# Towards Almost Global Synchronization on the Stiefel Manifold

Johan Markdahl, Johan Thunberg, and Jorge Gonçalves

*Abstract*— The Kuramoto model evolves on the circle, *i.e.*, the 1-sphere $S^1$. A graph $\mathcal{G}$ is referred to as $S^1$-synchronizing if the Kuramoto model on $\mathcal{G}$ synchronizes almost globally. This paper generalizes the Kuramoto model and the concept of synchronizing graphs to the Stiefel manifold $St(p, n)$. Previous work on generalizations of the Kuramoto model have largely been influenced by results and techniques that pertain to the original model. It was recently shown that all connected graphs are $S^n$-synchronizing for all $n \geq 2$. However, that does not hold for $n = 1$. Previous results on generalized models may thus have been overly conservative. The $n$-sphere is a special case of the Stiefel manifold, namely $St(1, n + 1)$. As such, it is natural to ask for the extent to which the results on $\mathcal{S}^n$ can be extended to the Stiefel manifold. This paper shows that all connected graphs are $St(p, n)$-synchronizing provided the pair $(p, n)$ satisfies $p \leq \frac{2n}{3} - 1$.

## I. INTRODUCTION

Scalability is a key advantage of distributed approaches to feedback control of multi-agent systems [1]. It is achieved by feedback laws with linear computational complexity that can be executed when interactions between most agents is indirect, *e.g.*, when the communications topology is given by a path or cycle graph. An often overlooked aspect of scalability is control performance at a large distance from nominal operating conditions. Consider the case of $N$ homogeneous agents, the state of each belonging to a compact manifold $\mathcal{M}$. Many results in the literature concerns the case when all agents belong to some convex subset of $\mathcal{M}$ [2]–[4]. A typical result is the guaranteed convergence of a synchronization algorithm [5]–[7]. If the initial states follow a uniform distribution on $\mathcal{M}$, then the probability that all agents are contained in such a convex set decreases exponentially with $N$. This is a case of poor scaling. Ideally, the probability of convergence should be $1$ independently of $N$. Such performance is achieved by almost globally convergent algorithms. This paper concerns the problem of establishing almost global convergence of a class of continuous time consensus protocols for multi-agent system that evolve on the compact, real Stiefel manifold $St(p, n)$. Each of the systems under consideration is an intrinsic gradient descent flow of a basic, quadratic potential function.

A graph $\mathcal{G}$ is $S^1$-synchronizing if, roughly speaking, the consensus manifold is an almost globally stable equilibrium manifold of a Kuramoto-like oscillator model with topology $\mathcal{G}$. Examples of such graphs include the complete graph, acyclic graphs, and sufficiently dense graphs [8]. The survey

[9] ponders the question of which combinations of graphs, manifolds, and consensus protocols lead to almost global synchrony. A key result is that not all undirected graphs are $S^1$-synchronizing, but there are protocols that yield other, almost globally synchronizing closed-loop systems on $S^1$ [1], [10]. A generalization to $SO(3)$ has been established [11]. Some protocols also converge in the case of quasi-strongly connected digraphs on the Stiefel manifold [12]. The control strategy requires that auxiliary variables are communicated between agents. As such, it is intended for use in engineering systems rather than as a model of self-organizing systems of coupled oscillators that are observed in nature.

Consensus protocols for agents whose dynamics evolve over linear spaces are well understood [13]. Moving forward, it is natural to look at homogeneous spaces [14]. All points on a homogeneous space are similar, making them well-suited for hosting multi-agent systems where relative and aggregate quantities are the main focus. This paper studies a consensus protocol on a homogeneous space: a generalization of the Kuramoto model to the compact, real Stiefel manifold. Special cases of $St(p, n)$ include the $(n − 1)$-sphere when $p = 1$, the special orthogonal group when $p = n − 1$, and the orthogonal group when $p = n$. As such, these results are of relevance in a number of applications including cooperative reduced and full attitude control in the cases of $(p, n) = (1, 3)$ [15], [16] and $(p, n) = (2, 3)$ [17] respectively. The Stiefel manifold for $p \in \{2, \ldots, n − 2\}$ is typically not used to model physical systems, which is the main area of application for continuous time control systems (although application of continuous time dynamical systems theory to numerical linear algebra problems do exists [18]).

The authors of this paper have showed that all connected graphs are $S^n$-synchronizing for all $n \in \mathbb{N}\backslash\{1\}$ [19]. This result is of interest since it is unexpected; it could not have been interpolated from the previous research concerning special cases such as $S^1$-synchronizing graphs for the Kuramoto model and more general findings concerning almost global synchrony on $SO(3)$ [11]. This paper generalizes those results to the Stiefel manifold where we show that all connected graphs are $St(p, n)$-synchronizing provided the pair $(p, n)$ satisfies $p \leq \frac{2n}{3} − 1$. This inequality is imposed for technical reasons and is unlikely to have any interesting interpretation. Previous research on almost global synchronization over homogeneous manifolds departs from the negative result that not all undirected graphs are $S^1$-synchronizing [1], [14]. We show that such pessimism may be unfounded since the circle is actually a pathological case, providing hope that results akin to those of this paper also apply to other Riemannian manifolds.

J. Markdahl and J. Gonçalves are with the Luxembourg Centre for Systems Biomedicine, University of Luxembourg (UL). J. Thunberg is with the School of Information Technology, Halmstad University.

J. Markdahl is the corresponding author: markdahl@kth.se.

## II. PROBLEM FORMULATION

### A. Preliminaries

The following notation is used in this paper. The Euclidean inner product of $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times p}$ is $g_e(\mathbf{X}, \mathbf{Y}) = \langle \mathbf{X}, \mathbf{Y} \rangle = \operatorname{tr} \mathbf{X}^\top \mathbf{Y}$. The norm of $\mathbf{X}$ is given by $\|\mathbf{X}\| = \langle \mathbf{X}, \mathbf{X} \rangle^{\frac{1}{2}}$, i.e., the Frobenius norm. The compact real Stiefel manifold is considered an embedded matrix manifold in $\mathbb{R}^{n \times p}$ [20],

$$\mathsf{St}(p, n) = \{\mathbf{S} \in \mathbb{R}^{n \times p} \mid \mathbf{S}^\top \mathbf{S} = \mathbf{I}\}.$$

The pair $(\mathsf{St}(p, n), g_e)$ forms a smooth Riemannian manifold. An important Stiefel manifold is $\mathsf{St}(1, n + 1)$, i.e., the $n$-sphere which we denote $\mathsf{S}^n$.

