Spring 2024

State Omniscience for Cooperative Local Catalog Maintenance of Close Proximity Satellite Systems

Chris Hays
Embry-Riddle Aeronautical University, haysc3@my.erau.edu

STATE OMNISCIENCE FOR COOPERATIVE LOCAL CATALOG MAINTENANCE OF CLOSE PROXIMITY SATELLITE SYSTEMS

By

Christopher W. Hays

A Dissertation Submitted to the Faculty of Embry-Riddle Aeronautical University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Aerospace Engineering

April 2024

Embry-Riddle Aeronautical University

Daytona Beach, Florida

This dissertation is dedicated to my community. Without whom, none of this would have been possible.

ACKNOWLEDGMENTS

First and foremost, I thank God for granting me the ability and opportunity to pursue my passions. I want to thank Dr. Troy Henderson for taking me under his wing and allowing me to work under his supervision and providing me with the opportunity to pursue exciting new research in the field of multi-agent distributed estimation. I would also like to thank Dr. Sean Phillips for his mentorship and guidance while working with the Air Force Research Laboratory - Space Vehicles Directorate, as well as each of the other members of my committee Dr. Morad Nazari, Dr. Riccardo Bevilacqua, and Dr. Sirani Perera for their continuous support throughout the past few years.

Working out of the Space Technologies Laboratory has introduced me to a number of wonderful people and allowed me to work on a number of exciting projects - namely, EagleCam. This is for the EagleCam team, thank you for making these past few years an unforgettable experience; I will never forget the late nights, laughs, frustrations, and triumphs we shared together. Not many can say they sent a camera to the moon during their college careers; but we can. To Daniel Posada, no words can express how thankful I am that I embarked on and get to end this journey with you.

Last and certainly not least, I want to thank my family. To my parents, Ron and Regina, thank you for all of the sacrifices you have made and the support you have given me to get me where I am today. I would not be the man that I am without you. To my wife, Torii, thank you for reminding me daily that life is not about the work I do, but the people I get to work with. Thank you for challenging me, encouraging me, and supporting me. You, more than anyone else, know the ups and downs of these past few years, and I cannot say thank you enough for choosing to stand beside me at every step. To my dogs, Leo and Luna, thank you for making this more difficult than it needed to be by insisting it was play time every time I sat at my office chair to write. You were the reason I could not work from home; but were the perfect distraction when I needed it most.

ABSTRACT

Resiliency in multi-agent system navigation is reliant on the inherent ability of the system to withstand, overcome, or recover from adverse conditions and disturbances. In large part, resiliency is achieved through reducing the impact of critical failure points to the success and/or performance of the system. In this view, decentralized multi-agent architectures have become an attractive solution for multi-agent navigation, but decentralized architectures place the burden of information acquisition directly on the agents themselves. In fact, the design of distributed estimators has been a growing interest to enable complex multi-sensor/multi-agent tasks. In such scenarios, it is important that each local estimator converges to the true global system state - a quality known as state omniscience. Most previous related work has focused on the design of such systems under varying assumptions on the graph topology with simplified information fusion schemes. Contrarily, this work introduces characterizations of state omniscience under generalized graph topologies and generalized information fusion schemes. State omniscience is discussed analogously to observability from classical control theory; and a collection of necessary and sufficient conditions for a distributed estimator to be state omniscient are presented. This dissertation discusses this phenomena in terms of an on-orbit scenarios dubbed the local catalog maintenance problem and the cooperative local catalog maintenance problem. The goal of each agent is to maintain their catalog of all bodies (objects and agents) within this neighborhood through onboard angles-only measurements and cooperative communication with the other agents. A central supervisor selects the target body for each agent, a local controller tracks the selected target body for each agent, and a local estimator coalesces both an agent's measurements and state estimates provided by neighboring agents within the communication graph. Numerical results are provided to demonstrate the supervisor's ability to select an appropriate target subject to an uncertainty threshold, the controller's ability to track the selected target, and the estimator's ability to maintain an accurate and precise estimate of each of the bodies in the local environment.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

1 Introduction

1.1 Introduction

Space is a key component to the global economy by facilitating trillions of dollars, annually, in the global market due to GPS alone [1]. Due to the lower barrier to entry, space is becoming increasingly congested and the current space surveillance network may not be able to fully maintain the entire catalog of small objects in the future [2]. Moreover, with the rise of proliferated low earth orbit constellations like SpaceX Starlink, OneWeb, and Amazon there is a need for commercial and government satellites to maintain an understanding of their local area to discern if an orbit change is required. Therefore, future satellites may, in addition to their operational mission, be required to autonomously maintain a local catalog of nearby space objects to assure the safety of the platforms and autonomously decide if any evasive actions would be required. Satellite autonomy is also increasing in popularity due to such situations, in fact, recently, readiness levels for space trusted autonomy are studied[3]. More generally, this problem requires high-level and coordinated autonomous decision-making.

Advances in autonomous system technology has allowed for multi-agent systems to become an enabling asset for a wide variety of applications - search and rescue [4], inspection [5], patrolling [6], mapping [7], etc. In order to accomplish such tasks, multi-agent systems have intentionally been designed to be robust to disturbances to prevent the disruption of operations with coordinating, cooperating, or collaborating networked teams of agents. Unfortunately, robustness comes with a price [8]. Often the price of robustness takes the form of over-conservatism of the optimal solution which could lead to off-nominal performance if the over-conservatism violates any system constraints. In [9], the authors suggest that resiliency replace robustness as the new central engineering paradigm for multi-agent system navigation. Resiliency within a multi-agent system is defined in [9] as the capability of withstanding or overcoming unexpected, adverse conditions or shocks and unknown, unmodeled disturbances.

An increasing interest in decentralized multi-agent architectures has sprouted in an effort to push for more resiliency in such systems. The lack of reliance on global information allows for decentralized frameworks to be more resilient to adverserial attacks, lack of trust, and communication failures. Unfortunately, this also leads to increased complexity for navigation algorithm design. Either infrastructure must be developed to aid collision avoidance, tracking, and path-planning, or each agent must be capable of accomplishing each task locally. The latter requires each agent to become aware of every other agent within the operating environment, and for most sensing modalities, information gain must inherently be built into all navigation algorithms. In many cases, development of an external infrastructure is infeasible and impractical. Thus this work considers multi-agent systems where the information gathering process must be accomplished locally. This places a unique challenge on the system as available tasks, feasible paths, collision constraints are harder to verify and implement at the agent level.

To effectively evaluate such criteria within a decentralized framework, every agent must have independent knowledge of the state of every other agent operating within the environment. This may be accomplished through distributed estimation. A distributed estimator is comprised of a number of local estimators and leverages both the measurements and communicated information. The convergence of each local estimator is dependent on *i.*) availability of intermittent measurements, *ii.*) the phenomena being tracked and measurements is observable globally, and *iii.*) measurement information is properly routed such that every local estimator converges. This work focuses on methods to guarantee convergence of a distributed estimator, constructing state omniscient distributed estimators, and present its utility in a practical example.

1.2 Literature Review

1.2.1 Space Situational Awareness

In the literature, the topic of satellite localization has been thoroughly studied for close proximity (relative motion) satellite systems [10–16], and ground-based sensor networks

[17, 18]. Due to the number of space objects compared to the number of ground based sensors, the tasking of ground sensors must be intelligently chosen [19, 20]. Situational awareness, in general, can be defined as "the perception of the elements in the environment within a volume of time and space, the (organizational) comprehension of their meaning, and the projection of their status in the near future" [21]. In particular, *space* situational awareness is a growing topic of interest [22]. Space situational awareness (a.k.a. space domain awareness, or SDA) has been expanded to encompass all space environmental impacts as well, which leads to a multidisciplinary domain of research that incorporates facets of information fusion, collection-tasking and exploitation to better assess anomaly identification and prediction. However, solutions to the SSA problem utilizing solely on-orbit assets has only recently been studied [14, 23–25] , and only even more recently has been referred to as the *local catalog maintenance problem* [26–28].

Transcending a traditional catalog maintenance problem, the cooperative local catalog maintenance problem is primarily one of decision-making and efficient communication. One that, due to its cooperative nature, is subjugated to communication constraints. Similar works have taken into account efficient communication for reinforcement-learning based schemes in recent years, but their focus has been decision-making algorithm design [29–31]. The work by Fedeler et. al. [32, 33] migrates these decentralized decision-making approaches to space domain awareness over random communication graphs. Many of these results assume connectivity of the underlying graph topology for effective communication and decision-making. However recent results show that, for estimation problems, connectivity is a sufficient but not necessary condition for state estimate convergence assuming certain conditions are met for the unconnected components. Because accurate and precise state estimate play a significant role in the efficacy of almost all decision-making algorithms, effective decisions may still be made without a connected graph. It is only necessary that each agent is able to gather the requisite information. It does not matter if that agent does by its own accord or through communication with other

agents. This work goes a step further than [32, 33] by strategically constructing a communication graph to support the transmission of information across the network.

1.2.2 Notions of Observability

The study of linear time-invariant (LTI) system observability has been fundamental in the development of control theory since its introduction by Kalman in the early 1960's [34]. Since that time, analogous forms of observability have been introduced to the frequency (via transfer functions) and graph domains [35–37] for LTI system analysis. Studying observability from a graph-based perspective has enabled the development of observability for large-scale complex networks [38, 39] and multi-agent systems [40–42]. Additionally, tools and concepts from algebraic graph theory have been readily applied to describe properties of observable systems. The spectrum of a graph is particularly influential in describing the observability and reconstructability of the graph [43, 44]. The majority of observability literature focuses on identifying observable systems and their properties. This is most commonly accomplished using a binary quantifier - a LTI system pair $(\mathbf{\Phi}, \mathbf{C})$ is either observable or it is not. Lin [35] discusses an extension to observable systems by introducing structural observability. Structural observability considers matrices of specified nonzero patterns to identify the observability properties of a system with a given pattern.

More recently, an interest in decentralized multi-agent systems necessitates alternative forms of the observability that focuses on the agent-level criterion rather than the system-level as a whole. The authors of [45] show that "centralized observability" is only a necessary condition for observability at the agent-level, and use a structured systems theoretic approach to present "distributed observability" to craft a notion of agent-level observability. Hays [27] demonstrates this phenomena in a practical example - each estimator only converges when its own knowledge of the states are observable.

1.2.3 Distributed Estimation

Employing multiple sensors/agents is a scalable and robust strategy for maintaining target tracking or estimation capability within a networked system. Particular within

recent decades utilizing multiple agents to complete such tasks has proliferated monitoring [46–48], fault-detection [49], target tracking, navigation [50–52], and space situational awareness [14, 23–25] arenas which has led to a rising interest in the development of technologies that may enable improved performance and enable autonomous operations. It was observed by He in [53] that there exist four primary strategies to fuse communicated information within a distributed estimator *i.*) Sequential *ii.*) Consensus *iii.*) Gossip and *iv.*) Diffusion. He further points out these strategies exist for both state vector and information vector based filtering and provides a thorough discussion as well as numerical comparison of each of the strategies. The work discusses each of the strategies in terms of their advantages and disadvantages in the fields of global optimality, local consistency, communication burden, and specific topology requirements. Sequential fusion algorithms [54–56] may be simple but require a specific topology and local observability at each node. Consensus algorithms [57–60] may achieve global convergence but may take long to converge; gossiping algorithms [61, 62] share the same problem. Diffusion algorithms [63–65] may have a low communication burden but have no guarantees of global convergence. However, irregardless of the fusion strategy chosen for distributed estimation some form of detectability/observability is required across the network for each local estimator to converge; a characterization that has, unfortunately, been scarce within the literature.

1.2.4 Convergence of Distributed Estimators

Accuracy and confidence in the state of the world is critical for well-resourced decision-making within the highly complex socio-political, economical, military, and emergency response arenas. With advances in automation, it is critical for systems with decision-making authority be endowed with the same knowledge and confidence in knowledge whether they be autonomous systems, multi-agent networks [4–7, 66], mechanical, cyber-physical or environmental observers [26, 27, 67], or power system state estimators [68, 69].

Particularly for networked systems, distributed estimation allows for each agent within the network to make informed decisions regardless of what the agent directly measures. The general distributed estimation problem is well-studied in literature for a variety of problem sets with equally many solutions. Early work in parallel Kalman filters opened the door [70, 71] to distributed dynamic system estimation [72, 73]. Distributed dynamic system estimation may be divided into consensus [57, 74] and data-fusion [73] based schemes that utilize external (to the agent) information to help the estimator to converge at each agent.

The convergence of any estimator is dependent on the ability of the system to reconstruct the system state based solely on its outputs. Often referred to as observability, this system quality is a sufficient condition for state reconstruction [34, 75]. For distributed networked systems, the notion of observability becomes convoluted as a networked system may be considered observable when all of the necessary information is being measured across the network, but it may not be properly routed to parts of the network. For the case of networked systems, the field of observability is partitioned into centralized observability and distributed observability [76, 77]. A system is centrally observable if it meets the classical definition of observability such that a centralized estimator may reconstruct the system states. A system is distributedly observable if it is centrally observable and the information is properly routed such that each independent agent within the network can reconstruct the system states. Therefore, analyzing and constructing communication networks to yield desirable properties have recently been a significant interest [52, 78, 79]. This analysis has commonly been approached through the lens of structured systems theory [35, 80] that is dependent on the placement of non-zero values rather than considering their exact values directly.

1.2.5 State Omniscient Systems

Motivated by increased demand for spatially disparate agents or sensors, the design and performance analysis of distributed state estimators has seen a surge of recent attention [81]. Comprised of a number of local state estimators each only capable of only perceiving

some subset of the global system's state, distributed estimators benefit from the exchange of information amongst the local estimators such that each local estimator may perceive the global system state; a quality Park and Martins [82, 83] refers to as state omniscience - a system property where each local estimator converges to the true system state. State omniscience is a designable property of distributed estimators encapsulating sensed, communicated, and fused information at each local estimator within a networked system. Moreover, state omniscience intertwines the typical discussion on the structural observability of network systems [35] and dynamic system observability [34].

Since its inception, many works have focused on state omniscient distributed estimator design with certain assumptions (typically requiring strong connectivity or strongly connected components) on the underlying communication graph and local output matrices [84–88]. Many such works utilize state-augmentation [82, 83, 89] to cast as a decentralized control problem; and/or utilize scalar weights which may limit the quality of the fusion solution [84, 86, 87, 90–94]. On the other hand, only few works have focused on the characterization of state omniscience as presented in this paper. Park and Martins in [82] present necessary and sufficient conditions for state omniscience of a distributed estimator that assumes bidirectional communication, augmented state variable, and a particular fusion scheme. The ensuing results are reliant on the system to have no unstable fixed modes; i.e. the distributed estimator is detectable. These results were later extended to include directional communication in [83]. Other similar works, focus on the observability of networked systems through the lens of the graph Laplacian implying consensus-based algorithms [95–97] which have been shown to be "unlikely" to be completely controllable and, dually, unlikely to be completely observable [98].

Unlike [82], we make no underlying assumptions on the structure of the communication graph, nor do we assume scalar fusion weights, nor do we consider state augmentation. Moreover, while the results presented in [99] may have similar implications to the result herein, they remain strongly rooted in structured systems and graph theory; where, in fact,

the main objective of this work is to draw clear and explicit connections between dynamic observability at the agent level and graph topology while maintaining a sharp distinction between the underlying dynamics and communication infrastructure. Drawing definitive connections between the results herein and the results from [99] may be an interesting topic for future investigations. Our primary focus is the construction of necessary and sufficient conditions for state omniscience in such a way that they are easily recognizable to anyone who has an introductory understanding of linear control theory. We accomplish this by making explicit connections to standard results, particularly the observability matrix. We expand on the standard results by incorporating network traversal; thus, becoming a strict generalization of the traditional notion of observability. Ultimately, it is the belief of the authors that this work clarifies and relaxes much of the results from the existing literature by uniquely tying structural and dynamical observability [97, 100].

1.2.6 Information Fusion

Fusing pieces of information (or data) has long been an outstanding problem within the scientific [101], industrial engineering [47, 49], and robotics communities [102]. Whether combining the information provided by multiple sensors on a single platform or through a networked team of agents [46], the objective of information fusion algorithms is to extract information from measurements and provide the most accurate, and precise, representation of the information provided and place it into a form effective for decision-making [103]. For systems with known correlation between statistics–e.g., mean, covariance, cross-correlation–the Kalman Filter provides a minimum mean squared error result.

However, in most practical applications it is impossible to know the cross-correlation between multiple sources, resulting in the Kalman Filter returning suboptimal solutions. As sensing networks become more distributed and complex, particularly within large, heterogeneous networks of sensors that operate at different frequencies [104], suboptimal solutions can quickly (within one sensing cycle) degrade estimation performance. It is imperative that fused solutions maintain a structure as close to the optimal solutions as

possible.

Additionally, the computational burden of fusing many pieces of information can make the problem computationally intractable (at least in the time it takes the dynamics or states of the system being observed to change). To address this issue, a number of distributed fusion algorithms have been introduced for networked systems [105]. This, unfortunately, compounds the problem of unnecessary covariance growth as conservative "over-approximations" of the covariance are added at every intersecting fusion node, rather than all at once. As distributed and decentralized sensor networks have seen significant usage within recent years, novel techniques in information fusion that preserve the structure of the optimal solution and maintain computational feasibility are warranted. Methods to provide such a solution have been considerably studied in the literature [106–112]. However, many of these solutions employ schemes that competitively weight these information sources and could potentially bias the fused solution towards a specific piece of information. In contrast, other solution methods attempt to cooperatively fuse the available information at the cost of maintaining the appropriate lower (consistent) and upper (tight) bounds.

