mathcal{E}_-} a_{ij}$ and is achieved only by $\mathcal{C}_{\mathrm{Gr}}$.

Proof: For all $\{i, j\} \in \mathcal{E}_-$,

$$\tfrac{1}{2}\|P_j - P_i\|^2 = p - \operatorname{tr} P_i P_j = p - \sum_{m=1}^{p} \cos^2 \theta_m \leq p.$$

The upper bound is achieved iff all principal angles are $\pi/2$ iff $P_j' P_i = 0$. For all $\{i, j\} \in \mathcal{E}_+$, the usual zero lower bound applies and is achieved iff $P_i = P_j$. Therefore $V_- \geq -p \sum_{\{i,j\}\in\mathcal{E}_-} a_{ij}$ is the global minimum achieved only by $\mathcal{C}_{\mathrm{Gr}}$. ∎

Remark 22: In our previous work on hypersurface polarization [11, Prop. 21], we showed the loss of stability when the two pimples are entirely outside $\mathcal{B}_{r_o}(x_o)$ except for the bottoms x_u and x_l. The proof relies on the existence of a perturbation $\tilde{x}_u \neq x_u$ and $\tilde{x}_l \neq x_l$ such that $\|\tilde{x}_u - \tilde{x}_l\| > 2r_o$ and that $\tilde{x}_u - \tilde{x}_l$ "passes through" x_o. Or to put the latter more formally, there exists $\lambda \in \mathbb{R}$ such that $\lambda \tilde{x}_u + (1 - \lambda)\tilde{x}_l = x_o$. In contrast, Prop. 21 implies that on $\mathrm{Gr}(n, p)$ for $p < n/2$, there cannot exist $\tilde{P}_u, \tilde{P}_l \notin \mathcal{B}_{r_o}(P_o)$ such that $\tilde{P}_u - \tilde{P}_l$ passes through P_o and its norm greater than $2r_o$. In fact, even the first requirement alone cannot be met. As noted, $P_o \in \mathrm{Gr}(n, 2p)$. For $\tilde{P}_u$ to be outside $\mathcal{B}_{r_o}(P_o)$, the subspace it represents must contain at least one basis outside the plane defined by every pair of bases from the $2p$ bases representing P_o (see the arguments following the second maximization problem in the proof of Prop. 20). Consequently, $\operatorname{rank}(P_o - \lambda \tilde{P}_u) > p$ and there is no such $(1 - \lambda)\tilde{P}_l = P_o - \lambda \tilde{P}_u$ for $\tilde{P}_l \in \mathrm{Gr}(n, p)$.

References

[1] J. Markdahl, J. Thunberg, and J. Gonçalves, "Almost global consensus on the n-sphere," *IEEE Transactions on Automatic Control*, vol. 63, no. 6, pp. 1664–1675, 2018.

[2] A. Sarlette, S. Bonnabel, and R. Sepulchre, "Coordinated motion design on Lie groups," *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1047–1058, 2010.

[3] L. DeVille, "Synchronization and stability for quantum kuramoto," *Journal of Statistical Physics*, vol. 174, no. 1, pp. 160–187, 2019.

[4] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[5] A. Bizyaeva, A. Franci, and N. E. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," *IEEE Transactions on Automatic Control*, 2022.

[6] D. A. Hormuth II, M. Farhat, C. Christenson, B. Curl, C. C. Quarles, C. Chung, and T. E. Yankeelov, "Opportunities for improving brain cancer treatment outcomes through imaging-based mathematical modeling of the delivery of radiotherapy and immunotherapy," *Advanced Drug Delivery Reviews*, p. 114367, 2022.

[7] R. Tron, B. Afsari, and R. Vidal, "Riemannian consensus for manifolds with bounded curvature," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 921–934, 2012.

[8] J. Markdahl, "Synchronization on Riemannian manifolds: Multiply connected implies multistable," *IEEE Transactions on Automatic Control*, vol. 66, no. 9, pp. 4311–4318, 2021.

[9] H. Hong and S. H. Strogatz, "Kuramoto model of coupled oscillators with positive and negative coupling parameters: An example of conformist and contrarian oscillators," *Physical Review Letters*, vol. 106, no. 5, p. 054102, 2011.

[10] S.-Y. Ha, D. Kim, J. Lee, and S. E. Noh, "Emergence of bicluster aggregation for the swarm sphere model with attractive-repulsive couplings," *SIAM Journal on Applied Dynamical Systems*, vol. 19, no. 2, pp. 1225–1270, 2020.

[11] L. Mi, J. Gonçalves, and J. Markdahl, "Asymptotically stable polarization of multi-agent gradient flows over manifolds," *arXiv:2301.04877*, 2023.

[12] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2012.

[13] U. Helmke and J. B. Moore, *Optimization and dynamical systems*. Springer, 1996.

[14] K. Kurdyka, T. Mostowski, and A. Parusinski, "Proof of the gradient conjecture of R. Thom," *Annals of Mathematics*, pp. 763–792, 2000.

[15] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303–353, 1998.

[16] J. Markdahl, J. Thunberg, and J. Goncalves, "High-dimensional Kuramoto models on Stiefel manifolds synchronize complex networks almost globally," *Automatica*, vol. 113, p. 108736, 2020.

[17] A. Sarlette and R. Sepulchre, "Consensus optimization on manifolds," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 56–76, 2009.

[18] T. Bendokat, R. Zimmermann, and P.-A. Absil, "A Grassmann manifold handbook: Basic geometry and computational aspects," *arXiv:2011.13699*, 2020.

[19] U. Helmke, K. Hüper, and J. Trumpf, "Newton's method on Grassmann manifolds," *arXiv:0709.2205*, 2007.

[20] J. H. Conway, R. H. Hardin, and N. J. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, 1996.