Define the projections skew : $\mathbb{R}^{n \times n} \to \mathsf{so}(n) : \mathbf{X} \mapsto \frac{1}{2}(\mathbf{X} - \mathbf{X}^\top)$ and sym : $\mathbb{R}^{n \times n} \to \mathsf{so}(n)^\perp : \mathbf{X} \mapsto \frac{1}{2}(\mathbf{X} + \mathbf{X}^\top)$. The tangent space of $\mathsf{St}(p, n)$ at $\mathbf{X}$ is given by

$$\mathsf{T_S St}(p, n) = \{\boldsymbol{\Delta} \in \mathbb{R}^{n \times p} \mid \operatorname{sym} \mathbf{S}^\top \boldsymbol{\Delta} = \mathbf{0}\}.$$

The projection on the tangent space, $\Pi : \mathbb{R}^{n \times p} \times \mathsf{St}(p, n) \to \mathsf{T_S St}(p, n)$, is given by

$$\Pi(\mathbf{X}, \mathbf{S}) = \mathbf{S} \operatorname{skew} \mathbf{S}^\top \mathbf{X} + (\mathbf{I}_n - \mathbf{S}\mathbf{S}^\top)\mathbf{X}.$$

From a computational perspective, it is sometimes preferable to use the equivalent expression

$$\Pi(\mathbf{X}, \mathbf{S}) = \mathbf{X} - \mathbf{S} \operatorname{sym} \mathbf{S}^\top \mathbf{X}.$$

The gradient on $\mathsf{St}(p, n)$ (in terms of the Euclidean inner product $g_e$) of a function $U : \mathsf{St}(p, n) \to \mathbb{R}$ is given by

$$\square U = \Pi \nabla V,$$

where $V$ is any smooth extension of $U$ on $\mathbb{R}^{n \times p}$, and $\nabla$ denotes the gradient in Euclidean space.

Each element $i \in \mathcal{V}$ also denotes an agent. Items associated with a specific agent $i$ carry the subindex $i$; we let $\mathbf{S}_i \in \mathsf{St}(p, n)$ denote the state of an agent, $\Pi_i$ the projection onto the tangent space $\mathbf{S}_i$, $\square_i U$ the gradient of $U$ with respect to $\mathbf{S}_i$, etc.

### B. Distributed Control Design

The consensus submanifold $\mathcal{C}$ of an analytic Riemannian manifold $(\mathcal{M}, g)$ is the set of equilibria

$$\mathcal{C} = \{(x_i)_{i=1}^N \in \mathcal{M}^N \mid x_i = x_j, \forall \{i, j\} \in \mathcal{E}\}.$$

The consensus set is a manifold, in fact it is diffeomorphic to $\mathcal{M}$ using the map $(x_i)_{i=1}^N \mapsto x_1$.

Given a graph $(\mathcal{V}, \mathcal{E})$, define the potential function $U : \mathcal{M}^N \to \mathbb{R}$ by

$$U = \sum_{e \in \mathcal{E}} f_{ij}(d_g(x_i, x_j)),$$

where $f_{ij} : \mathbb{R} \to [0, \infty)$, $d_g$ is the geodesic distance on $\mathcal{M}$ in terms of $g$, and $e$ is on the form $\{i, j\}$. The consensus seeking system on $\mathcal{M}$ obtained from $U$ is the gradient descent flow

$$(\dot{x}_i)_{i=1}^N = (-\square_i U)_{i=1}^N, \tag{1}$$

where $x_i(0) \in \mathcal{M}$ for all $i \in \mathcal{V}$.

Agent $i$ does not have access to $U$, but can calculate

$$U_i = \tfrac{1}{2} \sum_{j \in \mathcal{N}_i} f_{ij}(d_g(x_i, x_j))$$

at its current position. Symmetry of $d_g$ gives $U = \sum_{i \in \mathcal{V}} U_i$ whereby it follows that $\square_i U_i = \square_i U$. From a control design perspective, we can assume that the dynamics of each agent take the form $\dot{x}_i = u_i$ with $u_i \in \mathsf{T}_i \mathcal{M}$. Since agent $i$ can evaluate $U_i$ at its current position, it is reasonable to assume that it can also calculate $u_i = -\square_i U_i$.

### C. $\mathsf{St}(p, n)$-Synchronizing Graphs

*Definition 1:* A graph $\mathcal{G}$ is $\mathcal{M}$-synchronizing if all minimizers of $U$ belong to $\mathcal{C}$.

The property of being $\mathcal{M}$-synchronizing, which we have adopted from [1], does not explicitly reference the specific function $U$ under consideration. For the purpose of this paper, we limit consideration to potential functions $U$ on matrix manifolds $\mathcal{M} \subset \mathbb{R}^{n \times m}$ of the following form

$$U = \tfrac{1}{2} \sum_{e \in \mathcal{E}} a_{ij} \|\mathbf{X}_i - \mathbf{X}_j\|^2, \tag{2}$$

where $\mathbf{X}_i \in \mathbb{R}^{n \times m}$, and the constants $a_{ij}$ are strictly positive and symmetric, i.e., $a_{ji} = a_{ij}$. On the Stiefel manifold $\mathsf{St}(p, n) \subset \mathbb{R}^{n \times p}$, this reduces to

$$U = \sum_{e \in \mathcal{E}} a_{ij}(p - \langle \mathbf{S}_i, \mathbf{S}_j \rangle), \tag{3}$$

since $\|\mathbf{S}_i\|^2 = p$ for all $i \in \mathcal{V}$.

*Definition 2:* An equilibrium manifold $\mathcal{Q}$ of a dynamical system $\Sigma$ on an analytic Riemannian manifold $(\mathcal{M}, g)$ is referred to as almost globally asymptotically stable (AGAS) if it is stable and the flow $\Phi(t, x_0)$ of $\Sigma$ satisfies $\lim_{t \to \infty} d_g(\mathcal{Q}, \Phi(t, x_0)) = 0$ for all $x_0 \in \mathcal{M} \setminus \mathcal{N}$, where $\mathcal{N} \subset \mathcal{M}$ has measure zero on $\mathcal{M}$.

It is not immediately clear that $\mathcal{G}$ being $\mathcal{M}$-synchronizing implies that $\mathcal{C}$ is an AGAS equilibrium of (1). Since (1) is a gradient descent of $U$, it cannot converge to any maximum of $U$. Morover, any saddle point of $U$ is unstable. However, a set of saddle points may still have a region of attraction with positive measure, in which case $\mathcal{C}$ cannot be AGAS. In [19], we show that any connected graph is $\mathsf{S}^n$-synchronizing for the potential function $U = \sum_{\{i,j\} \in \mathcal{E}} a_{ij}(1 - \langle \mathbf{x}_i, \mathbf{x}_j \rangle)$. Moreover, we prove that $\mathcal{C}$ is AGAS. However, obtaining such results is not trivial. We limit the scope of this paper to characterizing $\mathsf{St}(p, n)$-synchronizing graphs. Sufficient conditions for $\mathcal{C}$ to be AGAS will be established in future work. To give the reader an idea of the current state of the art in terms of characterizing $\mathcal{M}$-synchronizing graphs, we provide the following examples.

*Example 3:* On $\mathbb{R}^n$, for $U$ given by the potential function (2) using the Euclidean vector norm, the system (1) becomes

$$\dot{\mathbf{x}}_i = \sum_{j \in \mathcal{N}_i} a_{ij}(\mathbf{x}_j - \mathbf{x}_i), \forall i \in \mathcal{V}.$$

The consensus manifold of this system is well-known to be globally asymptotically stable (which also implies AGAS),

although that fact is usually not expressed by saying that any connected graph is $\mathbb{R}^n$-synchronizing.