1.3 Contributions

This work makes three primary contributions related to multi-agent autonomous system guidance and navigation

1. Formally Characterize State Omniscience
2. Introduce and Address the Local Catalog Maintenance Problem
3. Improve information fusion techniques through matrix decomposition.

The following sections detail the nature of each of the listed contributions.

1.3.1 Formally Characterize State Omniscience

Ubiquitous Single-Output Observability

Rather than continuing the structured systems approach as is commonly done in the literature, this work explores a form of distributed controllability/observability using the classical tools in an attempt to shed some light on the underlying phenomena. Specifically, this work identifies a specific class of LTI systems that are observable independent of any nonzero input matrix $\mathbf{C}$. Rather than introducing this class of linear systems, this work formally characterizes this particular class of linear systems and endows them with a list of properties, and provides direct ties to other notions of observability. To the best of the author's knowledge, this is the first time this particular class of LTI systems has been formally quantified and characterized. For such systems, the authors use the term Ubiquitous Single-Output Observability (USOO). USOO systems have practical applications in the design fault tolerant systems [113]. Finally, Ubiquitous Single-Input Controllability, the dual form of USOO, system provides applications to the previously mentioned decentralized controller scenarios. If an actuator fails and it is important to maintain full controllability of the system, a USIC system will be able to maintain such a property regardless of which specific actuator failed. Moreover, USIC systems may be able to provide multiple actuation methods to achieve the same goal which can lead to intelligent actuator and sensor selection and optimization.

Heuristic and Set-theoretic State Omniscience

This work is primarily concerned with the design of a network system to yield distributed estimator convergence. In particular, we consider the interactions of sensed and communicated information as graphs to study necessary and sufficient conditions such that the estimator at each agent will converge to the truth states. One significant challenge this paper addresses is the modeling of sensing connections between vertices. Because a sensing-only network does not behave the same as traditionally considered graph networks, interpreting the traditional notions of networked system observability is not immediately

clear. This work attempts to clarify the role sensing connections have on networked system observability by merging the sensing connections and measurement matrix to form a single augmented measurement matrix.

A set theoretic approach introduces a binary quantifier for the convergence of a distributed estimator, something that has, as of yet, not been found in the literature. It is also shown that under certain design considerations this convergence criteria and observability are equivalent. Additionally, rather than restricting the estimator to scalar fusion weights, as has been traditionally done, this work generalizes the estimator notation such that more sophisticated fusion techniques [106, 109, 114] may be considered. In light of this, the design (graph topology) of networked distributed estimators is considered by establishing convergence conditions for each of the possible combinations of the sensed and communicated graph topologies.

State Omniscience as a Generalization of Classical Observability

Unlike [82], we make no underlying assumptions on the structure of the communication graph, or utilize scalar fusion weights, nor state augmentation. Moreover, while the results presented in [99] may have similar implications to the result herein, they remain strongly rooted in structured systems and graph theory; where, in fact, the main objective of this work is to draw clear and explicit connections between dynamic observability at the agent level and graph topology while maintaining a sharp distinction between the underlying dynamics and communication infrastructure. Drawing definitive connections between the results herein and the results from [99] may be an interesting topic for future investigations. Our primary focus is the construction of necessary and sufficient conditions for state omniscience in such a way that they are easily recognizable to anyone who has an introductory understanding of linear control theory. We accomplish this by making explicit connections to standard results, particularly the observability matrix. We expand on the standard results by incorporating network traversal; thus, becoming a strict generalization of the traditional notion of observability. Ultimately, it is the belief of the authors that this

work clarifies and relaxes much of the results from the existing literature by uniquely tying structural and dynamical observability [97, 100].

1.3.2 Introduce and Address the Cooperative Local Catalog Maintenance Problem

Additionally, Chapter 3 introduces a decision-making algorithm which supplies a control trajectory to the agent such that the uncertainties of all tracked deputies are guaranteed to lie underneath a predefined bound. The deputies are modeled to maneuver about a virtual chief according to the well-known Clohessy-Wiltshire dynamics, while each satellite's attitudinal dynamics are modeled according to Euler's rotational equations. The state estimation is performed using a distributed version of the extended Kalman filter, and the sequence of control inputs to the agent are computed via Model Predictive Control (MPC). Finally, a supervisory decision logic framework is constructed. It determines which sequence of deputies to observe based on the estimated relative positions and uncertainty effectively combining the decision-making and graph construction algorithms developed in [27, 28] with the control pipeline developed in [26].

Furthermore, Chapter 6 considers the same problem where the agents are capable of communication. It introduces a communication graph construction algorithm that maintains the desired state omniscient condition. Succinctly, this work presents a first-of-its-kind end-to-end algorithm that considers decision-making, estimation, and control to accomplish a multi-target tracking objective.

1.3.3 Contributions to Information Fusion

Reviewing much of the related literature from Section 1.2.6, CI is shown to maintain consistency but is not necessarily tight. EI is tight but does not guarantee consistency; and ICI, while it maintains both consistency and tightness, competitively weights the information provided which ultimately limits the best fused estimate the method can provide. To address each of these issues, the present work derives a fusion methodology that preserves both consistency and tightness while cooperatively fusing the information

provided by $\mathcal{A}$ and $\mathcal{B}$. Upper and lower bounds on a scaling parameter, $\mathbf{\Omega}$, are also derived and proved analytically to maintain consistency and tightness of the presented fusion solution. Finally a method to calculate the mutual mean between the information sources using the proposed methodology is also introduced.

1.4 Dissertation Outline

The remainder of this dissertation is outlined as follows: Chapter 2 lays out the notation and necessary preliminaries, Chapter 3 introduces and formalizes the local catalog maintenance problem. Chapter 4 transitions the train of thought to discussing state omniscient, a necessary development for the remainder of the work. Similarly, Chapter 5 lays out a novel information fusion scheme founded in matrix decomposition. Chapter 6 resumes the discussion on local catalog maintenance but considers communication amongst agents and refers to the problem as the *cooperative local catalog maintenance problem* and Chapter 7 presents numerical results for the *cooperative local catalog maintenance problem.* Finally Chapter 8 provides a summarized of the key results and recommendations for future work.

2 Notation and Preliminaries

2.1 Notations and Definitions

Let $\mathbb{R}$ be the set of real numbers and $\mathbb{N}$ be the set of natural numbers. Next, let $\mathbf{a} \in \mathbb{R}^n$ be a column vector of dimension n, $\mathbf{F} \in \mathbb{R}^{p \times n}$ be a matrix with p rows and n columns, and $\mathbf{A}$ denotes a set. Specifically, $\mathcal{X} \subseteq \mathbb{R}^n$ is the set of all vectors a state vector $\mathbf{x}$ may take. The operator $\mathbf{A} \backslash \mathcal{X}$ is the set difference of $\mathbf{A}$ and $\mathcal{X}$ and is the elements of $\mathbf{A}$ that are not in $\mathcal{X}$. We use the notation $(\cdot)_{lk}$ to indicate the element (or block element) of a matrix. The operator $\otimes$ is a block-matrix generalization of the standard Kronecker product denoted $\mathbf{W} \otimes \mathbf{\Phi}$ that may be applied to block matrices of $\mathbf{W}$, namely $\mathbf{W}_{ij}$, such that $(\mathbf{W} \otimes \mathbf{\Phi})_{ij} = \mathbf{W}_{ij}\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ given $\mathbf{W} \in \mathbb{R}^{nN \times nN}$, $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$, and $(\mathbf{W} \otimes \mathbf{\Phi}) \in \mathbb{R}^{nN \times nN}$ where n is the size of the state space and N is the number of vertices in the underlying graph considered. The operator $\text{diag}(\cdot)$ places all of the given elements on the block diagonal of the resulting matrix, and the operator $\text{diagm}(\cdot)$ only includes the elements on the block diagonal of the matrix. The matrix $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ represents the identity matrix, $\mathbf{0}_n \in \mathbb{R}^{n \times n}$ represents the zero matrix, and $\mathbf{J}_i = [\bar{\mathbf{J}}_1, \ldots, \bar{\mathbf{J}}_j, \ldots, \bar{\mathbf{J}}_N] \in \mathbb{R}^{nN \times nN^2}$ and

$$\bar{\mathbf{J}}_j = \begin{cases} \mathbf{0}_{nN}, \ j \neq i \\ \mathbf{I}_{nN}, \ j = i \end{cases}.$$

We say $\mathbf{1}_N$ is the N-dimensional vector of all ones. We notate a frame with a script $\mathscr{I}$. We also notate the image and null of a space $\mathbf{A}$ as $\text{Im}(\mathbf{A})$ and $\text{Null}(\mathbf{A})$, respectively.

A frame for an n-dimensional space is defined by n orthonormal vectors $\hat{\mathscr{I}}_i$, $i \in [1, n]$. In this paper, we work with spaces where $n = 3$. The Special Orthogonal Group SO(3) is the set of all real invertible 3×3 matrices that are orthogonal with determinant 1. We refer to individual elements of SO(3) as rotation matrices $\mathbf{R}$:

$$\text{SO}(3) := \{\mathbf{R} \in \mathbb{R}^{3 \times 3} \mid (\mathbf{R})^\top \mathbf{R} = \mathbf{I}_3, \ \det \mathbf{R} = 1\}$$

Given two coordinate frames $\mathscr{H}$ and $\mathscr{B}$, the rotation of vector $\mathbf{x}^{\mathscr{H}}$ as measured in frame $\mathscr{H}$ to $\mathbf{x}^{\mathscr{B}}$ as measured in frame $\mathscr{B}$ is denoted by $\mathbf{R}^{\mathscr{B}}_{\mathscr{H}}$, that is, $\mathbf{x}^{\mathscr{B}} = \mathbf{R}^{\mathscr{B}}_{\mathscr{H}}\mathbf{x}^{\mathscr{H}}$. Given three frames $\mathscr{H}$, $\mathscr{B}$, and $\mathscr{I}$, the angular velocity of $\mathscr{B}$ relative to $\mathscr{H}$ as measured in frame $\mathscr{I}$ is denoted by ${}^{\mathscr{I}}\omega^{\mathscr{B}}_{\mathscr{H}}$. For a vector $\mathbf{a} := (a_1\ a_2\ a_3)^\top \in \mathbb{R}^3$, the cross product operation (or skew operation) $(\cdot)^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$:

$$(\mathbf{a})^\times = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \tag{2.1}$$

If $\mathbf{b} \in \mathbb{R}^3$, then the cross-product $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ can be written $\mathbf{c} = (\mathbf{a})^\times\mathbf{b} = -(\mathbf{b})^\times\mathbf{a}$. Here, the Lie algebra $\mathfrak{so}(3)$ is the set of all real 3×3 skew-symmetric matrices and is the tangent space of SO(3) at the identity $\mathbf{I}_3$:

$$\mathfrak{so}(3) := \{((\mathbf{a})^\times)^\top = -(\mathbf{a})^\times \mid \mathbf{a} \in \mathbb{R}^3\}$$

For a vector $\mathbf{a} \in \mathbb{R}^3$, the $\mathcal{L}^2$-norm operator is defined as $||\mathbf{a}|| := \sqrt{\mathbf{a}^\top\mathbf{a}}$. A multivariate Gaussian distribution of an n-dimensional random variable is parameterized by a mean vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ and a symmetric, positive semidefinite covariance matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$. Given matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{p\times q}$, the Kroenecker product is denoted by $\otimes$ and is given by

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix}.$$

Given a vector $\mathbf{a} \in \mathbb{R}^m$, the projection of $\mathbf{a}$ onto $\mathbb{R}^n$, $n \leq m$ is denoted by $\mathbf{a} \downarrow \mathbb{R}^n$. Next, we specify a mapping between the vector $\mathbf{x}$ in Cartesian space to the azimuth-elevation or

(Az, El) space $g(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{S}^2$

$$g(\mathbf{x}) = \begin{bmatrix} \text{Az} \\ \text{El} \end{bmatrix} = \begin{bmatrix} \tan^{-1}(\frac{s_2}{s_1}) \\ \sin^{-1}(s_3) \end{bmatrix} \tag{2.2}$$

where $s = \frac{\mathbf{x}}{||\mathbf{x}||}$ and $\mathbb{S}^2$ is the unit sphere with angular coordinates $\text{Az} \in [-\pi, \pi]$ and $\text{El} \in [-\frac{\pi}{2}, \frac{\pi}{2})$ by astronomical convention.

2.2 Graph Theory

Consider an interconnected system of N vertices that may be modeled as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$ with $\mathcal{V} = \{1, 2, \ldots, N\}$ being the vertex set, $\mathcal{E} = \{(i, j) \in \mathcal{V} \times \mathcal{V}\}$ is the edge set, and $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is the adjacency matrix. A specific edge is defined by the tuple(j, i) exists if vertex i can access the information of vertex j. The neighbor set of vertices i is $\mathcal{N}_i = \{j \in \mathcal{V} \ : \ (j, i) \in \mathcal{E}\}$, constructed if a vertex j exists and its connection to i exists within the edge set. The adjacency matrix is constructed with $a_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. In other words, two vertices are adjacent if an edge exists between the two vertices. The incidence matrix of a digraph $\mathbf{E} = [\epsilon_{ij}] \in \mathbb{R}^{N \times |\mathcal{E}|}$ where $\epsilon_{ij} = 1$ if edge e_j leaves vertex i, $\epsilon_{ij} = -1$ if edge e_j enters vertex j and $\epsilon_{ij} = 0$ otherwise. Note that the sum of each column of $\mathbf{E}$ is equal to 0. We use the notation ${}^i\mathbf{E}^\top$ to indicate the incidence matrix of the collection of in-edges at agent i.

The directionality of the arrows in Figure 2.1 shows the flow of information through the network.

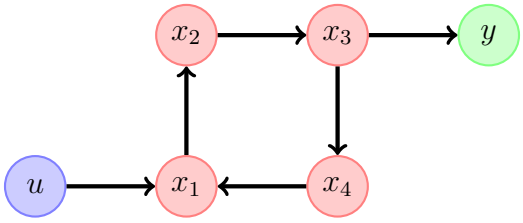


Figure 2.1 Example Graph Topology Modeling State Interactions, Inputs, and Outputs of a Linear Dynamic System.

Given a graph $\mathcal{G}$ with adjacency matrix $\mathbf{A}$ its characteristic polynomial is notated as

$p_{\mathbf{A}}(s) = |s\mathbf{I} - \mathbf{A}|$ where $|\cdot|$ indicates the matrix determinant. Later in the work, we will show that consideration for $\mathcal{G}$ with a specific vertex i removed will need to be made, therefore, we consider a set notation to indicate a vertex removed from $\mathcal{G}$ as $\mathcal{G} \setminus \{i\}$ where $\cdot \setminus \cdot$ notates the set minus. In a matrix case, this means that the i^{th} row and column is removed from the adjacency matrix, similarly, this case is notated as $\mathbf{A} \setminus \{i\}$. For more information on algebraic graphical topology see [44]. Algebraic interpretations of the adjacency matrix are particularly useful in analyzing the graph structure. The matrix powers of the adjacency matrix are known as the adjacency algebra.

Definition 1. ***(Adjacency Algebra)*** *The adjacency algebra, $\mathbf{A}(\mathcal{G})$, is the algebra of polynomials constructed by the matrix powers of $\mathbf{A}$.*

It will later be shown that the USIC property of the system is directly tied to the rank of the adjacency algebra. With that said, the adjacency algebra set $\{\mathbf{I},\ \mathbf{A},\ \mathbf{A}^2\ \ldots,\ \mathbf{A}^{n-1}\}$ is considered to be linearly independent if the vectorized forms of each matrix in the set is linearly independent from each other. nThe adjacency algebra also explicitly indicates the walks that exist within a graph.

Definition 2. ***(Walk)*** *A walk from vertex i to vertex j is said to exist if there is a sequence of vertices and directional edges that connects vertex i to vertex j.*

The number of walks between any two vertices may be specified as the powers of the adjacency matrix. The element $\mathbf{A}^l_{ij}$ indicates the number of l-length walks between vertex i and vertex j[115].

Definition 3. ***(Strongly Connected)*** *A graph $\mathcal{G}$ is called strongly connected if any two vertices are connected by a walk.*

Next, let us consider $\mathcal{S}, \mathcal{T} \subseteq \mathcal{V}$ to be a set of source and target vertices. The vectors

$\mathbf{s} = [\mathbf{s}_1, \ldots, \mathbf{s}_N]^\top \in \mathbb{R}^N$ and $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_N]^\top \in \mathbb{R}^N$ are constructed as follows

$$\mathbf{s}_i = \begin{cases} 1, \ i \in \mathcal{S} \\ 0, \ \text{otherwise} \end{cases} \qquad \mathbf{t}_i = \begin{cases} 1, \ i \in \mathcal{T} \\ 0, \ \text{otherwise} \end{cases} . \tag{2.3}$$

We then define the walk matrix for the set of all walks ending in $\mathcal{T}$ as

$$\mathbf{W} = \begin{bmatrix} \mathbf{t}^\top \\ \mathbf{t}^\top \mathbf{A} \\ \vdots \\ \mathbf{t}^\top \mathbf{A}^d \end{bmatrix}; \tag{2.4}$$

we may also define the walk matrix from $\mathcal{S}$ to $\mathcal{T}$ as

$$\mathbf{W}^{\mathcal{S}} = \begin{bmatrix} \mathbf{t}^\top \\ \mathbf{t}^\top \mathbf{A} \\ \vdots \\ \mathbf{t}^\top \mathbf{A}^d \end{bmatrix} \mathbf{s} \tag{2.5}$$

where d is known as the graph diameter and is the length of the longest shortest walk between two vertices [116]. We say an agent $j \in \mathcal{W}_i$ is in the walk set of agent i if and only if for $\mathbf{s}_j = \mathbf{i}_j$ and $\mathbf{t}_i = \mathbf{i}_i$ $\mathbf{W}^{\mathcal{S}} \mathbf{s}_j \neq \mathbf{0}$. This implies i to be in its own walk set. Furthermore, we say a digraph is *weakly connected* if its undirected counterpart is *connected*, i.e. there exists a walk $i \to j \ \forall \ i, j$ in the undirected graph.