*Example 4:* On $\mathsf{S}^1$, the dynamics (1) can be expressed in polar coordinates $\vartheta_i \in \mathbb{R}$ for all $i \in \mathcal{V}$ as

$$\dot{\vartheta}_i = \sum_{j \in \mathcal{N}_i} a_{ij} \sin(\vartheta_j - \vartheta_i), \; \forall\, i \in \mathcal{V}. \tag{4}$$

For a complete graph $\mathcal{G}$, (4) is equivalent to the Kuramoto model in the case of homogeneous agents, *i.e.*, identical oscillator frequencies. The cycle graph

$$\mathcal{G}_N = (\mathcal{V}, \{\{i, j\} \subset \mathcal{V} \,|\, |i - j| = 1\} \cup \{\{1, N\}\})$$

is $\mathsf{S}^1$-synchronizing for $N \le 4$ but not for $N \ge 5$ [1]. The problem of characterizing all $\mathsf{S}^1$-synchronizing graphs is open [1], [8].

*Example 5:* On $\mathsf{S}^n$, the dynamics are given by

$$\dot{\mathbf{x}}_i = \sum_{j \in \mathcal{N}_i} a_{ij}\mathbf{x}_j - \Big\langle \sum_{j \in \mathcal{N}_i} a_{ij}\mathbf{x}_j, \mathbf{x}_i \Big\rangle \mathbf{x}_i, \; \forall\, i \in \mathcal{V}.$$

Any connected graph is $\mathsf{S}^n$-synchronizing for all $n \ge 2$ [19].

### D. Problem Statement

The aim of this paper is to identify instances of the Stiefel manifold that are $\mathsf{St}(p, n)$-synchronizing. We know that all connected graphs are $\mathsf{St}(1, n)$-synchronizing for $n \ge 3$ [19]. We also know that not all connected graphs are $\mathsf{S}^1$-synchronizing [1], nor $\mathsf{SO}(3)$-synchronizing [11]. Note that $\mathsf{S}^1 \simeq \mathsf{St}(1, 2) \simeq \mathsf{SO}(2)$ and $\mathsf{SO}(3) \simeq \mathsf{St}(2, 3)$. The question, to which we give a partial answer, is how the results of [19] extends to the cases of $2 \le p \le n - 2$.

## III. Main Result

Theorem 6 is the main result of this paper. First we outline the proof. The details are provided in Section III-A to III-E.

*Theorem 6:* Let the pair $(p, n)$ satisfy $p \le \frac{2n}{3} - 1$. All minimizers of the potential function $U$ given by with $a_{ij} = 1$ for all $e \in \mathcal{E}$ belong to the consensus manifold; *i.e.*, all connected graphs are $\mathsf{St}(p, n)$-synchronizing.

*Proof:* Let $q : \mathsf{St}(p, n)^N \times \mathsf{TSt}(p, n)^N \to \mathbb{R}$ denote the quadratic form obtained from the intrinsic Hessian of $U$ evaluated at a critical point. It can either be calculated from the Lagrange optimality conditions or by linearizing the dynamics at an equilibrium; we take the latter approach. The second-order necessary conditions for unconstrained optimization over Riemannian manifolds imply that $q$ is weakly positive at any minimum $\{\mathbf{S}_i\}_{i=1}^N$ of $U$. Our goal is to exclude minimality of all equilibria $(\mathbf{S}_i)_{i=1}^N \notin \mathcal{C}$ by finding $(\boldsymbol{\Delta}_i)_{i=1}^N \in \mathsf{X}_{i=1}^N \mathsf{T}_i \mathsf{St}(p, n)$ such that $q((\mathbf{S}_i)_{i=1}^N, (\boldsymbol{\Delta}_i)_{i=1}^N)) < 0$. It is clear that all elements of $\mathcal{C}$ are global minimizers since $U \ge 0$ and $U|_{\mathcal{C}} = 0$.

Based on our previous work [19], we consider perturbations along the tangent space of $\mathcal{C}$. We do not need to determine an exact expression for the desired perturbation, it suffices to prove that it exists. Showing that $q$ is negative amounts to solving a nonconvex constrained optimization problem. This is done using the Lagrange optimality conditions. Here, we introduce the inequality $p \le \frac{2n}{3} - 1$

to fix a variable in the optimization problem. For this case, we can show that there is a perturbation such that $\mathcal{C}$ maximizes $q$ with an objective value of zero. Any other equilibrium configuration gives a strictly negative objective value. Throughout all these steps, we never utilize any particular property of the graph topology besides connectedness. Hence the final result applies to any connected graph. It follows that all connected graphs are $\mathsf{St}(p, n)$-synchronizing. ∎

### A. Critical Points

Let $V : (\mathbb{R}^{n \times p})^N \to [0, \infty)$ be the smooth extension of $U$ obtained by relaxing the requirement $(\mathbf{S}_i)_{i=1}^N \in \mathsf{St}(p, n)^N$ for all $i \in \mathcal{V}$. Then

$$\dot{\mathbf{S}}_i = -\Box_i U = -\Pi_i \nabla_i V = \Pi_i \sum_{j \in \mathcal{N}_i} \mathbf{S}_j$$

$$= \mathbf{S}_i \operatorname{skew}\Big(\mathbf{S}_i^\top \sum_{j \in \mathcal{N}_i} \mathbf{S}_j\Big) + (\mathbf{I}_n - \mathbf{S}_i \mathbf{S}_i^\top) \sum_{j \in \mathcal{N}_i} \mathbf{S}_j. \tag{5}$$

*Remark 7:* The system of homogeneous state-space oscillators on $\mathsf{St}(p, n)$ given by

$$\dot{\mathbf{S}}_i = \boldsymbol{\Omega}\mathbf{S}_i + \mathbf{S}_i\boldsymbol{\Xi} - \Box_i U, \tag{6}$$

where $\boldsymbol{\Omega} \in \mathsf{so}(n)$ and $\boldsymbol{\Xi} \in \mathsf{so}(p)$ can be reduced to (5) by a change of variables. To verify this, let $\mathbf{R} = \exp(-t\boldsymbol{\Omega}) \in \mathsf{SO}(n)$, $\mathbf{Q} = \exp(-t\boldsymbol{\Xi}) \in \mathsf{SO}(p)$, form $\mathbf{X}_i = \mathbf{R}\mathbf{S}_i\mathbf{Q} \in \mathsf{St}(p, n)$, and calculate

$$\dot{\mathbf{X}}_i = -\boldsymbol{\Omega}\mathbf{R}\mathbf{S}_i\mathbf{Q} + \mathbf{Q}\dot{\mathbf{S}}_i\mathbf{R} - \mathbf{R}\mathbf{S}_i\mathbf{Q}\boldsymbol{\Xi}$$
$$= \mathbf{X}_i \operatorname{skew}\Big(\mathbf{X}_i^\top \sum_{j \in \mathcal{N}_i} \mathbf{X}_j\Big) + (\mathbf{I}_n - \mathbf{X}_i\mathbf{X}_i^\top) \sum_{j \in \mathcal{N}_i} \mathbf{X}_j,$$

where we used that $[\boldsymbol{\Omega}, \mathbf{R}] = \mathbf{0}$, $[\boldsymbol{\Xi}, \mathbf{Q}] = \mathbf{0}$. Put $(p, n) = (1, 2)$ and let $\mathcal{G}$ be the complete graph to obtain the Kuramoto model of a system of homogeneous oscillators from (6).

The critical points of $U$ are the equilibria of (5). At an equilibrium,

$$\mathbf{0} = \mathbf{S}_i \operatorname{skew}\Big(\mathbf{S}_i^\top \sum_{j \in \mathcal{N}_i} \mathbf{S}_j\Big) + (\mathbf{I}_n - \mathbf{S}_i \mathbf{S}_i^\top) \sum_{j \in \mathcal{N}_i} \mathbf{S}_j.$$

Since these two terms are orthogonal, we get

$$\operatorname{skew} \mathbf{S}_i^\top \sum_{j \in \mathcal{N}_i} \mathbf{S}_j = \mathbf{0}, \; (\mathbf{I}_n - \mathbf{S}_i \mathbf{S}_i^\top) \sum_{j \in \mathcal{N}_i} \mathbf{S}_j = \mathbf{0}. \tag{7}$$

Assume (7) holds. Denote $\mathbf{V}_i = \sum_{j \in \mathcal{N}_i} \mathbf{S}_j$. Since $\mathbf{V}_i = \mathbf{S}_i \mathbf{S}_i^\top \mathbf{V}_i$, it follows that $\mathbf{V}_i \in \operatorname{Im} \mathbf{S}_i$. Hence $\mathbf{V}_i = \mathbf{S}_i \mathbf{P}_i$ for some $\mathbf{P}_i \in \mathbb{R}^{p \times p}$. Moreover, since $\operatorname{skew} \mathbf{S}_i^\top \mathbf{V}_i = \operatorname{skew} \mathbf{P}_i = \mathbf{0}$, we find that $\mathbf{P}_i$ is symmetric.

### B. The Intrinsic Hessian

Let $W_{i,st} : \mathbb{R}^{N \times n \times p} \to \mathbb{R}$ be the smooth extension of $(\Box_i U)_{st} = \langle \mathbf{e}_s, \Box_i U \mathbf{e}_t \rangle : (\mathsf{St}(p, n))^N \to \mathbb{R}$ obtained by relaxing the constraint $\mathbf{S}_i \in \mathsf{St}(p, n)$ to $\mathbf{S}_i \in \mathbb{R}^{n \times p}$ for all $i$. Using the rules governing derivatives of inner products with

respect to matrices [21], introducing $\mathbf{E}_{st} = \mathbf{e}_s \mathbf{e}_t^\top$, after a few calculations, we obtain

$$\nabla_k W_{i,st} = \begin{cases} -\Pi_i \mathbf{E}_{st} & \text{if } k \in \mathcal{N}_i, \\ \mathbf{E}_{st} \operatorname{skew}\left(\mathbf{S}_i^\top \sum_{j \in \mathcal{N}_i} \mathbf{S}_j\right) + \\ \sum_{j \in \mathcal{N}_i} \mathbf{S}_j \operatorname{sym}(\mathbf{S}_i^\top \mathbf{E}_{st}) + \\ \mathbf{E}_{st} \sum_{j \in \mathcal{N}_i} \mathbf{S}_j^\top \mathbf{S}_i & \text{if } k = i, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Evaluate at an equilibrium, where $\sum_{j \in \mathcal{N}_i} \mathbf{S}_j = \mathbf{S}_i \mathbf{P}_i$ and $\mathbf{P}_i \in \mathbb{R}^{p \times p}$ is symmetric by Section III-A, to find

$$\nabla_k W_{i,st} = \begin{cases} -\Pi_i \mathbf{E}_{st} & \text{if } k \in \mathcal{N}_i, \\ \mathbf{S}_i \mathbf{P}_i \operatorname{sym}(\mathbf{S}_i^\top \mathbf{E}_{st}) + \mathbf{E}_{st} \mathbf{P}_i & \text{if } k = i, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

The intrinsic Hessian is a $(N \times n \times p)^2$-tensor consisting of $N^2 np$ blocks $\mathbf{H}_{ki,st} \in \mathbb{R}^{n \times p}$, which are obtained by projecting the extrinsic Hesssian on the tangent space of $\mathbf{S}_k$

$$\mathbf{H}_{ki,st} = \Box_k(\Box_i U)_{st} = \Pi_k \nabla_k W_{i,st}$$
$$= \Pi_k \nabla_k (\Pi_i \nabla_i V)_{st}.$$

### C. The Quadratic Form

Consider the quadratic form $q : \mathsf{St}(p,n)^N \times \mathsf{TSt}(p,n)^N \to \mathbb{R}$ obtained from the intrinsic Hessian evaluated at an equilibrium for some perturbation $(\boldsymbol{\Delta}_i)_{i=1}^N \in \mathsf{T}\mathcal{C}$,

$$q = \sum_{i=1}^N \sum_{k=1}^N \langle \boldsymbol{\Delta}_i, [\langle \boldsymbol{\Delta}_k, \Box_k(\Box_i U)_{st}\rangle]\rangle$$
$$= \sum_{i=1}^N \sum_{k=1}^N \langle \Pi_i \boldsymbol{\Delta}, [\langle \Pi_k \boldsymbol{\Delta}, \Pi_k \nabla_k W_{i,st}\rangle]\rangle,$$

where $\boldsymbol{\Delta} \in \mathbb{R}^{n \times p}$.

Note that $\langle \Pi_k \mathbf{X}, \Pi_k \mathbf{Y}\rangle = \langle \Pi_k \mathbf{X}, \mathbf{Y}\rangle$. The quadratic form is hence

$$q = \sum_{i=1}^N \sum_{k=1}^N \langle \Pi_i \boldsymbol{\Delta}, [\langle \Pi_k \boldsymbol{\Delta}, \nabla_k W_{i,st}\rangle]\rangle.$$

Denote $p_{ki,st} = \langle \Pi_k \boldsymbol{\Delta}, \nabla_k W_{i,st}\rangle$. Then

$$p_{ki,st} = \begin{cases} \langle \Pi_k \boldsymbol{\Delta}, -\Pi_i \mathbf{E}_{st}\rangle \\ \langle \Pi_i \boldsymbol{\Delta}, \mathbf{S}_i \mathbf{P}_i \operatorname{sym}(\mathbf{S}_i^\top \mathbf{E}_{st}) + \mathbf{E}_{st} \mathbf{P}_i\rangle \\ \mathbf{0} \end{cases}$$

for the cases of $k \in \mathcal{N}_i$, $k = i$, and $k \notin \mathcal{N}_i \cup \{i\}$ respectively. Denote $p_{ki} = [p_{st}]$ and calculate

$$p_{ki} = \begin{cases} -\Pi_i \Pi_k \boldsymbol{\Delta} & \text{if } k \in \mathcal{N}_i, \\ \mathbf{S}_i \operatorname{sym} \mathbf{V}_i^\top \Pi_i \boldsymbol{\Delta} + \Pi_i(\boldsymbol{\Delta})\mathbf{P}_i & \text{if } k = i, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

To see this, consider each case separately. For $k \in \mathcal{N}_i$,

$$p_{ki,st} = \langle (\mathbf{I}_n - \Pi_i + \Pi_i)\Pi_k \boldsymbol{\Delta}, -\Pi_i \mathbf{E}_{st}\rangle$$
$$= -\langle \Pi_i \Pi_k \boldsymbol{\Delta}, \mathbf{E}_{st}\rangle = -(\Pi_i \Pi_k \boldsymbol{\Delta})_{st},$$

whereby $p_{ki} = -\Pi_i \Pi_k \boldsymbol{\Delta}$. For the case of $k = i$,

$$p_{ii,st} = \langle \Pi_i \boldsymbol{\Delta}, \mathbf{S}_i \mathbf{P}_i \operatorname{sym}(\mathbf{S}_i^\top \mathbf{E}_{st}) + \mathbf{E}_{st} \mathbf{P}_i\rangle$$
$$= \tfrac{1}{2}(\mathbf{S}_i \mathbf{V}_i^\top \Pi_i \boldsymbol{\Delta})_{st} + \tfrac{1}{2}(\mathbf{S}_i(\Pi_i \boldsymbol{\Delta})^\top \mathbf{V}_i)_{st} + (\Pi_i(\boldsymbol{\Delta})\mathbf{P}_i)_{st},$$

whereby $p_{ii} = \mathbf{S}_i \operatorname{sym} \mathbf{V}_i^\top \Pi_i \boldsymbol{\Delta} + \Pi_i(\boldsymbol{\Delta})\mathbf{P}_i$.