2.3 Heterogeneous Networks

Each vertex in the network has a set of underlying (not necessarily identical) dynamics; neither do the edges necessarily represent similar interactions between vertices; referred to as a heterogeneous network [117]. From this point forward, we will refer to the vertices of the heterogeneous network as "agents" because each vertex of the graph represent an

individual agent of a multi-agent system. For example, consider Figure 2.2, the directionality of the red arrows represent flow of communicated information and the directionality of the blue arrows represent the flow of sensed information.

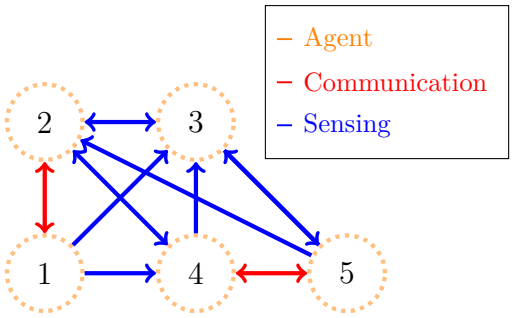
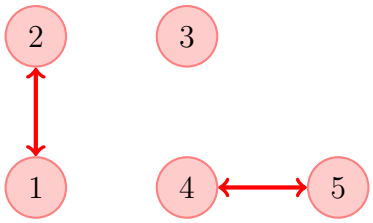


Figure 2.2 Example of modeling a heterogeneous network describing the system with coupling through communication and sensing interactions.

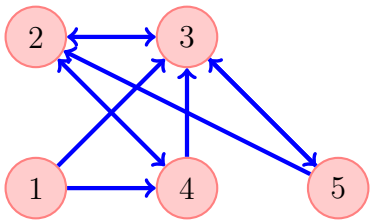
Note that these information types will take the form of distinct and possibly disjoint graphs, for example, these are depicted as two separate graphs in Figure 2.3. Therefore, the overall information flow over the heterogeneous graph is constructed by the union of the graphs $\mathcal{G}_s$ and $\mathcal{G}_c$ describing the sensing and communication networks, respectively, i.e., $\mathcal{G} \leftarrow \mathcal{G}_s \bigcup \mathcal{G}_c$[1].

Due to the inherent behavior of these separate types of information, a distinction must be made. The communication graph in Figure 2.3a allows for the information to pass through vertices, while the sensing graph in Figure 2.3b does not allow for the information to be relayed through a vertex; the arrows indicate the direction of information flow. This distinction is a reflection on how these types of information may flow through the graph; an agent cannot sense what another agent is sensing, but an agent may communicate what it senses to another agent.

[1]The subscripts s and c notate sensing and communication graphs, respectively.



(a) Communication $\mathcal{G}_c$



(b) Sensing $\mathcal{G}_s$

Figure 2.3 Graph models of communication and sensing network topologies where the arrows indicate the directionality of information flow. The communication graph in 2.3a allows for the information to pass through vertices, while the sensing graph in 2.3b does not allow for the information to pass through a vertex.

2.4 Multi-agent Linear Time-Invariant Systems

We consider each inidvidual agent is independently governed by a discrete-time linear time-invariant system

$$\begin{aligned} \mathbf{x}_i(k+1) &= \mathbf{\Phi}_i \mathbf{x}_i(k) \\ \mathbf{y}_i(k) &= \mathbf{C}_i \mathbf{x}(k) \end{aligned} \tag{2.6}$$

where $\mathbf{x}_i(k)$ is the state of agent i at timestep k, and $\mathbf{\Phi}_i \in \mathbb{R}^{n \times n}$ is the system matrix, $\mathbf{C}_i \in \mathbb{R}^{p_i \times nN}$ is the output matrix, $\mathbf{x}_i(k) \in \mathbb{R}^n$ and $\mathbf{y}_i(k) \in \mathbb{R}^{p_i}$ are the state and output vectors of agent i for each $i \in \mathcal{V}$ and $k \in \mathbb{N}$ is the discrete-time step. The output vector $\mathbf{y}_i(k)$ in (2.6) is dependent on the full state vector, $\mathbf{x}^\top(k) = [\mathbf{x}_1^\top(k), \ldots, \mathbf{x}_N^\top(k)]$. The full output vector $\mathbf{y}^\top(k) = [\mathbf{y}_1^\top(k), \ldots, \mathbf{y}_N^\top(k)]$ can then be written as

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{\Phi} \mathbf{x}(k) \\ \mathbf{y}(k) &= \mathbf{C} \mathbf{x}(k) \end{aligned} \tag{2.7}$$

where $\mathbf{\Phi} = \text{diag}([\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_N]) \in \mathbb{R}^{nN \times nN}$ and $\mathbf{C} = [\mathbf{C}_1^\top, \ldots, \mathbf{C}_N^\top]^\top \in \mathbb{R}^{p \times nN}$, where $p = \sum_{i=1}^{N} p_i$. The operator $\text{diag}(\cdot)$ places all of the given elements on the block diagonal of the resulting matrix, and $\mathbf{\Phi}_i$ is the system matrix of agent i. Consequently, $\mathbf{C}_i$ represents all of the measured information available at agent i and $\mathbf{C}$ represents all of the measured information across the network. The full network can be modeled by taking the stack of $\mathbf{x}_i$ which leads to (2.7). For a homogeneous agent system, $\mathbf{\Phi} = (\mathbf{I}_N \otimes \bar{\mathbf{\Phi}})$ where $\mathbf{\Phi}_1 = \mathbf{\Phi}_2 = \cdots = \mathbf{\Phi}_N = \bar{\mathbf{\Phi}}$. In this setting, an agent may be capable of measuring other agents, we can model this using a measurement matrix $\mathbf{H} \in \mathbb{R}^{q \times n}$ such that

$$\mathbf{y}_i(k) = ({}^i\mathbf{E}_s^\top \otimes \mathbf{H})\mathbf{x}(k) = \ \mathbf{C}_i\mathbf{x}(k) \tag{2.8}$$

where $({}^i\mathbf{E}_s^\top \otimes \mathbf{H}) = \mathbf{C}_i \in \mathbb{R}^{p_i \times nN}$ is the augmented output matrix.

2.5 Observability

For linear systems, a state representation of a full dynamic network may be constructed using techniques adapted from [118] and [97]. For example the underlying dynamics of each agent in a heterogeneous network can be modeled using a system matrix ($\mathbf{\Phi}_i$) for each $i \in \mathcal{V}$, the sensing graph ($\mathcal{G}_s$), and the communication graph ($\mathcal{G}_c$). In this section, we provide the general notions of observability considered in this paper. Observability has historically been studied and well understood in the literature [34, 35, 97]. The traditional notion of observability where the initial state and trajectory may be inferred only from the system outputs is given as follows

Definition 4. ***(Observability)*** *A dynamic system in* (2.6) *is considered to be observable if the initial state of the system* $\mathbf{x}_i(0)$ *is able to be reconstructed from a time history of the system outputs* $\{\mathbf{y}_i(0), \mathbf{y}_i(1), \ldots, \mathbf{y}_i(k)\}$.

2.5.1 Binary Observability Quantifiers

The notions of controllability and observability in dynamic systems were originated by Kalman [34] and have since played a pivotal role within the realm of control theory. In

general, a dynamic system is considered to be controllable if there exists a control signal $\mathbf{u}$ that drives the system to a desired state in finite time. A dynamic system is considered to be observable if given an observation signal $\mathbf{y}$ the initial state of the system may be reconstructed. Proving whether a given dynamic system is controllable or observable is a well studied problem and takes a variety of forms depending on the dynamic system representation. Consider the state-space representation of a linear time-invariant dynamical system in (2.9)

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{\Phi x} + \mathbf{Bu} \\ \mathbf{y} &= \mathbf{Cx} \end{aligned} \tag{2.9}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the set of states, $\mathbf{u} \in \mathbb{R}^p$ is a vector of control inputs, and $\mathbf{y} \in \mathbb{R}^q$ is the measurement vector, $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ is the continuous time state transition matrix, $\mathbf{B} \in \mathbb{R}^{n \times p}$ is the input matrix that maps the inputs to states, and $\mathbf{C} \in \mathbb{R}^{q \times n}$ is the output matrix that maps the states to outputs. Each relate the states of the system and external inputs to how the system changes over time and how it is observed.

For LTI systems, controllability and observability are well-studied problems where the following statements are equivalent:

- The pair $(\mathbf{\Phi}, \mathbf{C})$ is observable.
- The observability matrix $\mathbf{O}$ is full column rank.

$$\mathbf{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C\Phi} \\ \mathbf{C\Phi}^2 \\ \vdots \\ \mathbf{C\Phi}^{n-1} \end{bmatrix}$$

- The observability Gramian $\mathbf{G}_o(t)$ is nonsingular for any $t > 0$.

$$\mathbf{G}_o(t) = \int_0^t e^{\mathbf{\Phi}^T\tau}\mathbf{C}^T\mathbf{C}e^{\mathbf{\Phi}\tau}d\tau$$

- The matrix $[(\mathbf{\Phi} - \lambda I)^T,\ \mathbf{C}^T]^T$ is full column rank for every eigenvalue λ of $\mathbf{\Phi}$, known as the Popov-Belevitch-Hautus (PBH) observability test.

- There does not exist a right-eigenvector of $\mathbf{\Phi}$ such that $\mathbf{Cv} = 0$ with

- The system is in or can be placed in observable canonical form, where α_i is the i^{th} coefficient of the characteristic polynomial of $\mathbf{\Phi}$.

$$\mathbf{\Phi}' = \begin{bmatrix} -\alpha_n & 1 & 0 & \dots & 0 \\ -\alpha_{n-1} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\alpha_2 & 0 & 0 & \dots & 1 \\ -\alpha_1 & 0 & 0 & \dots & 0 \end{bmatrix},\ \mathbf{C}' = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

All the aforementioned observability tests are also applicable to linearized versions of nonlinear systems and indicate local observability of the system. Another widely accepted for evaluating the local observability condition of a nonlinear vector field $f(\mathbf{x})$ with a nonlinear measurement function $h(\mathbf{x})$ is the Lie derivative shown in (2.10) [119–121].

$$\mathbf{O} = \frac{\partial}{\partial \mathbf{x}} \begin{bmatrix} L_f^0 h(\mathbf{x}) \\ L_f^1 h(\mathbf{x}) \\ \vdots \\ L_f^{N-1} h(\mathbf{x}) \end{bmatrix};\ L_f^N h(\mathbf{x}) = \frac{\partial L_f^{N-1} h(\mathbf{x})}{\partial \mathbf{x}} f(\mathbf{x});\ L_f^1 h(\mathbf{x}) = \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} f(\mathbf{x}) \tag{2.10}$$

Other approaches exist for the global observability of nonlinear systems but are reliant on Lie bracket analysis and are considered outside the scope of this work [122, 123].

These metrics are not strictly binary quantifiers of observability, like the rank condition of the observability matrix, rather they provide an idea about how much state information is available to the system.

2.5.2 Observability Quality

Friedland [124] adapts the condition number of a matrix to be a quantifier of the observability of the system. He then defines this observability quantifier as shown in (2.11) and coins it as the "coefficient of observability", ζ.

$$\zeta = \frac{\lambda_{min}(\mathbf{F})}{\lambda_{max}(\mathbf{F})} \tag{2.11}$$

where $\mathbf{F} = \mathbf{O}^T\mathbf{O}$ or $\mathbf{F} = \mathbf{G}_o$. By definition, $0 \leq \zeta \leq 1$, because $\lambda_{min}(\mathbf{F}) \leq \lambda_{max}(\mathbf{F})$ except for when $\lambda_{max} = 0$ which implies the dimension of the observable space is negligible. Therefore as ζ approaches 1, the better the observability of the dynamic system. However, unlike the rank condition of the observability matrix, the matrix condition number is sensitive to similarity transforms [125]. An immediate extension to a multi-agent system would be to directly use the augmented dynamical system to construct an augmented observability matrix using the dynamics defined by (2.7).

From a graph topology persepctive, an analogous observability matrix can be constructed in Equation 2.12

$$\mathbf{O}_i = \begin{bmatrix} c_i \\ c_i\mathbf{A} \\ c_i\mathbf{A}^2 \\ \vdots \\ c_i\mathbf{A}^{n-1} \end{bmatrix} \tag{2.12}$$

where $\mathbf{A}$ is the adjacency matrix of the graph and c_i is the i^{th} row of the identity matrix. Physically, the row vector c_i queries the information available at Agent i, therefore the matrix $\mathbf{F}_i = \mathbf{O}_i^T\mathbf{O}_i$ quantifies the quality of the observability condition at Agent i. This

may also be adapted for a system wide approach as shown in Equation 2.13.

$$\bar{\mathbf{O}} = \text{diag}([\mathbf{O}_i]) \; \forall \; i = 1, \ldots, n \tag{2.13}$$

Consequently, $\bar{\mathbf{F}} = \bar{\mathbf{O}}^T \bar{\mathbf{O}}$. However this direct approach still suffers from a hefty computational burden, as the size of $\bar{\mathbf{F}}$ grows quadratically with n. So, may only be of practical use for systems of small dimension.

2.5.3 Structural Controllability

Linear dynamic systems may take the form of directed graphs (digraphs) for both continuous and discrete time systems. For such systems, we define the state $\mathbf{x} \in \mathbb{R}^n$, the input $\mathbf{u} \in \mathbb{R}^q$ and output $\mathbf{y} \in \mathbb{R}^p$. In this work, we consider continuous-time systems of the form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{\Phi x} + \mathbf{B u} \\ \mathbf{C} &= \mathbf{C x} \end{aligned} \tag{2.14}$$

where $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ is the system matrix, $\mathbf{B} \in \mathbb{R}^{n \times q}$ is the input matrix, and $\mathbf{C} \in \mathbb{R}^{p \times n}$ is the output (or measurement) matrix. By interpreting the system matrix $\mathbf{\Phi}$ as an adjacency matrix $\mathbf{A}$ for a graph network, many of the tools from graph theory may be leveraged to describe and analyze the equivalent linear dynamic system. For example, consider the graph in Figure 2.1 which can equivalently be represented as a linear dynamic system (2.14) where $\mathbf{x} = [x_1, x_2, x_3, x_4]^\top$ with matrices given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; \mathbf{C} = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.15}$$

In this work, we consider a graph describing a continuous-time LTI system with following assumptions.

Assumption 1. *There is no edge pointing to a vertex that represents a control input.*

Assumption 2. *The input matrix* $\mathbf{B} \in \mathbb{R}^{n\times 1}$ *and the output matrix* $\mathbf{C} \in \mathbb{R}^{1\times n}$ *are limited to having a singular non-zero element. This directly indicates each input only has direct control over and each output only receives information from a single state rather than a linear combination of multiple states.*

On the surface, these assumptions may seem overly restrictive. However, many of the following results will generalize to almost any non-zero n-dimensional vector $\mathbf{B}$ the main results are agnostic to the structure of $\mathbf{B}$. Similar assumptions have previously been made by Liu and Morse in [126] with generalizable results.

Lin [35] provides a graph-theoretic definition of what is called "structural" controllability which generalizes controllability to "structured" matrices which have fixed zero entries and free parameters everywhere else. A structurally controllable system is defined by the following equivalent statements:

1. The pair $(\mathbf{A}, \mathbf{B})$ is structurally controllable.

2. There is no permutation matrix, $\mathbf{P}$, that can bring the pair $(\mathbf{A}, \mathbf{B})$ into (2.16).

$$\mathbf{PAP}^{-1} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \ \mathbf{PB} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_2 \end{bmatrix} \tag{2.16}$$

   equivalently, (2.16) is also a reducible matrix[2].

3. The graph of $(\mathbf{A}, \mathbf{B})$ contains no nonaccessible vertices and no dilations[35].

4. The graph of $(\mathbf{A}, \mathbf{B})$ is spanned by a cactus[35].

[2]To avoid confusion with reducible and irreducible polynomials later in the work, we refer to reducible matrices as having Form (2.16).