This gives us the quadratic form

$$q = \sum_{i=1}^N \sum_{k=1}^N \langle \Pi_i \boldsymbol{\Delta}, p_{ki}\rangle$$
$$= \sum_{e \in \mathcal{E}} \langle \Pi_i \boldsymbol{\Delta}, p_{ki}\rangle + \langle \Pi_k \boldsymbol{\Delta}, p_{ik}\rangle + \sum_{i \in \mathcal{V}} \langle \Pi_i \boldsymbol{\Delta}, p_{ii}\rangle$$

For ease of notation, let $q = 2 \sum_{e \in \mathcal{E}} q_{ik} + \sum_{i \in \mathcal{V}} q_i$, where

$$q_{ik} = \langle \Pi_i \boldsymbol{\Delta}, p_{ki}\rangle = -\langle \Pi_i \boldsymbol{\Delta}, \Pi_k \boldsymbol{\Delta}\rangle = q_{ki},$$
$$q_i = \langle \Pi_i \boldsymbol{\Delta}, p_{ii}\rangle.$$

Calculate

$$q_{ik} = -\langle \Pi_i \boldsymbol{\Delta}, \Pi_k \boldsymbol{\Delta}\rangle$$
$$= \operatorname{tr}(-\boldsymbol{\Delta}^\top \boldsymbol{\Delta} + \tfrac{1}{2}\boldsymbol{\Delta}^\top \mathbf{S}_i \mathbf{S}_i^\top \boldsymbol{\Delta} + \tfrac{1}{2}\mathbf{S}_i^\top \boldsymbol{\Delta} \mathbf{S}_i^\top \boldsymbol{\Delta} +$$
$$\tfrac{1}{2}\boldsymbol{\Delta}^\top \mathbf{S}_k \mathbf{S}_k^\top \boldsymbol{\Delta} + \tfrac{1}{2}\boldsymbol{\Delta}^\top \mathbf{S}_k \boldsymbol{\Delta}^\top \mathbf{S}_k -$$
$$\tfrac{1}{4}\boldsymbol{\Delta}^\top \mathbf{S}_i \mathbf{S}_i^\top \mathbf{S}_k \mathbf{S}_k^\top \boldsymbol{\Delta} - \tfrac{1}{4}\boldsymbol{\Delta}^\top \mathbf{S}_i \mathbf{S}_i^\top \mathbf{S}_k \boldsymbol{\Delta}^\top \mathbf{S}_k -$$
$$\tfrac{1}{4}\mathbf{S}_i^\top \boldsymbol{\Delta} \mathbf{S}_i^\top \mathbf{S}_k \mathbf{S}_k^\top \boldsymbol{\Delta} - \tfrac{1}{4}\mathbf{S}_i^\top \boldsymbol{\Delta} \mathbf{S}_i^\top \mathbf{S}_k \boldsymbol{\Delta}^\top \mathbf{S}_k).$$

Use the identity $\operatorname{tr} \mathbf{ABCD} = \langle \operatorname{vec} \mathbf{A}^\top, (\mathbf{D}^\top \otimes \mathbf{B}) \operatorname{vec} \mathbf{C}\rangle$ [22] and the notation $\mathbf{d}_1 = \operatorname{vec} \boldsymbol{\Delta}$, $\mathbf{d}_2 = \operatorname{vec} \boldsymbol{\Delta}^\top$ to write $q_{ik} = \langle \mathbf{d}, \mathbf{Q}_{ik} \mathbf{d}\rangle$, where $\mathbf{Q}_{ik}$ is given in Table I and $\mathbf{d} = [\mathbf{d}_1^\top \, \mathbf{d}_2^\top]^\top$.

Furthermore,

$$q_i = \langle \Pi_i \boldsymbol{\Delta}, \mathbf{S}_i \operatorname{sym} \mathbf{V}_i^\top \Pi_i \boldsymbol{\Delta} + \Pi_i(\boldsymbol{\Delta})\mathbf{P}_i\rangle = \langle \mathbf{d}, \mathbf{Q}_i \mathbf{d}\rangle,$$

where

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{P}_i \otimes \mathbf{I}_n - \tfrac{3}{4}\mathbf{P}_i \otimes \mathbf{S}_i \mathbf{S}_i^\top & \mathbf{0} \\ -\tfrac{1}{2}\mathbf{S}_i \mathbf{P}_i \otimes \mathbf{S}_i^\top & \tfrac{1}{4}\mathbf{S}_i \mathbf{P}_i \mathbf{S}_i^\top \otimes \mathbf{I}_p \end{bmatrix}.$$

There is a constant permutation matrix $\mathbf{K} \in \mathsf{O}(np)$ such that $\operatorname{vec} \boldsymbol{\Delta}^\top = \mathbf{K} \operatorname{vec} \boldsymbol{\Delta}$ for all $\operatorname{vec} \boldsymbol{\Delta} \in \mathbb{R}^{np}$ [22]. Hence

$$\mathbf{d} = \begin{bmatrix} \operatorname{vec} \boldsymbol{\Delta} \\ \operatorname{vec} \boldsymbol{\Delta}^\top \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{np} \\ \mathbf{K} \end{bmatrix} \operatorname{vec} \boldsymbol{\Delta} = \begin{bmatrix} \mathbf{I}_{np} \\ \mathbf{K} \end{bmatrix} \mathbf{d}_1.$$

The quadratic form $q$ satisfies $q = \langle \mathbf{d}_1, \mathbf{M} \mathbf{d}_1\rangle$, where

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_{np} & \mathbf{K}^\top \end{bmatrix} \mathbf{Q} \begin{bmatrix} \mathbf{I}_{np} \\ \mathbf{K} \end{bmatrix}, \quad \mathbf{Q} = \sum_{i \in \mathcal{V}} \mathbf{Q}_i + \sum_{k \in \mathcal{N}_i} \mathbf{Q}_{ik}.$$

We wish to show that $q$ assumes positive values for some $\boldsymbol{\Delta} \in \mathbb{R}^{n \times p}$ at all equilibria $(\mathbf{S}_i)_{i=1}^N \notin \mathcal{C}$. If $\operatorname{tr} \mathbf{M}$ is positive, then the symmetric part of $\mathbf{M}$ has at least one positive eigenvalue. Hence calculate

$$\operatorname{tr} \mathbf{M} = \operatorname{tr} \begin{bmatrix} \mathbf{I}_{np} & \mathbf{K}^\top \end{bmatrix} \mathbf{Q} \begin{bmatrix} \mathbf{I}_{np} \\ \mathbf{K} \end{bmatrix} = \operatorname{tr}\left(\mathbf{Q} \begin{bmatrix} \mathbf{I}_{np} & \mathbf{K}^\top \\ \mathbf{K} & \mathbf{I}_{np} \end{bmatrix}\right)$$
$$= \operatorname{tr}\left(\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{np} & \mathbf{K}^\top \\ \mathbf{K} & \mathbf{I}_{np} \end{bmatrix}\right)$$
$$= \operatorname{tr}(\mathbf{A} + \mathbf{B}\mathbf{K} + \mathbf{C}\mathbf{K}^\top + \mathbf{D}).$$

$$\mathbf{Q}_{ik} = \begin{bmatrix} -\mathbf{I}_{np} + \frac{1}{2}\mathbf{I}_p \otimes (\mathbf{S}_i\mathbf{S}_i^\top + \mathbf{S}_k\mathbf{S}_k^\top) - \frac{1}{4}\mathbf{I}_p \otimes \mathbf{S}_i\mathbf{S}_i^\top\mathbf{S}_k\mathbf{S}_k^\top & \frac{1}{2}\mathbf{S}_k^\top \otimes \mathbf{S}_k - \frac{1}{4}\mathbf{S}_k^\top \otimes \mathbf{S}_i\mathbf{S}_i^\top\mathbf{S}_k \\ \frac{1}{2}\mathbf{S}_i \otimes \mathbf{S}_i^\top - \frac{1}{4}\mathbf{S}_i \otimes \mathbf{S}_i^\top\mathbf{S}_k\mathbf{S}_k^\top & -\frac{1}{4}\mathbf{S}_i\mathbf{S}_k^\top \otimes \mathbf{S}_i^\top\mathbf{S}_k \end{bmatrix}$$

**TABLE I.** The matrix $\mathbf{Q}_{ik}$.

Omitting the details of calculations, it holds that

$$\operatorname{tr}\mathbf{A} = 2\sum_{e\in\mathcal{E}}\left(n - \tfrac{3p}{4}\right)\langle\mathbf{S}_k, \mathbf{S}_i\rangle - (n-p)p - \tfrac{p}{4}\|\mathbf{S}_k^\top\mathbf{S}_i\|^2,$$

$$\operatorname{tr}\mathbf{D} = 2\sum_{e\in\mathcal{E}}\tfrac{p}{4}\langle\mathbf{S}_k, \mathbf{S}_i\rangle - \tfrac{1}{4}\langle\mathbf{S}_k, \mathbf{S}_i\rangle^2.$$

To deal with terms involving $\mathbf{K}$, we utilize that $\mathbf{K} = \sum_{a=1}^n\sum_{b=1}^p \mathbf{E}_{ab} \otimes \mathbf{E}_{ba}$, where the elemental matrix $\mathbf{E}_{ab} \in \mathbb{R}^{n\times p}$ is given by $\mathbf{E}_{ab} = \mathbf{e}_a \otimes \mathbf{e}_b$ for all $a \in \{1,\ldots,n\}$, $b \in \{1,\ldots,p\}$ [22]. After some calculations we obtain

$$\operatorname{tr}\mathbf{BK} = 2\sum_{e\in\mathcal{E}}\tfrac{p}{2} - \tfrac{1}{4}\|\mathbf{S}_k^\top\mathbf{S}_i\|^2,$$

$$\operatorname{tr}\mathbf{CK}^\top = 2\sum_{e\in\mathcal{E}} -\tfrac{1}{2}\langle\mathbf{S}_k, \mathbf{S}_i\rangle + \tfrac{p}{2} - \tfrac{1}{4}\|\mathbf{S}_i^\top\mathbf{S}_k\|^2.$$

Adding up all four terms gives

$$\tfrac{1}{2}\operatorname{tr}\mathbf{M} = \sum_{e\in\mathcal{E}}\left(n - \tfrac{p+1}{2}\right)\langle\mathbf{S}_k, \mathbf{S}_i\rangle - \tfrac{p+2}{4}\|\mathbf{S}_k^\top\mathbf{S}_i\|^2 -$$
$$\tfrac{1}{4}\langle\mathbf{S}_k, \mathbf{S}_i\rangle^2 + (1 - n + p)p. \quad (8)$$

At a consensus we get $\operatorname{tr}\mathbf{M}|_{\mathcal{C}} = 0$. This is expected since $U$ is constant over $\mathcal{C}$ and $\mathcal{C}$ is invariant under any perturbation that belongs to its tangent space. We also note that the result (8) is consistent with that of [19].

### D. Nonlinear Programming Problem

Having determined $q = \operatorname{tr}\mathbf{M}$ in (8), it remains to show that $\operatorname{tr}\mathbf{M}$ is strictly negative for any configuration $(\mathbf{S}_i)_{i=1}^N \notin \mathcal{C}$. To that end, consider

$$\tfrac{1}{2}\operatorname{tr}\mathbf{M} \le |\mathcal{E}|\max_{\mathbf{X},\mathbf{Y}} f(\mathbf{X}, \mathbf{Y}), \quad (9)$$

$$f(\mathbf{X}, \mathbf{Y}) = \left(n - \tfrac{p+1}{2}\right)\langle\mathbf{X}, \mathbf{Y}\rangle - \tfrac{p+2}{4}\|\mathbf{X}^\top\mathbf{Y}\|^2 -$$
$$\tfrac{1}{4}\langle\mathbf{X}, \mathbf{Y}\rangle^2 + (1 - n + p)p, \quad (10)$$

where $f : \operatorname{St}(p,n) \times \operatorname{St}(p,n) \to \mathbb{R}$. Hence, if we can show that the image of $f$ is negative for all $\mathbf{X} \ne \mathbf{Y}$, then we are done. Note that the inequality is sharp in the case of two agents and that $f(\mathbf{X}, \mathbf{X}) = 0$ since this corresponds to consensus in a system of two agents.

Consider the nonlinear, non-convex optimization problem

$$\max f(\mathbf{X}, \mathbf{Y}) \text{ s.t. } \mathbf{X}, \mathbf{Y} \in \operatorname{St}(p,n), \quad (11)$$

where $f(\mathbf{X}, \mathbf{Y})$ is given by (10). It follows from (9) that (11) is a relaxation of the problem $\max\operatorname{tr}\mathbf{M}$ such that $(\mathbf{S}_i)_{i=1}^N \in \operatorname{St}(p,n)^N$ and the equations (7) hold.

Problem (11) can be solved through use of the Lagrange conditions for optimality. To that end, introduce the functions $g_{st}(\mathbf{X}) = \langle\mathbf{X}\mathbf{e}_s, \mathbf{X}\mathbf{e}_t\rangle - \delta_{st}$, where $\delta_{\cdot,\cdot}$ denotes the Kronecker delta. The constraints in (11) can be summarized

as $g_{st}(\mathbf{X}) = 0$, $g_{st}(\mathbf{Y}) = 0$ for all $s, t \in \{1,\ldots,p\}$. Form the Lagrangian

$$\mathcal{L} = f(\mathbf{X}, \mathbf{Y}) + \sum_{s,t}\lambda_{st}g_{st}(\mathbf{X}) + \xi_{st}g_{st}(\mathbf{Y}),$$

where $\lambda_{st}, \xi_{st}, s, t \in \{1,\ldots,p\}$, are Lagrange multipliers.