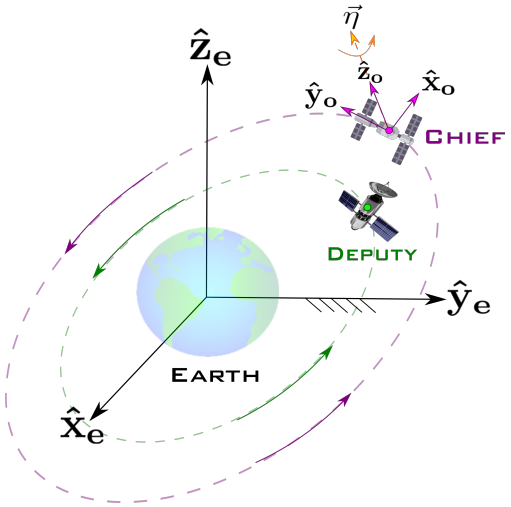
5. The generalized rank of the matrix $[\mathbf{A}\ \mathbf{B}]$ is n.

The generic rank, s-rank, or term rank is the maximum number of elements contained in at least one set of independent entries of $\mathbf{A}$ [127]. Physically, the generic rank places an upper-bound on the rank of a matrix of that particular structure.

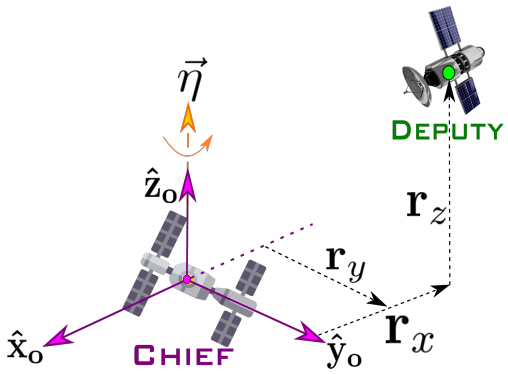
2.6 Relative Orbital Mechanics

2.6.1 Clohessy-Wiltshire-Hill Equations

The Clohessy-Wiltshire-Hill (CWH) equations [128] have become fundamental in describing the relative orbital dynamics of multi-satellite configurations operating within close proximity of a circular orbital trajectory of a "chief" satellite, whether virtual or real. See Figure 2.4. By leveraging the non-inertial Hill's frame [129], the CWH equations



(a) The Earth-Centered Inertial frame.



(b) The CWH frame.

Figure 2.4 A chief satellite and a deputy satellite both orbit about the Earth relative to the inertial frame $\{\hat{\mathbf{x}}_\mathbf{e}, \hat{\mathbf{y}}_\mathbf{e}, \hat{\mathbf{z}}_\mathbf{e}\}$ as shown in (a). The dashed lines correspond to the closed orbital trajectories of both spacecraft. The orbit-fixed CWH frame $\{\hat{\mathbf{x}}_\mathbf{o}, \hat{\mathbf{y}}_\mathbf{o}, \hat{\mathbf{z}}_\mathbf{o}\}$ is attached to the chief and rotates relative to the inertial frame at η rad/s. In (b), the position of the deputy relative to the chief is shown in this moving CWH frame.

provide a linear approximation of the dynamics of such systems and allow for associated linear analysis techniques to be applied. The state of a satellite as expressed in the Hill's frame, $\mathbf{x} := [\mathbf{r}_x,\ \mathbf{r}_y,\ \mathbf{r}_z,\ \mathbf{v}_x,\ \mathbf{v}_y,\ \mathbf{v}_z]^\top$ are respectively *relative* positions and velocities from

the state of the chief satellite. The full CWH equations of motion are shown in Equation (2.17):

$$\begin{aligned}\dot{\mathbf{v}}_x &= 2\eta\mathbf{v}_y + 3\eta^2\mathbf{r}_x + \mathbf{u}_x \\ \dot{\mathbf{v}}_y &= -2\eta\mathbf{v}_x + \mathbf{u}_y \\ \dot{\mathbf{v}}_z &= -\eta^2\mathbf{r}_z + \mathbf{u}_z\end{aligned} \tag{2.17}$$

where η, denoted the mean motion, is the constant angular velocity of the Hill's frame with respect to some inertial frame attached the celestial body (i.e. Earth), while $\mathbf{u} := [\mathbf{u}_x,\ \mathbf{u}_y,\ \mathbf{u}_z]^\top$ is the acceleration input applied to the spacecraft as measured in the Hill's frame. These equations may be placed in the form of a standard linear system, $\dot{\mathbf{x}} = \mathbf{\Phi}\mathbf{x} + \mathbf{B}\mathbf{u}$, $\mathbf{y} = \mathbf{C}\mathbf{x}$ where the state matrix $\mathbf{\Phi}$, control matrix $\mathbf{B}$ are defined as

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3\eta^2 & 0 & 0 & 0 & 2\eta & 0 \\ 0 & 0 & 0 & -2\eta & 0 & 0 \\ 0 & 0 & -\eta^2 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.18}$$

For multi-satellite systems, the above linear description may be expanded as $\dot{\mathbf{x}} = \bar{\mathbf{\Phi}}\mathbf{x} + \bar{\mathbf{B}}\mathbf{u}$ and $\mathbf{y} = \bar{\mathbf{C}}\mathbf{x}$, where, for homogeneous graph networks, $\bar{\mathbf{\Phi}} = (\mathbf{I}_N \otimes \mathbf{\Phi})$, $\bar{\mathbf{B}} = (\mathbf{I}_N \otimes \mathbf{B})$, and $\bar{\mathbf{C}} = (\mathbf{I}_N \otimes \mathbf{C})$ Where $\otimes$ denotes the Kronecker product, $\mathbf{I}_N$ is the N-dimensional identity matrix, and now, the state vector $\mathbf{x}$ is the augmentation of each of the individual satellite state vectors, $\mathbf{x} = [\mathbf{x}_1^\top,\ \ldots,\ \mathbf{x}_N^\top]^\top$.

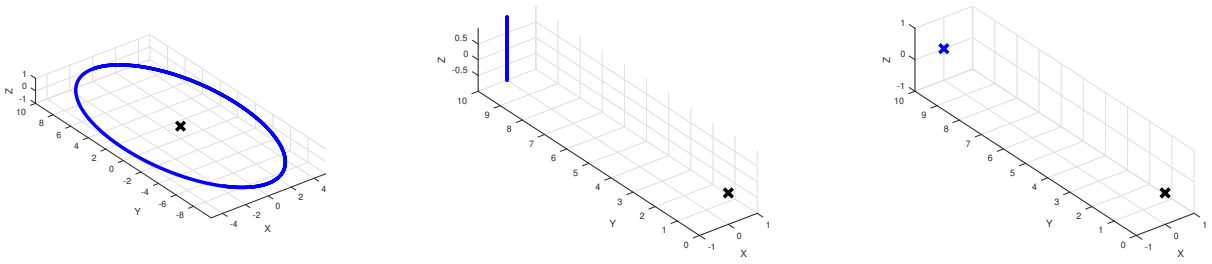
2.6.2 Natural Motion Trajectories

A class of periodic solutions exist to the CWH equations for when $\mathbf{u} = 0$. This particular class of solutions are often referred to as Natural Motion Trajectories (NMTs) or

Passive Relative Orbits (PROs) and have been extensively used for a wide variety of relative orbital motion problems [66, 130–136]. First pointed out in [137] and expanded upon in [138], a specific set of closed NMTs can take the form of stationary points, lines, ellipses, or spirals depending on the inital conditions of the satellite. All of the forms must first meet the initial condition $\mathbf{v}_y = -2\eta\mathbf{r}_x$. Additional specific initial conditions for each NMT form is specified below:

1. *Stationary Point*: $\mathbf{r}_x = \mathbf{r}_z = \mathbf{v}_x = \mathbf{v}_z = 0$
2. *Line Segment*: $\mathbf{r}_x = \mathbf{v}_x = 0$, $\mathbf{r}_z = c\sin(\psi)$, and $\mathbf{v}_z = \eta c\cos(\psi)$
3. *Ellipse*: $\mathbf{v}_x = \frac{\eta}{2}\mathbf{r}_y$

Where c is the amplitude of the length segment in the $\hat{z}$ direction, and ψ is the initial angle between the agent and the $\mathbf{x} - y$ plane. Examples of each can be found in Figure 2.5.



(a) Elliptical *(b)* Periodic Line Segment *(c)* Stationary

Figure 2.5 Types of Natural Motion Trajectories

2.7 Attitude dynamics

Consider each agent spacecraft to be on a NMT around a virtual chief that is in a circular orbit centered on the earth. While the agents do not have thrusters for translational motion, they can change their orientation using gas jet thrusters that produce external torques about each agent's center of mass. Each agent's attitudinal state is given by the Euler angle set $\mathbf{r}_i(k) := [\phi,\ \psi,\ \gamma]^\top$, respectively denoted yaw, pitch, and roll, that parameterizes the rotation matrix from the body frame to the inertial frame

$\mathbf{R}_{\mathscr{B}}^{\mathscr{H}} = \mathbf{R}_{\mathscr{B}}^{\mathscr{H}}(\gamma))\mathbf{R}_{\mathscr{B}}^{\mathscr{H}}(\psi)\mathbf{R}_{\mathscr{B}}^{\mathscr{H}}(\phi)$ using the 3-2-1 rotation sequence. The analogous angular velocities are $\boldsymbol{\omega}_{\mathscr{I}\mathscr{B}}^{\mathscr{B}} := [\omega_\phi,\ \omega_\psi,\ \omega_\gamma]^\top$. The agent's attitude control inputs are given by the torques in the body frame $\boldsymbol{\tau}^{\mathscr{B}} := [\tau_\phi,\ \tau_\psi,\ \tau_\gamma]^\top$. The rotation kinematics for the agent's body-fixed frame $\mathscr{B}$ is given by

$$\dot{\mathbf{R}}_{\mathscr{B}}^{\mathscr{H}} = \mathbf{R}_{\mathscr{B}}^{\mathscr{H}}(\boldsymbol{\omega}_{\mathscr{H}\mathscr{B}}^{\mathscr{B}})^\times. \tag{2.19}$$

The Euler angle set (locally) evolves according to

$$\dot{\Gamma} = \begin{bmatrix} -\cos(\phi)\tan(\psi) & -\sin(\phi)\tan(\psi) & -1 \\ \sin(\phi) & -\cos(\phi) & 0 \\ -\cos(\phi)\sec(\psi) & -\sin(\phi)\sec(\psi) & 0 \end{bmatrix} \boldsymbol{\omega}_{\mathscr{I}\mathscr{B}}^{\mathscr{B}}. \tag{2.20}$$

Assuming that the agent spacecraft can be modeled as a rigid body and has an inertia matrix $\boldsymbol{\mathcal{J}}$ as measured in the body-fixed frame, the application of Euler's second law of motion gives the attitude dynamics

$$\dot{\boldsymbol{\omega}} = -\boldsymbol{\mathcal{J}}^{-1}(\boldsymbol{\omega}_{\mathscr{I}\mathscr{B}}^{\mathscr{B}})^\times[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}_{\mathscr{I}\mathscr{B}}^{\mathscr{B}}] + \boldsymbol{\mathcal{J}}^{-1}\boldsymbol{\tau}^{\mathscr{B}}. \tag{2.21}$$

The angular velocity of the $\mathscr{B}$ frame relative to Hill's frame $\mathscr{H}$ using the property of relative angular velocity summation is

$$\boldsymbol{\omega}_{\mathscr{H}\mathscr{B}}^{\mathscr{I}} = \boldsymbol{\omega}_{\mathscr{I}\mathscr{B}}^{\mathscr{I}} - \boldsymbol{\omega}_{\mathscr{I}\mathscr{H}}^{\mathscr{I}}$$

with the time rate derivatives given as

$$\dot{\boldsymbol{\omega}}_{\mathscr{H}\mathscr{B}}^{\mathscr{I}} = \dot{\boldsymbol{\omega}}_{\mathscr{I}\mathscr{B}}^{\mathscr{I}} - \dot{\boldsymbol{\omega}}_{\mathscr{I}\mathscr{H}}^{\mathscr{I}}. \tag{2.22}$$

Using (2.19) and $\boldsymbol{\omega}^{\mathcal{B}}_{\mathcal{A}\mathcal{B}} = \mathbf{R}^{\mathcal{B}}_{\mathcal{A}}\boldsymbol{\omega}^{\mathbf{A}}_{\mathbf{A}\mathcal{B}}$ a manipulation of (2.22) yields

$$\dot{\boldsymbol{\omega}}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}} = \dot{\boldsymbol{\omega}}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}} - \mathbf{R}^{\mathscr{B}}_{\mathscr{H}}\dot{\boldsymbol{\omega}}^{\mathscr{H}}_{\mathscr{I}\mathscr{H}} - (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}})^{\times}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}}. \tag{2.23}$$

Note that from the circular chief orbit assumption we have $\dot{\boldsymbol{\omega}}^{\mathscr{H}}_{\mathscr{I}\mathscr{H}} = \mathbf{0}$ and that $\boldsymbol{\omega}^{\mathscr{H}}_{\mathscr{I}\mathscr{H}} = (0\ 0\ \eta)^{\top}$ is the constant angular velocity of the Hill's frame relative to the inertial frame where η is the constant mean motion parameter. From the assumption that the chief-fixed inertia matrix is aligned with its principal axes we have

$$\boldsymbol{\mathcal{J}} = \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{bmatrix}.$$

The combination of (2.21) and (2.23) yields the angular velocity dynamics as

$$\dot{\boldsymbol{\omega}}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}} = -\boldsymbol{\mathcal{J}}^{-1}(\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}})^{\times}[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}}] + \boldsymbol{\mathcal{J}}^{-1}\tau^{\mathscr{B}} - (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}})^{\times}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}}.$$

Using the relation $\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{B}} = \boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}} + \boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}}$ and cross product properties the above expression becomes

$$\begin{aligned}\dot{\boldsymbol{\omega}}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}} = -\boldsymbol{\mathcal{J}}^{-1}\big((\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}})^{\times}[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}}] + (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}})^{\times}[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}}] + \\ (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}})^{\times}[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}}] + (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}})^{\times}[\boldsymbol{\mathcal{J}}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}}]\big) + (\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{H}\mathscr{B}})^{\times}\boldsymbol{\omega}^{\mathscr{B}}_{\mathscr{I}\mathscr{H}} + \boldsymbol{\mathcal{J}}^{-1}\boldsymbol{\tau}^{\mathscr{B}}.\end{aligned} \tag{2.24}$$

2.8 Model Predictive Control

Consider the following control system with state $\mathbf{x}(k) \in \mathbb{R}^n$ at timestep k that evolves according to a discrete-time system with process function $f : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ and

measurement function $h : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^p$ over timesteps $\{k, k+1, k+2, \dots\}$, $k \geq 0$ as

$$\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k)) \tag{2.25}$$

$$\mathbf{y}(k) = h(\mathbf{x}(k), \mathbf{u}(k)) \tag{2.26}$$

where $\mathbf{u}(k) \in \mathbb{R}^q$ is the input applied to the plant while $\mathbf{y}(k) \in \mathbb{R}^p$, $p \leq n$, is the output of the system. Without loss of generality, we assume $f(0,0) = 0$ and $c(0,0) = 0$ for all $k \geq 0$. Let $T \in \mathbb{N}$, $T \geq 2$ be the finite horizon length. We define a control policy $\bar{\mathbf{u}}(k) := \{\mathbf{u}(0), \mathbf{u}(1), \dots, \mathbf{u}(s), \dots, \mathbf{u}(T-1)$ at time k with an additional index $s = 0, \dots, T-1$. The output trajectory resulting from $\bar{\mathbf{u}}(k)$ according to (2.25)-(2.26) is $\bar{\mathbf{y}}(k) := \{\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(s), \dots, \mathbf{y}(T-1)\}$. Let $\bar{\mathbf{y}}_i^s(k)$ and $\bar{\mathbf{u}}_i^s(k)$ denote particular elements of $\bar{\mathbf{y}}(k)$ and $\bar{\mathbf{u}}(k)$ at index s and agent i given an initial condition $\mathbf{x}(k)$, respectively. Let there be some nonnegative function $h : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}_{\geq 0}$, known as a stage cost, and a control objective function $J : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}_{\geq 0}$ defined as

$$J(\mathbf{y}(k), \mathbf{u}(k)) := \sum_{s=0}^{T-1} h(\mathbf{y}_s(k), \mathbf{u}_s(k)). \tag{2.27}$$

The goal of a model predictive controller is to, at timestep k, generate an *optimal* control sequence $\bar{\mathbf{u}}^*(k) := \{\mathbf{u}^*(0), \mathbf{u}^*(1), \dots, \mathbf{u}^*(s), \dots, \mathbf{u}^*(T-1)\}$ such that the objective function J from (2.27) is minimized. The first input in this sequence, $\mathbf{u}^*(k)(0)$, is applied to the system. At timestep $k+1$, this minimization process is repeated and $\mathbf{u}^*(k)(0)$ is applied.

This is formally written as

$$\begin{aligned}
&\underset{\mathbf{u}(k)}{\text{minimize}}\ J(\mathbf{y}(k), \mathbf{u}(k)) \\
&\text{subject to } \mathbf{x}(0) = \mathbf{x}(k) \\
&\qquad \mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k)), \quad k = 0, \ldots, T-1 \\
&\qquad \mathbf{y}(k) = c(\mathbf{x}(k), \mathbf{u}(k)), \quad k = 0, \ldots, T-1 \\
&\qquad (\mathbf{x}(k), \mathbf{u}(k)) \in \mathcal{X} \times \mathcal{U}, \quad k = 0, \ldots, T-1
\end{aligned} \tag{2.28}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is the closed set of state constraints and $\mathcal{U}$ is the closed set of input constraints. Generally, the stage cost function h is chosen to represent a distance of the output from some predefined desired output trajectory, i.e. $\mathbf{y}^*(k)$.