Partial derivatives of the objective function are given by

$$\tfrac{\partial}{\partial\mathbf{X}}\mathcal{L} = \left(n - \tfrac{p+1}{2}\right)\mathbf{Y} - \tfrac{p+2}{2}\mathbf{Y}\mathbf{Y}^\top\mathbf{X} - \tfrac{1}{2}\langle\mathbf{X}, \mathbf{Y}\rangle\mathbf{Y} + \mathbf{X}\Lambda,$$

$$\tfrac{\partial}{\partial\mathbf{Y}}\mathcal{L} = \left(n - \tfrac{p+1}{2}\right)\mathbf{X} - \tfrac{p+2}{2}\mathbf{X}\mathbf{X}^\top\mathbf{Y} - \tfrac{1}{2}\langle\mathbf{X}, \mathbf{Y}\rangle\mathbf{X} + \mathbf{X}\Xi,$$

where

$$\Lambda = \sum_{s,t}\lambda_{st}(\mathbf{e}_s\mathbf{e}_t^\top + \mathbf{e}_t\mathbf{e}_s^\top) = [\lambda_{st}] + [\lambda_{ts}] = [\lambda_{st} + \lambda_{ts}],$$

$$\Xi = \sum_{s,t}\xi_{st}(\mathbf{e}_s\mathbf{e}_t^\top + \mathbf{e}_t\mathbf{e}_s^\top) = ([\xi_{st}] + [\xi_{ts}]) = [\xi_{st} + \xi_{ts}],$$

are symmetric. The critical points of $\mathcal{L}$ satisfy

$$\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\langle\mathbf{X}, \mathbf{Y}\rangle\right)\mathbf{Y} - \tfrac{p+2}{2}\mathbf{Y}\mathbf{Y}^\top\mathbf{X} + \mathbf{X}\Lambda = \mathbf{0}, \quad (12)$$

$$\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\langle\mathbf{X}, \mathbf{Y}\rangle\right)\mathbf{X} - \tfrac{p+2}{2}\mathbf{X}\mathbf{X}^\top\mathbf{Y} + \mathbf{Y}\Xi = \mathbf{0}, \quad (13)$$

$[g_{st}(\mathbf{X})] = \mathbf{X}^\top\mathbf{X} - \mathbf{I} = \mathbf{0}$, and $[g_{st}(\mathbf{Y})] = \mathbf{Y}^\top\mathbf{Y} - \mathbf{I} = \mathbf{0}$.

Solve this system for $\Lambda$ and $\Xi$. Introduce $\mathbf{Z} = \mathbf{X}^\top\mathbf{Y}$. Multiply (12) and (13) from the left by $\mathbf{X}^\top$ and $\mathbf{Y}^\top$ respectively, to find

$$\Lambda = -\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}\mathbf{Z}\right)\mathbf{Z} + \tfrac{p+2}{2}\mathbf{Z}\mathbf{Z}^\top, \quad \Xi = \Lambda^\top.$$

Since $\Lambda$ and $\Xi$ are symmetric, it holds that

$$\operatorname{skew}\Lambda = -\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}\mathbf{Z}\right)\operatorname{skew}\mathbf{Z} = -\operatorname{skew}\Xi = \mathbf{0}.$$

Note that $n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}\mathbf{Z} \ge n - p - \tfrac{1}{2} > 0$ for $p < n$ since $\operatorname{tr}\mathbf{Z} \le p$, wherefore $\operatorname{skew}\mathbf{Z} = \mathbf{0}$ (the case of $p = n$ is excluded from consideration since $\operatorname{St}(n,n) \simeq \operatorname{O}(n)$ is not path connected). This implies that $\Xi = \Lambda$.

Substitute

$$\Lambda = -\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}(\mathbf{Z})\right)\mathbf{Z} + \tfrac{p+2}{2}\mathbf{Z}^2$$

into the equation $\mathbf{Y}^\top\nabla_\mathbf{X}\mathcal{L} = \mathbf{0}$ to find

$$\mathbf{Y}^\top\nabla_\mathbf{X}\mathcal{L} = \left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}\mathbf{Z}\right)\mathbf{I} - \tfrac{p+2}{2}\mathbf{Z} +$$
$$\mathbf{Z}\left(-\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}(\mathbf{Z})\right)\mathbf{Z} + \tfrac{p+2}{2}\mathbf{Z}^2\right) = \mathbf{0}$$

By simplifying, we obtain

$$\mathbf{Y}^\top\nabla_\mathbf{X}\mathcal{L} = \left(\left(n - \tfrac{p+1}{2} - \tfrac{1}{2}\operatorname{tr}\mathbf{Z}\right) - \tfrac{p+2}{2}\mathbf{Z}\right)(\mathbf{I} + \mathbf{Z})(\mathbf{I} - \mathbf{Z}).$$

It follows that

$$p(z) = \left(\tfrac{2n - p - 1 - \operatorname{tr}\mathbf{Z}}{p+2} - z\right)(1 + z)(1 - z) \quad (14)$$

is the minimal polynomial of $\mathbf{Z}$, up to the exclusion of any factors corresponding to non-singular matrices.

## E. Integer Programming Problem

Reformulate the nonlinear programming problem (11) as an integer programming problem in terms of the algebraic multiplicities of the eigenvalues of $\mathbf{Z}$. By (14), the spectrum of $\mathbf{Z}$ satisfies $\sigma \subset \{-1, \lambda_*, 1\}$, where

$$\lambda_* = \frac{2n - p - 1 - \operatorname{tr} \mathbf{Z}}{p+2}. \tag{15}$$

Let $m_-, m_*$, and $m_+ \in \{0, \ldots, p\}$ denote the algebraic multiplicities of the eigenvalues $-1, \lambda_*$, and $1$ respectively. First note that $\lambda_* \leq \|\mathbf{Z}\|_2 \leq \|\mathbf{X}\|_2 \|\mathbf{Y}\|_2 = 1$, wherefore $m_* = 0$ if $\lambda_* > 1$. We know an expression for $\lambda_*$ in terms of $\operatorname{tr} \mathbf{Z}$. We also know that $\operatorname{tr} \mathbf{Z} = -m_- + \lambda^* m_* + m_+$ and $p = m_- + m_* + m_+$. Solving for $\lambda_*$ in terms of $m_-$ and $m_*$, we find

$$\lambda_* = \frac{2n - 2p - 1 + 2m_- + m_*}{p + 2 + m_*}.$$

From the inequality $\lambda^* \leq 1$ we obtain $m_- \leq \frac{3}{2}(p+1) - n$. This contradicts $m_- \geq 0$ if $p < \frac{2n}{3} - 1$, in which case $m_* = 0$. For the case of $p = \frac{2n}{3} - 1$ we learn that either $m_* = 0$ or $m_- = 0$. The case of $p > \frac{2n}{3} - 1$ is less informative.

We consider two cases: that of $m_* = 0$ for general $p$ and that of $m_- = 0$ for $p = \frac{2n}{3} - 1$. These two cases exhaust all instances of $\mathsf{St}(p, n)$ for which the pair $(p, n)$ satisfies $p \leq \frac{2n}{3} - 1$. Start with $m_* = 0$. Then

$$\operatorname{tr} \mathbf{Z} = -m_- + m_+ = 2m_+ - p,$$
$$\operatorname{tr} \mathbf{Z}^2 = m_- + m_+ = p,$$
$$(\operatorname{tr} \mathbf{Z})^2 = 4m_+^2 - 4pm_+ + p^2.$$

Recast the nonlinear program (11) as the equivalent quadratic integer program

$$\max (2n+1)m_+ - m_+^2 \quad \text{s.t. } m_+ \in \{0, \ldots, p\} \tag{16}$$

where we used $\mathbf{Z} = \mathbf{X}^\top \mathbf{Y}$, $\langle \mathbf{X}, \mathbf{Y} \rangle = \operatorname{tr} \mathbf{Z}$, $\|\mathbf{Z}\|^2 = \operatorname{tr} \mathbf{Z}^2$, and removed the constant terms.

Denote the objective function of Problem (16) by $h : \mathbb{N} \to \mathbb{R}$. Note that $h(m_+ + 1) = h(m_+) + 2(n - m_+)$. Since $m_+ \leq p \leq n$, the optimization problem is solved by maximizing $m_+$, i.e., $m_+ = p$. This corresponds to $\mathbf{Z} = \mathbf{I}$ and the optimal value of zero in (11). Any suboptimal solution gives a strictly lower objective value.

It remains to consider the case of $p = \frac{2n}{3} - 1$, $m_- = 0$. Calculate

$$\operatorname{tr} \mathbf{Z} = \lambda_* m_* + m_+ = p + (\lambda_* - 1)m_*$$
$$\operatorname{tr} \mathbf{Z}^2 = p + (\lambda_* + 1)(\lambda_* - 1)m_*$$
$$(\operatorname{tr} \mathbf{Z})^2 = p^2 + 2(\lambda_* - 1)pm_* + (\lambda_* - 1)^2 m_*^2.$$

Recast the nonlinear program (11) as the equivalent mixed integer nonlinear program

$$\max -(\tfrac{n}{6} + \tfrac{1}{4} + \tfrac{1}{4}m_*)(1 - \lambda_*)^2 m_*,$$
$$\text{s.t.} \quad \lambda_* \in [0, 1], \ m_* \in \{0, \ldots, p\}, \tag{17}$$

where the constant term has been removed. It is clear that the optimal solution has $\lambda_* = 1$ or $m_* = 0$. It follows that $\sigma(\mathbf{Z}) = \{1\}$, i.e., $\mathbf{Z} = \mathbf{I}$.

## IV. Conclusions

This paper extends the results of [19] concerning almost global stability of a state space oscillator on $\mathsf{S}^n$ for $n \geq 2$ to the case of $\mathsf{St}(p, n)$. We characterize all connected graphs as $\mathsf{St}(p, n)$-synchronizing provided that the pair $(p, n)$ satisfies $p \leq \frac{2n}{3} - 1$. This inequality is sharp with respect to known results: some connected graphs are not $\mathsf{S}^1$-synchronizing, $\mathsf{S}^1 \simeq \mathsf{St}(1, 2)$ and $1 \nleq \frac{1}{3}$ [1]; all connected graphs are $\mathcal{S}^n$-synchronizing since $\mathcal{S}^n \simeq \mathsf{St}(1, n+1)$ and $1 \leq \frac{2n-1}{3}$ for $n \geq 2$ [19]; some connected graphs are not $\mathsf{SO}(3)$-synchronizing, $\mathsf{SO}(3) \simeq \mathsf{St}(2, 3)$ and $2 \nleq 1$ [11]. Still, these results are likely to be conservative for $n \geq 4$ due to the derivations involving a number of inequalities.

## References

[1] A. Sarlette. *Geometry and symmetries in coordination control*. PhD thesis, Liège University, 2009.

[2] B. Afsari. Riemannian $L^p$ center of mass: Existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society*, 139(2):655–673, 2011.

[3] R. Tron, B. Afsari, and R. Vidal. Riemannian consensus for manifolds with bounded curvature. IEEE *Transactions on Automatic Control*, 58(4):921–934, 2013.

[4] R. Hartley, J. Trumpf, Y. Dai, and H. Li. Rotation averaging. *International journal of computer vision*, 103(3):267–305, 2013.

[5] J. Zhu. Synchronization of Kuramoto model in a high-dimensional linear space. *Physics Letters A*, 377(41):2939–2943, 2013.

[6] J. Thunberg, W. Song, E. Montijano, Y. Hong, and X. Hu. Distributed attitude synchronization control of multi-agent systems with switching topologies. *Automatica*, 50(3):832–840, 2014.

[7] C. Lageman and Z. Sun. Consensus on spheres: Convergence analysis and perturbation theory. In *Proceedings of the 55th* IEEE *Conference on Decision and Control*, pages 19–24, 2016.

[8] F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014.

[9] R. Sepulchre. Consensus on nonlinear spaces. *Annual reviews in control*, 35(1):56–64, 2011.

[10] L. Scardovi, A. Sarlette, and R. Sepulchre. Synchronization and balancing on the $N$-torus. *Systems & Control Letters*, 56(5):335–341, 2007.

[11] R. Tron, B. Afsari, and R. Vidal. Intrinsic consensus on SO(3) with almost-global convergence. In *Proceedings of the 51st* IEEE *Conference on Decision and Control*, pages 2052–2058, 2012.

[12] J. Thunberg, J. Markdahl, and J. Goncalves. Dynamic controllers for column synchronization of rotation matrices: a QR-factorization approach. *Automatica*, 93:20–25, 2018.

[13] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multi-Agent Networks*. Princeton University Press, 2010.

[14] A. Sarlette and R. Sepulchre. Consensus optimization on manifolds. *SIAM Journal on Control and Optimization*, 48(1):56–76, 2009.

[15] R. Olfati-Saber. Swarms on the sphere: A programmable swarm with synchronous behaviors like oscillator networks. In *Proceedings of the 45th* IEEE *Conference on Decision and Control*, pages 5060–5066, 2006.

[16] W. Li and M.W. Spong. Unified cooperative control of multiple agents on a sphere for different spherical patterns. IEEE *Transactions on Automatic Control*, 59(5):1283–1289, 2014.

[17] A. Sarlette, R. Sepulchre, and N.E. Leonard. Autonomous rigid body attitude synchronization. *Automatica*, 45(2):572–577, 2009.

[18] U. Helmke and J.B. Moore. *Optimization and dynamical systems*. Springer, 2012.

[19] J. Markdahl, J. Thunberg, and J. Gonçalves. Almost global consensus on the $n$-sphere. IEEE *Transactions on Automatic Control*, 63:1664–1675, 2018.

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

[21] K.B. Petersen and M.S. Pedersen, *et al. The Matrix Cookbook*. Denmark Technical University, 2008.

[22] A. Graham. *Kronecker Products and Matrix Calculus: With Applications*. Wiley, 1981.