2.9 Information Fusion

The simplest information fusion problem consists of at least two pieces of information, which this work considers, labelled $\mathcal{A}$ and $\mathcal{B}$. Now suppose each of the pieces of information are corrupted by noise to form Gaussian random variable $\mathbf{a}$ and $\mathbf{b}$ from $\mathcal{A}$ and $\mathcal{B}$ respectively. The statistics of $\mathbf{a}$ and $\mathbf{b}$ are given as the mean ($\bar{\mathbf{a}}$, $\bar{\mathbf{b}}$) and covariance ($\mathbf{P}_a$, $\mathbf{P}_b$) and sometimes, the cross-correlation term ($\mathbf{P}_{ab}$) where

$$\begin{aligned}
\mathbf{P}_a &= \mathbb{E}\big[(\mathbf{a} - \bar{\mathbf{a}})(\mathbf{a} - \bar{\mathbf{a}})^T\big] \\
\mathbf{P}_b &= \mathbb{E}\big[(\mathbf{b} - \bar{\mathbf{b}})(\mathbf{b} - \bar{\mathbf{b}})^T\big] \\
\mathbf{P}_{ab} &= \mathbb{E}\big[(\mathbf{a} - \bar{\mathbf{a}})(\mathbf{b} - \bar{\mathbf{b}})^T\big]
\end{aligned} \tag{2.29}$$

where $\mathbb{E}[\cdot]$ is the expectation operator with $\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i p_i$ and x_i is an element of X and p_i is the probability of occurrence for x_i. It is important to note that the cross-correlation term, $\mathbf{P}_{ab}$, is generally unknown. If the cross-correlation term is known, the optimal information fusion solution can be found through the Kalman filtering algorithm. So effectively, the information fusion problem and Kalman filtering problems are focusing on disjoint cases of the estimation/fusion problem as indicated by [139].

The objective of the information fusion problem is to find the fused vector $\mathbf{c}$ that is the combination of the information provided by $\mathbf{a}$ and $\mathbf{b}$, with partially or unknown cross-correlations, such cases where the Kalman filter will be inconsistent.

Definition 5. ***(Consistency)*** *- A covariance ellipsoid is said to be consistent if* $\boldsymbol{P} \geq \mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T]$*, where* $\mathbb{E}[\boldsymbol{e}\boldsymbol{e}^T]$ *is the actual error covariance matrix and provides a lower-bound on the known covariance matrix.*

Where consistency provides a lower-bound on the fused solution, the notion of tightness provides an upper-bound.

Definition 6. ***(Tightness)*** *- The known covariance* $\boldsymbol{P}$ *is said to be tight if* $\mathcal{E}(0, \boldsymbol{P}) \supseteq \mathcal{E}(0, \boldsymbol{\Lambda}) \supseteq \mathcal{E}(0, \boldsymbol{P}_a) \cap \mathcal{E}(0, \boldsymbol{P}_b) \Rightarrow \boldsymbol{P} = \boldsymbol{\Lambda}$*. In other words, the ellipse formed by* $\boldsymbol{P}$ *encompasses the intersection of the ellipses formed by* $\boldsymbol{P}_a$ *and* $\boldsymbol{P}_b$ *so closely that no other ellipse formed by* $\boldsymbol{\Lambda}$ *may be slipped in between.*

where $\mathcal{E}(0, \mathbf{P})$ denotes the ellipse centered at 0 and shaped by the matrix $\mathbf{P}$. Together, consistency and tightness restrict the class of possible solutions around the optimal fused solution and provide useful notions that constrain the overall problem.

Information fusion of two information pieces is most often accomplished using the linear combination

$$\bar{\mathbf{c}} = \mathbf{W}_a \bar{\mathbf{a}} + \mathbf{W}_b \bar{\mathbf{b}} \tag{2.30}$$

where $\mathbf{W}$ is the fusion gains for $\mathbf{a}$ and $\mathbf{b}$, respectively. The corresponding error covariance matrix is given by

$$\begin{aligned} \mathbf{P}_c = &\mathbf{W}_a \mathbf{P}_a \mathbf{W}_a^T + \mathbf{W}_a \mathbf{P}_{ab} \mathbf{W}_b^T + \\ &\mathbf{W}_b \mathbf{P}_{ba} \mathbf{W}_a^T + \mathbf{W}_b \mathbf{P}_b \mathbf{W}_b^T \end{aligned} \tag{2.31}$$

For random vectors with no cross correlation, $\mathbf{P}_{ab} = 0$, Equation 2.31 is guaranteed to maintain consistency. Alternatively, if the cross-correlation of the random vectors is known,

the optimal solution is provided by a convex combination of the covariance matrices and cross-correlations as

$$\begin{aligned} \mathbf{P}_c^{-1} &= \mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1} \\ \bar{\mathbf{c}} &= \mathbf{P}_c\Big(\mathbf{P}_a^{-1}\mathbf{a} + \mathbf{P}_b^{-1}\mathbf{b} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1}\boldsymbol{\gamma}\Big) \end{aligned} \tag{2.32}$$

where $\boldsymbol{\gamma}$ is the mutual mean that represents the estimate that $\mathbf{a}$ and $\mathbf{b}$ agree on the most. This formulation may be viewed as a direct implementation eliminating double-counting of information, and is used as the optimal solution throughout the rest of the work. However, the same guarantee does not hold for when $\mathbf{P}_{ab} \neq 0$. To address this issue, parameterizations of the mutual information provided by $\mathbf{P}_{ab}$ have been developed that guarantee consistency.

Tangentially, consider the joint distribution between the two random variables $\mathbf{a}$ and $\mathbf{b}$.

$$\mathbf{P}_j = \begin{bmatrix} \mathbf{P}_a & \mathbf{P}_{ab} \\ \mathbf{P}_{ba} & \mathbf{P}_b \end{bmatrix} \tag{2.33}$$

It is assumed that there exists matrix parameters $\mathbf{A},\ \mathbf{B},\ \mathbf{M}$ such that $\mathbf{P}_{ab}$ is delimited by

$$\mathbf{P}_{ab} \in \{\mathbf{M} + \mathbf{A}\boldsymbol{\Omega}\mathbf{B}^T \mid \boldsymbol{\Omega}\boldsymbol{\Omega}^T \leq \mathbf{I}\} \tag{2.34}$$

where $\boldsymbol{\Omega}$ is an unknown admissible[3] tuning parameter, and $\mathbf{M}$ encapsulates any known cross-correlation components. In this work, we take advantage of the currently available state-of-the-art from which we begin the construction of $\boldsymbol{\Omega}$. Therefore, we are guaranteed at least admissible $\boldsymbol{\Omega}$ exists because, as is shown later in this work, that $\boldsymbol{\Omega}$ may be specifically constructed to result in one of the current state-of-the-art solutions. This approach follows directly from [139–141] that apply decomposition and factorization

[3]Admissibility refers to any matrix $\boldsymbol{\Omega}$ that results in the term $\mathbf{A}\boldsymbol{\Omega}\mathbf{B}^T$ meeting the criterion for a covariance matrix, positive semi-definite and symmetric. Calculating admissible $\boldsymbol{\Omega}$ is outside the scope of this work and will be the subject of future investigation, however it is assumed for the remainder of this work that $\boldsymbol{\Omega}$ is admissible.

approaches to the joint distribution to define a family of upper-bounds on the joint solution. It should also be noted that this particular family of covariance matrices have known upper ($\mathbf{\Pi}_u$) and lower-bounds ($\mathbf{\Pi}_l$) of the joint distribution

$$\begin{aligned}\mathbf{\Pi}_u &= \begin{bmatrix}\mathbf{P}_a + \mathbf{A}\mathbf{A}^T\mu & \mathbf{M} \\ \mathbf{M}^T & \mathbf{P}_b + \mathbf{B}\mathbf{B}^T\mu^{-1}\end{bmatrix} \\ \mathbf{\Pi}_l &= \begin{bmatrix}\mathbf{P}_a - \mathbf{A}\mathbf{A}^T\gamma & \mathbf{M} \\ \mathbf{M}^T & \mathbf{P}_b - \mathbf{B}\mathbf{B}^T\gamma^{-1}\end{bmatrix}\end{aligned} \tag{2.35}$$

where μ and γ are free positive scalar parameters, such that $\mathbf{\Pi}_u \geq \mathbf{P}_j$ and $\mathbf{P}_j \geq \mathbf{\Pi}_l$ where $\geq$ is considered in the sense that the difference between the two terms is positive semi-definite. This inequality may also be written as $\mathbf{\Pi}_u - \mathbf{P}_j \geq 0$ and $\mathbf{P}_j - \mathbf{\Pi}_l \geq 0$. Analysis of these bounds can be found abundantly in literature for both partially and unknown cross-correlations [140, 142–146]. In particular, these scalar parameterizations are possible when $\mathbf{A}$ and $\mathbf{B}$ are fractions of the square root, so $k_a\mathbf{P}_a = k_a\mathbf{L}_a\mathbf{L}_a^T = \mathbf{A}\mathbf{A}^T$ and $k\mathbf{P}_b = k_b\mathbf{L}_b\mathbf{L}_b^T = \mathbf{B}\mathbf{B}^T$, with $0 \leq k_a, k_b \leq 1$ which induces $\mathbf{A} = \sqrt{k_a}\mathbf{L}_a$ and $\mathbf{B} = \sqrt{k_b}\mathbf{L}_b$.

2.9.1 Covariance Intersection

Alternatively to Equation 2.31, as first pointed out by Julier and Uhlmann [106], this same calculation may be completed on the information space (which is the inverse of the covariance matrix) and provides a consistent result without explicit knowledge of $\mathbf{P}_{ab}$. This method, known as Covariance Intersection (CI), has become a seminal piece of information fusion,

$$\begin{aligned}\mathbf{P}_c^{-1} &= \omega\mathbf{P}_a^{-1} + (1-\omega)\mathbf{P}_b^{-1} \\ \bar{\mathbf{c}} &= \mathbf{P}_c(\omega\mathbf{P}_a^{-1}\bar{\mathbf{a}} + (1-\omega)\mathbf{P}_b^{-1}\bar{\mathbf{b}})\end{aligned} \tag{2.36}$$

with $0 \leq \omega \leq 1$. This method of using a convex combination of information $\mathcal{A}$ and $\mathcal{B}$ is also proven to be tight [139, 147] because it prevents double-counting of the mutual information

[148]. The effect that ω has on the CI fused solutions is shown in Figure 2.6. As ω transitions from 0 to 1, the fused solution is being pulled to either the information provided by $\mathcal{A}$ or to the information provided by $\mathcal{B}$. This effect limits the tightness the CI solution is capable of achieving as evident in Figure 2.6.

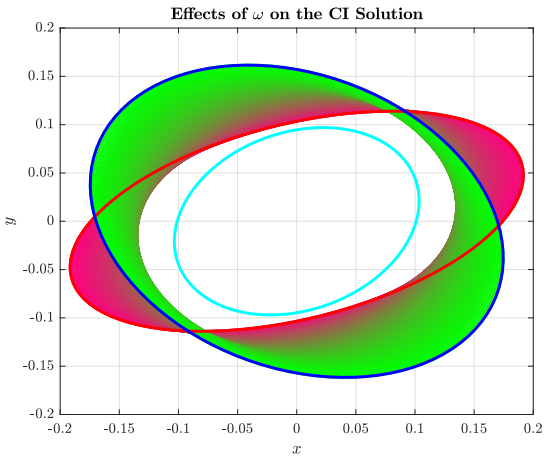


Figure 2.6 Demonstration of how ω influences the CI solution by fusing Covariance Ellipse 1 with Covariance Ellipse 2. ω is allowed to vary between 0 and 1. The Optimal Solution is included for reference.

Similarly, a criticism of traditional CI, while consistent and tight, may be too conservative [107] meaning it over-approximates the true upper-bound on the uncertainty. In light of this, alternatives to CI have been developed to define an upper-bound on the fused solution.

2.9.2 Ellipsoidal Intersection

In an effort to solve this issue of over-conservatism of CI, Sijs, Lazar, and Bosch in [108] present a parameterized form of the information fusion problem

$$\begin{aligned} \mathbf{P}_c^{-1} &= \mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1} \\ \bar{\mathbf{c}} &= \mathbf{P}_c(\mathbf{P}_a^{-1}\bar{\mathbf{a}} + \mathbf{P}_b^{-1}\bar{\mathbf{b}} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1}\boldsymbol{\gamma}) \end{aligned} \tag{2.37}$$

with the goal to find a construction of the mutual information matrix $\mathbf{P}_{ab}^{-1}$ and mutual information mean $\boldsymbol{\gamma}$ from known information. The Ellipsoidal Intersection (EI) procedures

provides explicit approximations for $\mathbf{P}_{ab}$ and $\boldsymbol{\gamma}$.

$$\begin{aligned}\mathbf{P}_{ab} + \mathbf{P}_{ba} &= \mathbf{T}\mathbf{D}_{ab}\mathbf{T}^T \\ \boldsymbol{\gamma} &= (\mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1} + 2\eta\mathbf{I})^{-1} \\ &\times \left((\mathbf{P}_b^{-1} - \mathbf{P}_{ab}^{-1} + \eta\mathbf{I})\bar{\mathbf{a}} + (\mathbf{P}_a^{-1} - \mathbf{P}_{ab}^{-1} + \eta\mathbf{I})\bar{\mathbf{b}}\right)\end{aligned} \tag{2.38}$$

where $\mathbf{T} = \mathbf{V}_a\mathbf{D}_a^{1/2}\mathbf{V}_b$ with $\mathbf{V}$ and $\mathbf{D}$ being the eigenvectors and matrix of eigenvalues of the covariance matrices $\mathbf{P}_a$ and $\mathbf{P}_b$, respectively. η is a scaling parameter. The authors refer the interested reader to [108] for more detailed explanations of $\mathbf{T}$, $\mathbf{D}$, and η.

An example of EI is provided in Figure 2.7. From the figure, it is evident that the EI solution does not maintain consistency as it lies within the optimal solution.

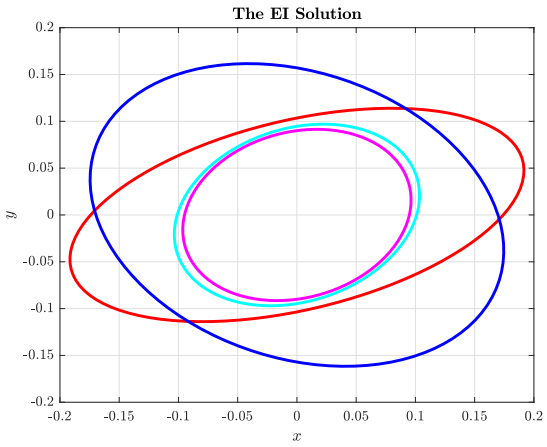


Figure 2.7 Demonstration the EI solution by fusing Covariance Ellipse 1 with Covariance Ellipse 2. The Optimal Solution is included for reference.

This point is analytically shown in [109, 110] that the EI result does not necessarily guarantee consistency. This fact prompts Ajgl in [111] to derive a mathematical tool for fusion performance analysis under partially known correlations. This tool allows for explicit expressions for upper and lower bounds on the fused solution to be defined–a fact taken advantage of in the present work.

2.9.3 Inverse Covariance Intersection

Where CI is consistent, yet too conservative, and EI does not maintain consistency, Noack [109] derives an alternative fusion methodology termed Inverse Covariance Intersection (ICI) and shows it to be less conservative but still maintains consistency.

$$\begin{aligned} \mathbf{P}_c^{-1} &= \mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1} \\ \bar{\mathbf{c}} &= \mathbf{W}_a\bar{\mathbf{a}} + \mathbf{W}_b\bar{\mathbf{b}} \end{aligned} \tag{2.39}$$

where,

$$\begin{aligned} \mathbf{W}_a &= \mathbf{P}_c(\mathbf{P}_a^{-1} - \omega(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1}) \\ \mathbf{W}_b &= \mathbf{P}_c\Big(\mathbf{P}_b^{-1} - (1-\omega)(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1}\Big) \end{aligned} \tag{2.40}$$

While ω can be set to minimize any given optimality criterion, such as the trace or determinant of the bounding covariance matrix, evaluating and selecting the optimal value of ω is a cumbersome task and may have to be repeated for evolving systems. Consequently, selecting a non-optimal value of ω will bias the solution towards either the information provided by $\mathcal{A}$ or $\mathcal{B}$, skewing the fused estimate away from the optimal solution. It could also place an over-consistent lower-bound on the attainable solutions from ICI.

The potential bias can be seen in Figure 2.8 as the fused solutions transition from magenta to green. As $\omega \to 1$ the fused solution swings from the information provided by $\mathcal{A}$ to the information provided by $\mathcal{B}$. Because the fused solution is constantly being pulled towards one of the two information pieces, the best fused solution is limited to the "equilibrium" of the two pieces of information.

Comparing the ICI solution directly to the CI solution, the ICI solution is capable of providing an overall tighter approximation of the optimal solution. There are also a wider range of values for ω that provide tighter fused solutions from the decreased density of ellipses present near the original covariance matrices.

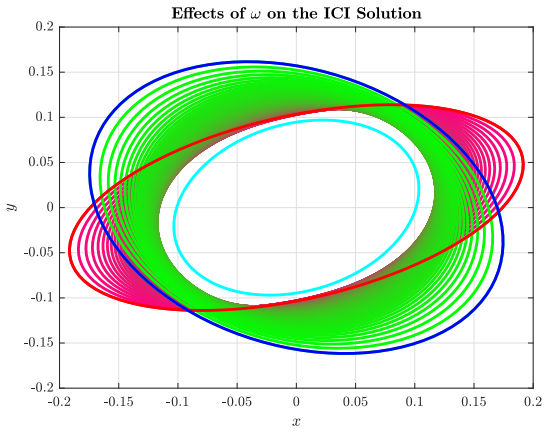


Figure 2.8 Demonstration of how ω influences the ICI solution by fusing Covariance Ellipse 1 with Covariance Ellipse 2. ω is allowed to vary between 0 and 1. The Optimal Solution is included for reference.

3 Local Catalog Maintenance of Close Proximity Satellite Systems

3.1 Introduction

In this chapter, we consider the case of a single controllable chief satellite equipped with an electro-optical sensor and on-board estimating capabilities that must track and maintain the states of several uncontrollable deputy satellites, each of which lay on a closed, elliptical natural motion trajectory [138] about the chief. The local catalog we wish to maintain is the concatenated list of all the deputies' positional and velocity states relative to the chief. However, a deputy's state is inherently uncertain. Generally, these state uncertainties are represented through stochastic parameters, such as a mean and state covariance matrix pair for a Gaussian probability distribution, also known as the "belief state." Through the use of a Bayesian filter [149], the fusion between a deputy's prior belief state and an observation of that deputy results in a posterior belief state with reduced uncertainty. Crucially, the chief cannot observe the states of all deputies simultaneously and is only able to make measurements of deputies when they lie within the chief's field-of-view. Thus, the uncertainties of unobserved agents will grow accordingly with the amount of time elapsed without measurement. This engenders the core of the local catalog maintenance problem: how does the chief know when and where to look in order to provide a catalog of deputy states that are closest to their true values?

3.2 Problem Statement

Consider a chief spacecraft with state space $\mathcal{X}_0 \subseteq \mathbb{R}^n$ and input space $\mathcal{U} \subseteq \mathbb{R}^q$. The state $\mathbf{x}_0 \in \mathcal{X}_0$ evolves according to $\dot{\mathbf{x}}_0 = f(\mathbf{x}_0, \mathbf{u})$ where $\mathbf{u} \in \mathcal{U}$ is the chief's input, and $f : \mathcal{X}_0 \mapsto \mathcal{X}_0$ is the chief's dynamic function. Additionally, consider some constraint function $c : \mathcal{X}_0 \times \mathcal{U} \to \mathbb{R}^r$. A state $\mathbf{x}_0 \in \mathcal{X}_0$ and input $\mathbf{u} \in \mathcal{U}$ satisfies the constraints when $c_l(\mathbf{x}_0, \mathbf{u}) \leq 0$ for each row $l \in \{1, \ldots, r\}$ in the constraint vector. Given an initial state $\mathbf{x}_0(0) \in \mathcal{X}_0$ and an input function $u_c : \mathbb{R}_{\geq 0} \to \mathcal{U}$, a *trajectory* of the chief is given by $\xi_{u_c, \mathbf{x}_0(0)} : \mathbb{R}_{\geq 0} \to \mathcal{X}$ and satisfies $\xi_{u_c, \mathbf{x}_0(0)}(0) = \mathbf{x}_0(0)$ and $\frac{d}{dt}\xi_{u_c, \mathbf{x}_0(0)}(t) = f(\xi_{u_c, \mathbf{x}_0(0)}(t), \mathbf{u}(t))$.

The chief is tasked with maintaining a catalog of N deputies. Each deputy

$i \in \{1, \ldots, N\}$ has a state space $\mathcal{X}_i \subseteq \mathbb{R}^n$. A deputy's state $\mathbf{x}_i \in \mathcal{X}_i$ evolves according to $\dot{\mathbf{x}}_i = \mathbf{\Phi}\mathbf{x}_i$ where $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ is the system matrix. The state of all deputies, or system state, is given by $\mathbf{x} := [\mathbf{x}_1^\top \ \ldots \ \mathbf{x}_N^\top]^\top$ and evolves according to $\dot{\mathbf{x}} = (\mathbf{I}_N \otimes \mathbf{\Phi})\mathbf{x}$.

3.2.1 Chief Estimator

The chief is equipped with a recursive state estimator to track N deputy belief states, parameterized by concatenated mean $\hat{\mathbf{x}} := [\mathbf{x}_1^\top \ \cdots \ \mathbf{x}_N^\top]^\top$ and covariance block matrix $\mathbf{P} := \text{diag}\{\mathbf{P}_1, \ldots, \mathbf{P}_N\}$. These constructs evolve via their assumed dynamics models according to

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= (\mathbf{I}_N \otimes \mathbf{\Phi})\hat{\mathbf{x}} \\ \dot{\mathbf{P}} &= (\mathbf{I}_N \otimes \mathbf{\Phi})\mathbf{P}(\mathbf{I}_N \otimes \mathbf{\Phi})^\top + \mathbf{B}_\mathbf{Q}\mathbf{Q}^{proc}\mathbf{B}_\mathbf{Q}^\top, \end{aligned} \tag{3.1}$$

where, $\mathbf{Q}^{proc} \in \mathbb{R}^{nN \times nN}$ is the process noise matrix and $\mathbf{B}_\mathbf{Q} \in \mathbb{R}^{nN \times nN}$ is the noise input matrix. We assume $\mathbf{B}_\mathbf{Q} = \mathbf{I}_{nN}$ for simplicity.

The chief is equipped with a sensor parameterized by its pointing vector in the chief's body frame $\mathbf{p}^{\mathscr{B}}$ and its angle of view α. The sensor can observe the states of the deputies within some sensing field-of-view (FOV) dependent on $\mathbf{p}^{\mathscr{B}}$, α, and the chief's state $\mathbf{x}_0$. The linear measurement model of the sensor is given by $\mathbf{y} = \mathbf{C}_\mathbf{A}\mathbf{x}$ where $\mathbf{y}$ is the observed state of deputies, and $\mathbf{C}_\mathbf{A}$ is the observation matrix dependent on the system state $\mathbf{x}$ and the chief's state $\mathbf{x}_0$. The observation matrix $\mathbf{C}_\mathbf{A}$ is a block matrix given as

$$\mathbf{C}_\mathbf{A} = \begin{bmatrix} \mathbf{C}_1(\mathbf{x}_1, \mathbf{x}_0) & \emptyset_{n\times n} & \cdots & \emptyset_{n\times n} \\ \emptyset_{n\times n} & \mathbf{C}_2(\mathbf{x}_2, \mathbf{x}_0) & \cdots & \emptyset_{n\times n} \\ \vdots & \vdots & \ddots & \vdots \\ \emptyset_{n\times n} & \emptyset_{n\times n} & \cdots & \mathbf{C}_N(\mathbf{x}_N, \mathbf{x}_0) \end{bmatrix}, \tag{3.2}$$

where each $\mathbf{C}_i(\mathbf{x}_i, \mathbf{x}_0)$ is given as

$$\mathbf{C}_i(\mathbf{x}_i, \mathbf{x}_0) = \begin{cases} \mathbf{I}_n & \text{if } \mathbf{x}_i \text{ is observed} \\ \\ \emptyset_{n\times n} & \text{else.} \end{cases} \tag{3.3}$$

It should be noted here that $\emptyset_{n\times n}$ is an overloaded operator. In the case that $\mathbf{C}_i(\mathbf{x}_i, \mathbf{R}^{\mathscr{I}}_{\mathscr{B}}) = \mathbf{I}_n$, $\emptyset_{n\times n}$ takes on the value of $\mathbf{0}_n$, and in the case when $\mathbf{C}_i(\mathbf{x}_i, \mathbf{R}^{\mathscr{I}}_{\mathscr{B}}) = \emptyset_{n\times n}$, $\emptyset_{n\times n}$ is the null set taking the form of an empty matrix.

Whenever the chief makes an observation of a deputy $i \in \{1, \ldots, N\}$ at a time instant k, the measurement data is incorporated by the on-board estimator to impulsively update the belief state such that

$$(\hat{\mathbf{x}}_i(k^-), \mathbf{P}_i(k^-)) \to (\hat{\mathbf{x}}_i(k^+), \mathbf{P}_i(k^+)).$$

As a result, the size of the covariance matrices of observed deputies will necessarily decrease $v(\mathbf{P}_i(k^+)) \leq v(\mathbf{P}_i(k^-))$ and the updated belief state will be recursively supplied to the propagator until the next available measurement update. Here $v : \mathbb{R}^{n\times n} \mapsto \mathbb{R}$ is any metric quantifying the size of the covariance matrix or its representative hyperellipsoid. However, because the sensor has a limited FOV, it may not be possible for the chief to observe every deputy at once. Therefore, the chief must maneuver to view each deputy to maintain a level of certainty in the state estimate, that is $v(\mathbf{P}_i(k)) \leq \epsilon$ for all $k \geq 0$, where ϵ is some positive value determined by the sensor FOV.

Then, the local catalog maintenance problem is the following. Consider a chief satellite with dynamic function f and constraint function c, a sensor parametrized by pointing vector $\mathbf{p}^{\mathscr{B}}$ and area of view α. Given N deputies with initial estimate $\hat{\mathbf{x}}(0) \in \mathcal{X}$ and initial covariances $\mathbf{P}(0) \in \mathbb{R}^{nN\times nN}$, and initial chief state $\mathbf{x}_0(0) \in \mathcal{X}_0$, find a control function $u_c : \mathcal{X}_0 \times \mathcal{X} \times \mathbb{R}^{nN\times nN} \times \mathbb{R}_{\geq 0} \to \mathcal{U}$ such that for all $k \geq 0$, $c_l(\xi_{u_c,\mathbf{x}_0(0)}(k), u_c(\xi_{u_c,\mathbf{x}_0(0)}(t))) \leq 0$ for all $l \in \{1, \ldots, r\}$, and $v(\mathbf{P}_j(k)) \leq \epsilon$.

3.2.2 Agent constraints

Each agent also has some constraints it must follow while performing the tracking. Typically, the torque inputs cannot exceed some $\tau_{\max}$, and the angular velocity must be maintained under some $\omega_{\max}$. To maintain safety, the sensor cannot point in the direction of the (assumed static) sun vector $\mathbf{s}$ given in $\mathscr{I}$. Recall that $\mathbf{R}_{\mathscr{B}}^{\mathscr{I}}$ denotes the rotation matrix from the agent's body-fixed frame $\mathscr{B}$ to the inertial frame $\mathscr{I}$. Then, the constraint function is given as

$$c(\mathbf{x}_i(k), \boldsymbol{\tau}) = \begin{bmatrix} c_1(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_2(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_3(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_4(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_5(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_6(\mathbf{x}_i(k), \boldsymbol{\tau}) \\ c_7(\mathbf{x}_i(k), \boldsymbol{\tau}) \end{bmatrix} = \begin{bmatrix} \mathbf{s}^\top(\mathbf{R}_{\mathscr{B}}^{\mathscr{I}}\mathbf{p}_{\mathscr{B}}) - \cos(\alpha) \\ \|\omega_1\| - \omega_{\max} \\ \|\omega_2\| - \omega_{\max} \\ \|\omega_3\| - \omega_{\max} \\ \|\tau_1\| - \tau_{\max} \\ \|\tau_2\| - \tau_{\max} \\ \|\tau_2\| - \tau_{\max} \end{bmatrix}. \tag{3.4}$$

If $c_1(\mathbf{x}_i(k), \boldsymbol{\tau}) \leq 0$, then the angle between $\mathbf{s}$ and $\mathbf{p}^{\mathscr{I}} = \mathbf{R}_{\mathscr{B}}^{\mathscr{I}}\mathbf{p}^{\mathscr{B}}$ is at least α, meaning the sun vector is not pointing in the sensor FOV. If $c_l(\mathbf{x}_i(k), \boldsymbol{\tau}) \leq 0$, $l \in \{2, 3, 4\}$, then $\|\omega_j\| \leq \omega_{\max}$ for $j \in \{1, 2, 3\}$. If $c_l(\mathbf{x}_i(k), \boldsymbol{\tau}) \leq 0$, $l \in \{5, 6, 7\}$, then $\|\tau_j\| \leq \tau_{\max}$ for $j \in \{1, 2, 3\}$.

3.2.3 Observation condition

Recall from before that the deputy belief states are updated whenever an observation is made. Given a deputy state $\mathbf{x}i$, chief state $\mathbf{x}_i(k)$, with a sensor characterized by the pointing vector in the body frame $\mathbf{p}^{\mathscr{B}}$ and the viewing angle α. An observation occurs if

$$\cos^{-1}(\mathbf{x}_i^\top \mathbf{R}_{\mathscr{B}}^{\mathscr{H}} \mathbf{p}^{\mathscr{B}})) \leq \alpha$$

where $\mathbf{R}_{\mathscr{B}}^{\mathscr{H}} = \mathbf{R}_{\mathscr{I}}^{\mathscr{H}} \mathbf{R}_{\mathscr{B}}^{\mathscr{I}\top}$.

3.3 Approach

To solve the local catalog maintenance problem, this work proposes a multi-layered closed-loop control strategy (see Fig.3.1) that incorporates continuous dynamics with impulsive zero-order hold controls applied at discrete time instances. That is, at a given time instant k, the state estimator fuses predicted deputy belief states $(\hat{\mathbf{x}}^{-}(k), \mathbf{P}^{-}(k))$ with available observation data $\mathbf{y}(k)$ to yield posterior deputy belief states $(\hat{\mathbf{x}}^{+}(k), \mathbf{P}^{+}(k)$ through a Bayesian update scheme such as the commonly-used extended Kalman filter or unscented Kalman filter.

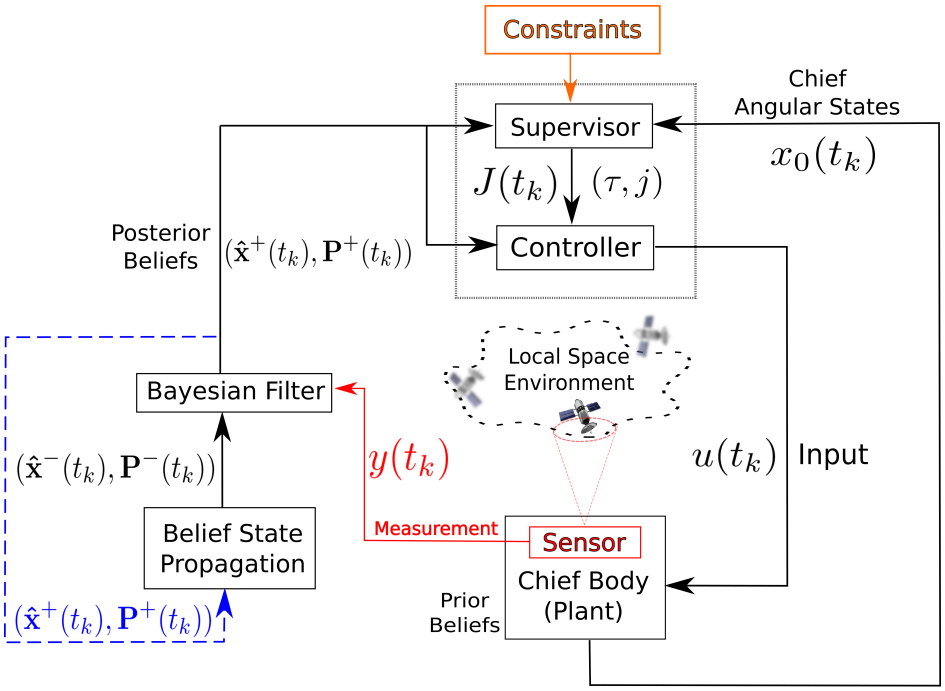


Figure 3.1 Visual depiction of the proposed catalog maintenance operation. Note that the sensor component is embedded within the plant component. The blue dashed line indicates the recursive nature of the Bayesian filter, where updated belief states are supplied to the propagation block for the next time step.

3.3.1 Supervisor Algorithm

These updated beliefs are fed into a higher-level autonomous "supervisor" which accounts for (i) deputy belief state information gaps (ii) prespecified constraints and (iii) the current angular states of the chief. The supervisor amalgamates this information and

provides decision-making capabilities to a lower-level controller, which drives the chief orientation trajectory by administering torque inputs at discrete timesteps. Recall the local catalog maintenance problem where a chief, with a sensor parametrized by $\mathbf{p}^{\mathscr{B}}$ and α, must maintain state estimates of N deputies. The objective of the supervisor algorithm is to select the "optimal" target deputy to track based on some cost function notated by $\mathcal{J}$. A detailed breakdown of the supervisor algorithm can be seen in Algorithm 1. Given the state estimates-covariance pairs of the deputies $\{(\hat{\mathbf{x}}_i, \mathbf{P}_i)\}_{i=1}^{N}$ and the current deputy j. Additionally, we introduce a new variable Δt^* which "remembers" the time that the supervisor switched to deputy j. The output is the deputy j_{new} to be viewed as well as Δt^*_{new}. Then, the supervisor algorithm is the following. If some Δ time has not passed since the last switching time, or the entropy of the current deputy has not fallen below some threshold ϵ, the current deputy j and current Δt^* is returned. If not, then the entropy of all the deputies is computed. The deputy to be viewed is chosen to be the one with maximum entropy, and Δt^* is updated to the current time. The supervisor cost function in this work is considered to be the Shannon entropy of the covariance matrix, which is assumed to be Gaussian; other metrics such as measures of uncertainty relating to the deputies' belief states, such as the det or trace() of the covariance matrix, or a measure of information gain post-fusion, such as the posterior Fisher information [150] may also be utilized. The cost value for each deputy is calculated, and then the deputy with the highest cost value is selected to be the target deputy j. However, to prevent the possibility of instantaneous switching, a form of hysteresis is applied. A target switch is only triggered when the cost of the current target deputy falls below some threshold, the potential other target deputy crosses over the same threshold, and two target switches cannot occur within a time threshold, $\delta \geq \gamma$, of each other, where δ is the time differential between switches, and γ is the desired time hysteresis threshold.

Algorithm 1: Supervisor algorithm

Input: $\{(\hat{\mathbf{x}}_i, \mathbf{P}_i)\}_{i=1}^{N}$, j, Δt^*
Output: j_{new}, Δt^*_{new}

$\leftarrow$ CurrentTime
if $(t - \Delta t^* \leq \Delta) \ \vee \ (\frac{k}{2}(1 + \log(2\pi)) + \log(|\mathbf{P}_i|) > \epsilon)$ **then**
    $j_{new} \leftarrow j$
    $\Delta t^*_{new} \leftarrow \Delta t^*$
else
    for $i = 1, \ldots, N$ **do**
        $\mathcal{J}_i \leftarrow \frac{k}{2}(1 + \log(2\pi)) + \log(|\mathbf{P}_i|)$
    $j_{new} \leftarrow \arg\max_{i \in \{1,\ldots,N\}} \mathcal{J}_i$
    $\Delta t^*_{new} \leftarrow t$
return j_{new}, Δt^*_{new}

3.3.2 Controller

We implement a state tracking version of the model predictive controller outlined in (2.28). Here we let $h(\mathbf{x}(k), \mathbf{u}(k)) := \mathbf{x}(k)$. Let there be two positive definite square matrices $\mathbf{W}_1 \in \mathbb{R}^{2\times 2}$ and $\mathbf{W}_2 \in \mathbb{R}^{3\times 3}$ and an Euler angle vector $\mathbf{z} := [\phi, \ \psi]^\top$. Note that the angle pair ϕ and ψ were chosen due to the mapping from the azimuth-elevation space. We denote a reference angle trajectory (i.e. angle sequence associated with a current designated target) is denoted as $\bar{\mathbf{z}}^r_k := \{\mathbf{z}^r(0), \mathbf{z}^r(1), \ldots, \mathbf{z}^r(T-1)\}$. The MPC problem is defined with a quadratic stage cost as

$$\begin{aligned} \underset{\mathbf{u}(k)}{\text{minimize}} \ & J(\mathbf{z}(k), \mathbf{u}(k)) := \sum_{i=k}^{k+T-1} (\mathbf{z}(i) - \mathbf{z}^r(i)) \\ \text{subject to } & \mathbf{z}(0) = \mathbf{z}(kk) \\ & \mathbf{z}(i+1) = f_z(\mathbf{z}(i), \mathbf{u}(i)), \quad i = k, \ldots, k+T-1 \\ & (\mathbf{z}(i), \mathbf{u}(i)) \in \mathscr{X} \times \mathscr{U}, \quad i = k, \ldots, k+T-1. \end{aligned} \tag{3.5}$$

Here, f_z is equivalent to the dynamics expressed in (2.20) and (2.24).

3.4 Results

This section presents a case when the agent is on the origin and tracks 5 agents and another case when the agent tracks 10 agents. Furthermore, this section also presents a case when multiple agents are on elliptical NMTs and want to track themselves without communication.

3.4.1 Agent at Origin of LVLH Frame

To demonstrate the efficacy of the proposed methodology, results of two example scenarios are presented - a 3 deputy case and a 10 deputy case. For this sample, the target deputy is selected via the Supervisor Algorithm detailed in Algorithm 1. The MPC algorithm then uses the azimuth-elevation tracks of deputy j' as the reference trajectory. The CasADi optimization toolbox is used to solve the nonlinear optimal control problem [151] using a direct single shooting method. The MPC cost function weighting matrices were set as $\mathrm{W}_1 = \mathbf{I}_2$ and $\mathrm{W}_2 = \mathbf{I}_3$. The torque constraint was set to be $u_{max} = 2\pi$ Nm, the angular velocity constraint was set as $\omega_{max} = \pi$ rad/s The example simulation consists of three deputy agents following elliptical NMTs about the chief. The chief's initial attitude and angular velocity were randomly generated with the initial angular velocity being restricted to a bounded box of $\pi/2$ m/s about the origin. Because the focus of this work is on target selection and control, the initial estimates were set to the truth values and the associated covariances were set to be $\mathbf{P}_0 = \beta\mathbf{I}_n$ where β is a randomly generated parameter $10 \leq \beta \leq 100$. If discrepancies in the initial truth state and state estimate were considered, the complexity of the problem increases dramatically. Elements of target search are introduced if the discrepancy between truth position and estimated position grows sufficiently large. This alone introduces the question: what bound on initial estimate error is feasible for estimate convergence? This question along with related questions are too broad in scope to be considered here and have consequently been left for future work. However, the relevancy and significance of tackling such a problem cannot be overstated.

Figure 3.2 depicts azimuth-elevation tracks of the deputy spacecraft and the ability of

the chief to track each deputy. The orientation of the chief switches rather frequently at the beginning of the simulation while the covariance matrices begin to settle below the specified threshold of $\epsilon = 0$ on the covariance entropy seen in Figure 3.3. To prevent instantaneous switching, a switch is only triggered when the entropy of the associated covariance matrix falls below this threshold, in and the entropy of the covariance of any other deputy exceeds the same threshold. A time hysteresis is also applied indicating the target agent cannot switch within 100 seconds of one another such that $\Delta = 100$ from Algorithm 1. It is shown that as the simulation progresses, the chief switches less frequently because these threshold are crossed less often as the chief becomes more certain of the trajectory of each deputy. Figure 3.4 is the resulting control torque to track the reference trajectory. Uncharacteristically large spikes are present at about 1,750 seconds when the azimuth of the reference trajectory crosses from π to $-\pi$, causing the chief's attitude to wrap around. This particular cases is present due to the parameterization of the attitude, and for this reason future work is needed to adapt the presented methodology to attitude parameterizations that do not present this discontinuity such as rotation matrices and unit quaternions.

The performance of the algorithm was also tested on a 10 deputy case, shown in Figure 3.6, Figure 3.7, Figure 3.8. In Figure 3.6 the MPC is able to correctly track the reference trajectory specified by the supervisor algorithm even with increased deputy count. However, because of the increased number of deputies, the chief is not able to linger on each deputy for as long and must quickly move on to track the next deputy. Despite the increased workload, the chief is still able to achieve the desired level of confidence in each state estimate indicated by Figure 3.7. But, each entropy stays closer to the threshold when compared to the 3 deputy case. Furthermore, the chief must apply a greater amount of torque throughout the duration of the the mission seen in Figure 3.8. This is a direct result of the supervisor algorithm greedily selecting the next reference deputy based on the information entropy. It is possible to include control effort in the supervisor algorithm to

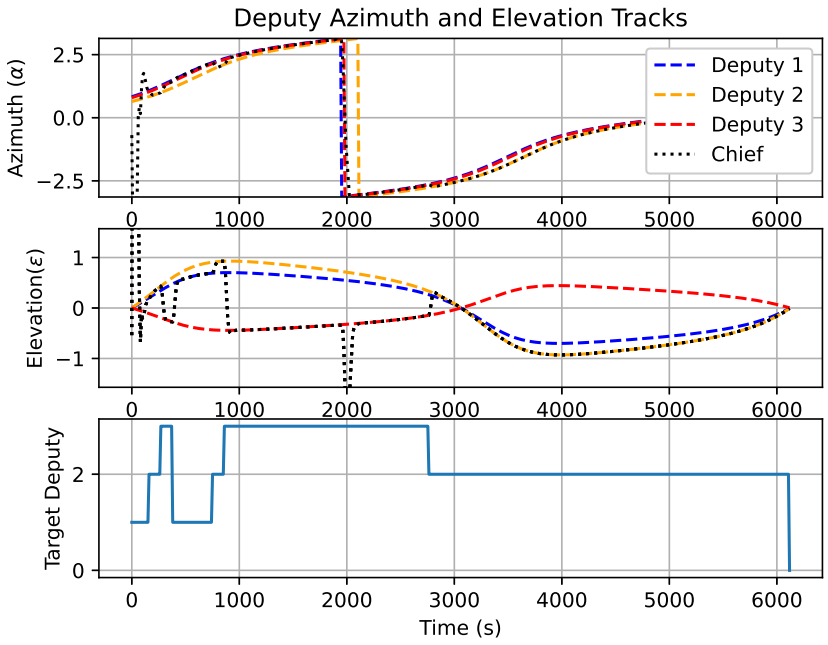


Figure 3.2 The MPC tracks the reference azimuth and elevation from the deputy selected by the supervisory algorithm.

help mitigate the effort applied by the chief.

3.4.2 Agent on an Elliptical NMT

The first case is a 5 agent case with communication disabled. This case demonstrates the ability of each agent to track all of the other agents on their own without the aid of communication; this would also resemble a single agent tracking 4 noncooperative bodies. Figure 3.9a depicts azimuth-elevation tracks of one of the agents and demonstrates the ability of the agent to track each of the other agents. We notice that the covariance entropy begins to settle below the specified threshold of $\epsilon = 0$ on the covariance entropy but is unable to during the time allotted. To prevent instantaneous switching, a switch is only triggered when the entropy of the associated covariance matrix falls below this threshold, in and the entropy of the covariance of any other deputy exceeds the same threshold. To evaluate the performance of the algorithm, the time hysteresis is set such that $\mu = 100$s from Algorithm 1. Each agent, on its own, is unable to maintain the desired

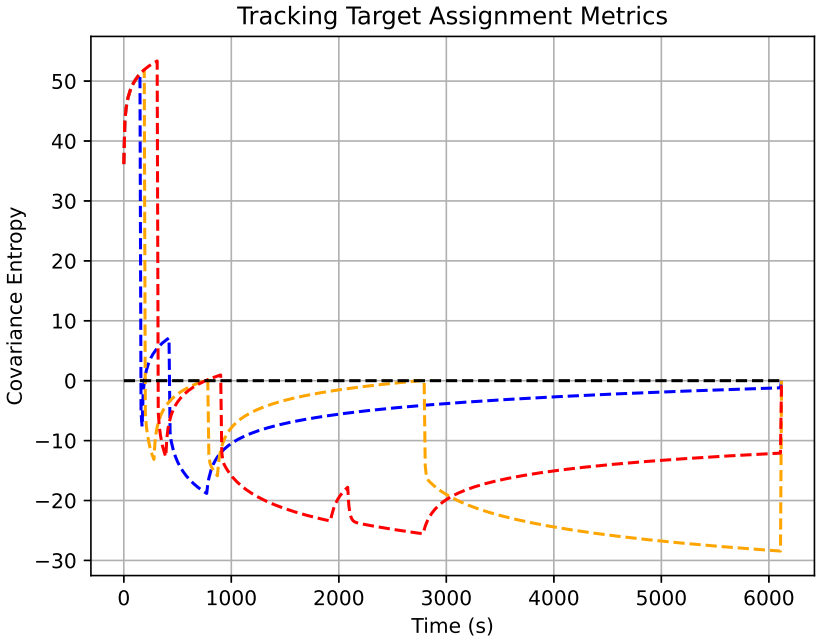


Figure 3.3 Covariance entropy as the chief observes each target deputy. It is shown that once each entropy value falls below the threshold, the control algorithm is able to maintain the entropy below the threshold.

covariance threshold due varying factors i.e. increased transverse time to view another agent, and the number of agents requiring tracking. In Figure 3.9b, the estimate errors for each agent collapse when a measurement is taken, but continues to grow until that point. The supervisor is able to select an appropriate target for each agent and the local controller is able to maintain the reference trajectory of the target.

Similarly, the estimate errors in Figure 3.10b continue to grow until each agent takes a measurement of another; at which point the error collapses. However, because of the increased number of objects, each agent is not able to linger on a target for as long and must quickly move on to track the next target. It is expected that the inclusion of state and output communication amongst agents will improve the convergence of the local estimators.

3.5 A Note Moving Forward

What happens when we allow each of the agents in Section 3.4.2 to communicate with one another? Each of the agents may communicate a local copy of their state estimates

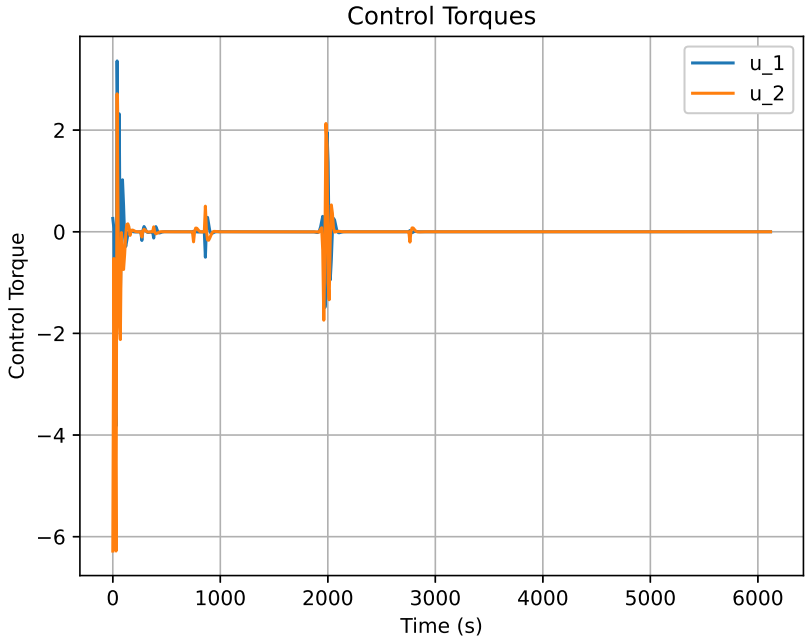


Figure 3.4 Applied torque for the 3 deputy case.

and/or their measurements to some other subset of the agents. Then, it is up to each agent to fuse any incoming state estimates with its local copy to update, and ideally, improve its own state estimate. With this in mind, how can we guarantee that information is routed appropriately such that each agent's own state estimate? Here, we must pause our current train of thought and discuss a tool that will allow us to analyze, and later design, a networked estimator to achieve the desired convergence property.

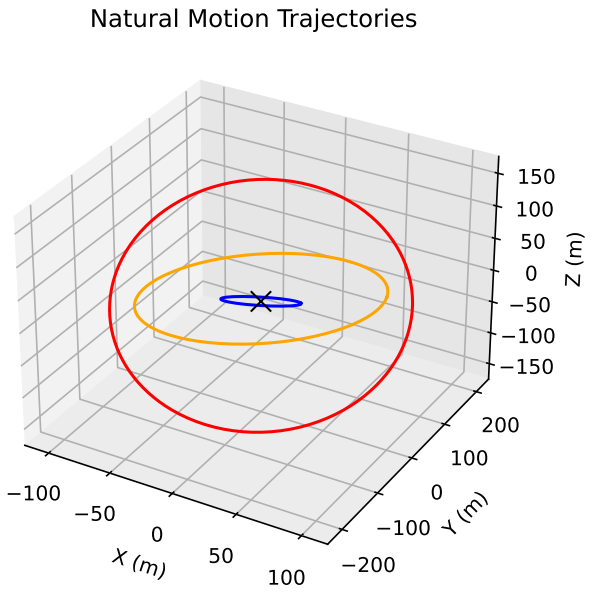


Figure 3.5 NMTs that generate the azimuth-elevation tracks.

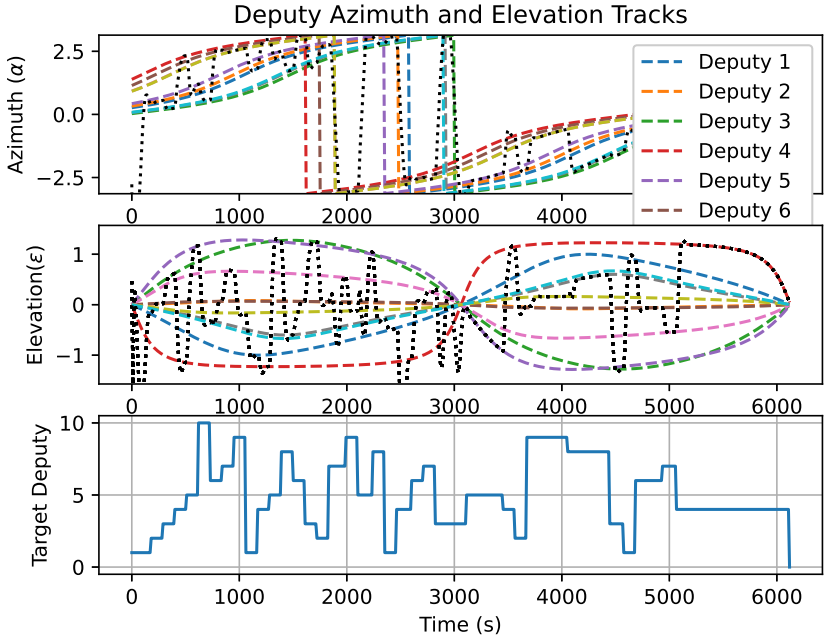


Figure 3.6 The algorithms are capable of tracking the reference azimuth and elevation track when more deputies require tracking.

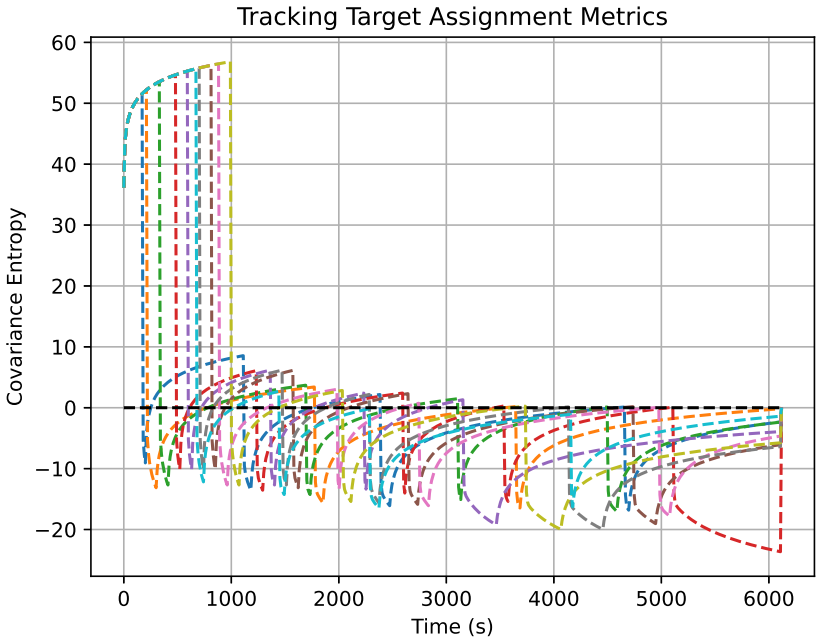


Figure 3.7 The algorithms are also able to maintain the covariance entropy below the desired threshold, even though it takes longer to achieve this level of estimator confidence.

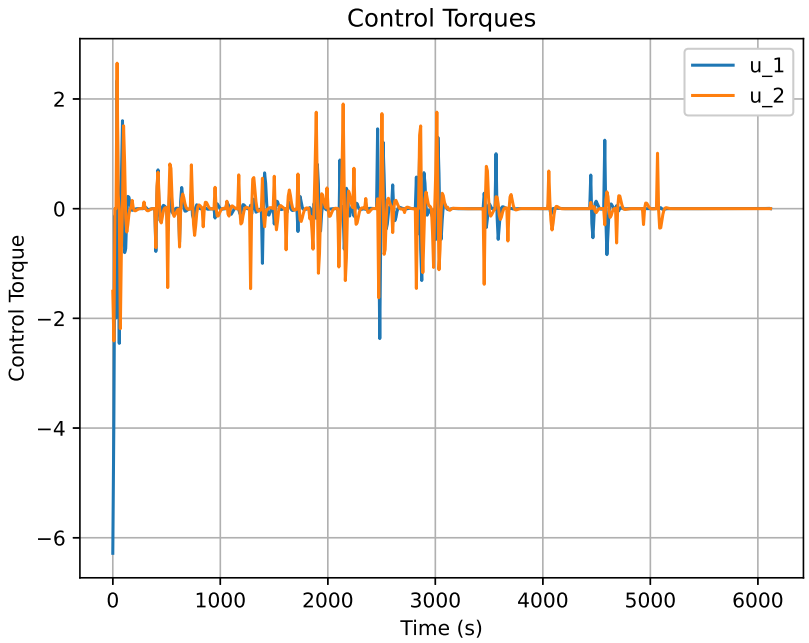
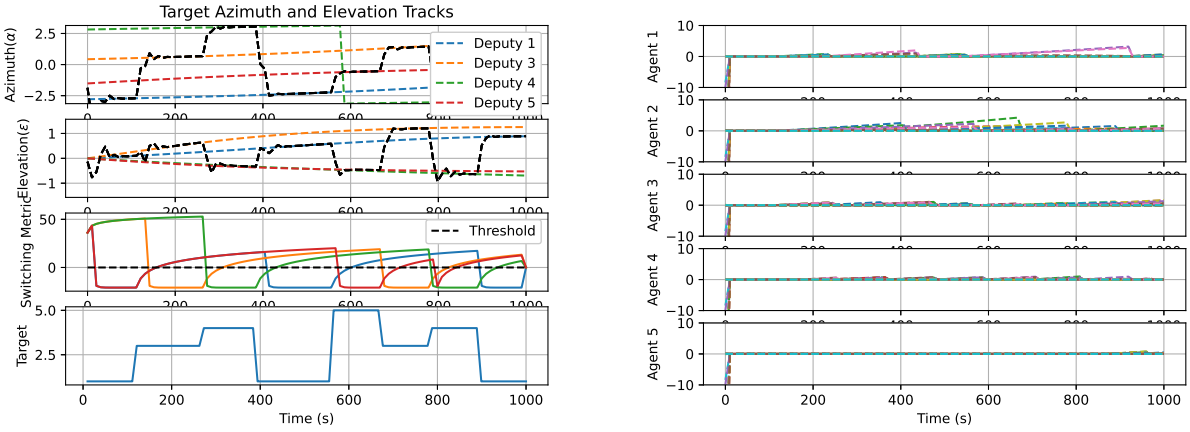
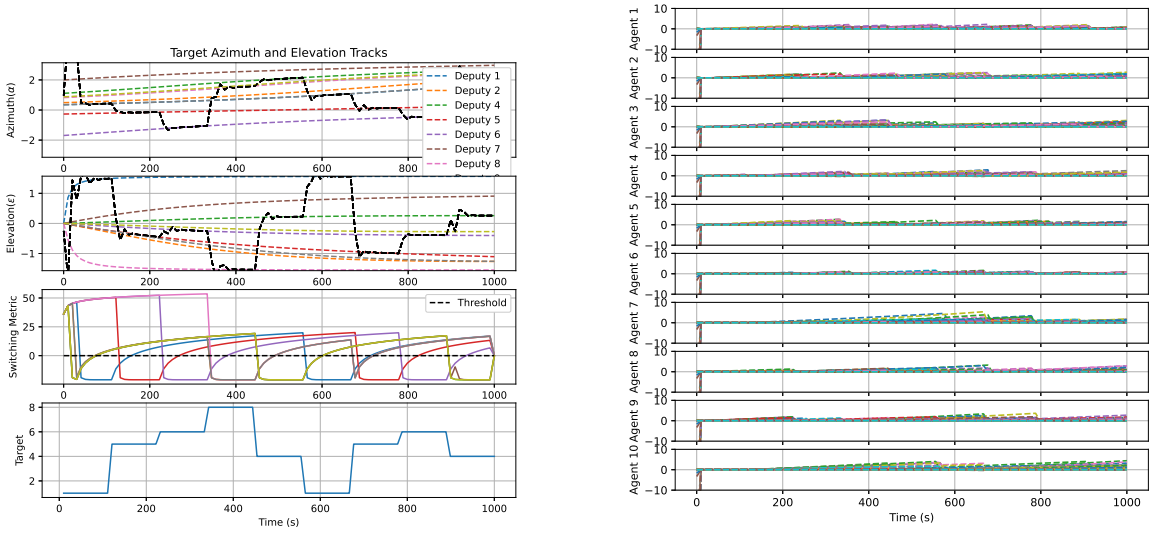


Figure 3.8 Applied torque for the 10 deputy case.



(a) Simulation Example *(b)* Estimate Error

Figure 3.9 5 Agents with Communication Disabled



(a) Simulation Example *(b)* Estimate Error

Figure 3.10 10 Agents with Communication Disabled

4 State Omniscience of Linear Time-Invariant Distributed Estimators

In this chapter, we consider N discrete-time autonomous dynamic systems of the form

$$\mathbf{x}_i(k+1) = f_i(\mathbf{x}_i(k), \mathbf{u}_i(k)) \tag{4.1}$$

where $\mathbf{x}_i \in \mathbb{R}^{n_i}$ is the local state vector, $\mathbf{x} = [\mathbf{x}_1^\top, \ldots, \mathbf{x}_N^\top]^\top \in \mathbb{R}^n$ where $n = \sum_i^N n_i$, $\mathbf{u}_i \in \mathbb{R}^{q_i}$ is the local control vector with $\mathbf{u} = [\mathbf{u}_1^\top, \ldots, \mathbf{u}_N^\top]^\top \in \mathbb{R}^q$ where $q = \sum_i^N q_i$, and $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ represents the dynamics. There exists N local estimators of the form

$$\hat{\mathbf{x}}_i(k+1) = \hat{f}_i(\hat{\mathbf{x}}(k), \mathbf{y}(k), \mathbf{u}(k)) \tag{4.2}$$

where $\hat{\mathbf{x}}_i \in \mathbb{R}^n$ is the local state estimate at each estimator with $\hat{\mathbf{x}} := [\hat{\mathbf{x}}_1^\top, \ldots, \hat{\mathbf{x}}_N^\top]^\top \in \mathbb{R}^{nN}$. It is assumed that each estimator is equipped with a measurement function

$$\mathbf{y}_i = h_i(\mathbf{x}) \tag{4.3}$$

where $\mathbf{y}_i \in \mathbb{R}^{p_i}$ is the local output vector at each estimator with $\mathbf{y} := [\mathbf{y}_1^\top, \ldots, \mathbf{y}_N^\top]^\top \in \mathbb{R}^p$ where $p = \sum_i^N p_i$. We allow for each of the local estimators to communicate either $\hat{\mathbf{x}}_i$, $\mathbf{y}_i$, or both to their neighbors according to a network graph $\mathcal{G}_c$. The collection of estimators and $\mathcal{G}_c$ comprise what we call a distributed estimator. Formally, a distributed estimator is uniquely defined by the tuple $\mathcal{DE} := (\mathcal{G}_c, \{\hat{\mathbf{x}}_i\}_{i=1}^N)$ and each estimator represents a vertex of $\mathcal{G}_c$. Furthermore, we define $\mathbf{e}_i := \mathbf{x} - \hat{\mathbf{x}}_i$ to be the estimate error at the i^{th} state estimator.

It is the goal of each local estimator to correctly estimate the true state of the dynamics such that $\hat{\mathbf{x}}_i \to \mathbf{x}$. Consequently, this indicates that $\mathbf{e}_i \to 0$ for each $i \in \mathcal{V}$.

Definition 7. *(**State Omniscient** (SO)) A $\mathcal{DE}$ is SO if $\mathbf{e}_i$ for each $i \in \mathcal{V}$ is asymptotically stable i.e. for $\delta > 0$*

$$||\mathbf{e}_i(0)|| < \delta \Rightarrow \lim_{k\to\infty} \mathbf{e}_i(k) = 0.$$

for each $i \in \mathcal{V}$.

Definition 7 implies the existence of N Lyapunov functions $V_i(\mathbf{e}_i, k) : \mathbb{R}^n \to \mathbb{R}$ such that $V_i(0, k) = 0$, $V_i(\mathbf{e}_i, k) > 0$, and $\Delta V_i(\mathbf{e}_i, k+1) = V_i(\mathbf{e}_i, k+1) - V_i(\mathbf{e}_i, k) < 0 \; \forall \; k \geq 0$.

4.1 Problem Statement

Consider the multi-agent LTI system described in Section 2.4. A standard Luenberger observer at each agent i can be defined as

$$\hat{\mathbf{x}}_i(k+1) = \mathbf{\Phi}\hat{\mathbf{x}}_i(k) + \mathbf{K}(\mathbf{y}_i(k) - \mathbf{C}_i\hat{\mathbf{x}}_i(k)) \tag{4.4}$$

where $\hat{\mathbf{x}}_i(k) \in \mathbb{R}^{nN}$ is the full system state estimate of agent i at timestep k, and $\mathbf{K}$ is a tunable gain matrix. The Luenberger observer in (4.4) assumes a single state observer or a centralized observer, in the case of multi-agent systems, and does not explicitly allow for a distributed approach when each agent within a network may have its own estimator.

We adopt a collection of local estimators within a networked system that takes advantage of all available information at agent i, including sensed and communicated. This collection of estimators forms a networked state estimator given by

$$\hat{\mathbf{x}}_i(k+1) = \underbrace{\sum_{j \in \mathcal{N}_i} \mathbf{W}_{ij}\mathbf{\Phi}\hat{\mathbf{x}}_j(k)}_{\text{Predictor Fusion Term}} + \mathbf{K}_i \underbrace{\sum_{j \in \mathcal{N}_i} \mathbf{U}_{ij}\mathbf{C}_j^\top(\mathbf{y}_j(k) - \mathbf{C}_j\hat{\mathbf{x}}_i(k))}_{\text{Output Fusion Term}} \tag{4.5}$$

which incorporates both predictor fusion weights through a positive-definite matrix $\mathbf{W}_{ij} \in \mathbb{R}^{nN \times nN}$ and output fusion weights through a positive-definite matrix $\mathbf{U}_{ij} \in \mathbb{R}^{nN \times nN}$ for each $i, j \in \mathcal{V}$ where $\mathbf{W}_{ij} = \mathbf{0}$ if and only if $a_{ij} = 0$ and $\mathbf{U}_{ij}$ is populated in a similar manner; for more information see [76, 77, 80, 152]. The innovation gain matrix $\mathbf{K}_i \in \mathbb{R}^{nN \times nN}$ tunes the state estimate for agent i based on the available information. The fusion elements are almost exclusively governed by the communication links described by the adjacency matrix $\mathbf{A}_c$. The state estimator may also be placed into an augmented form using the full state estimator vector $\hat{\mathbf{x}}(k) = [(\hat{\mathbf{x}}_1(k))^\top, \ldots, (\hat{\mathbf{x}}_N(k))^\top]^\top \in \mathbb{R}^{nN^2}$ with each

$\hat{\mathbf{x}}_i(k) \in \mathbb{R}^{nN}$

$$\hat{\mathbf{x}}(k+1) = (\mathbf{W} \otimes \mathbf{\Phi})\hat{\mathbf{x}}(k) + \mathbf{K}\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})(\hat{\mathbf{x}}(k) - \mathbf{x}(k)) \tag{4.6}$$

where the augmented innovation gain and augmented output matrix are $\mathbf{K} = \text{diag}([\mathbf{K}_1, \ldots, \mathbf{K}_N])$ and $\mathbf{C} = \text{diag}([\mathbf{C}_1^\top, \ldots, \mathbf{C}_N^\top])^\top$, respectively. The values of each $\mathbf{W}_{ij}$ and $\mathbf{U}_{ij}$ may be computed via an information fusion scheme [106, 109, 114].

Considering the case when $\mathbf{U}_{ij}$ is a scalar, as is the case in [76, 80], yields

$$\Big(\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})\Big)_{ij} = \sum_{l=1}^{N} u_{il}u_{lj}\mathbf{C}_l^\top\mathbf{C}_l. \tag{4.7}$$

The block diagonal elements of (4.7) are then

$$\Big(\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})\Big)_{ii} = \sum_{l=1}^{N} u_{il}^2\mathbf{C}_l^\top\mathbf{C}_l. \tag{4.8}$$

Also following [76] and [80], it is evident that the estimators at each agent in (4.5) may be written in the form

$$\hat{\mathbf{x}}(k) = \big(\mathbf{W} \otimes \mathbf{\Phi}\big)\hat{\mathbf{x}}_{k-1} + \mathbf{K}\mathbf{C}_\mathbf{A}(\bar{\mathbf{x}}(k) - \hat{\mathbf{x}}(k)) \tag{4.9}$$

with $\mathbf{C}_\mathbf{A} = \text{diag}([\mathbf{C}_\mathbf{A}^1, \ldots, \mathbf{C}_\mathbf{A}^N])$ where $\mathbf{C}_\mathbf{A}^i = \sum_{j\in\mathcal{N}_i} u_{ij}^2\mathbf{C}_j^\top\mathbf{C}_j$ for each $i \in \mathcal{V}$ This section demonstrates that the $\mathbf{C}_\mathbf{A}$ of (4.9) is equivalent to $\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})$ of (4.6).

The notation presented herein generalizes the notation used by Khan and Doostmohammadian in [80] allowing for more general approaches to information fusion schemes to be applied such as $\mathbf{W}$ and $\mathbf{U}$ used herein. Explicitly describing the graph interconnections for both the predictor and output fusion steps allows us to directly analyze the role communication links play in networked system observability.

To analyze the system, we consider the error dynamics by substituting (2.7) and (4.6) into $\mathbf{e}(k) := \bar{\mathbf{x}}(k) - \hat{\mathbf{x}}(k)$ with $\bar{\mathbf{x}}(k) = \mathbf{1}_N \otimes \mathbf{x}(k)$ where $\mathbf{1}_N$ is the N-dimensional vector of

all ones. An augmented version of the truth dynamics may then be written as

$$\begin{aligned} \bar{\mathbf{x}}(k+1) &= \big(\mathbf{W} \otimes \mathbf{\Phi}\big)\bar{\mathbf{x}}(k) \\ \mathbf{y}(k) &= (\mathbf{E}_s^\top \otimes \mathbf{H})\bar{\mathbf{x}}(k) \end{aligned}. \tag{4.10}$$

Assuming an admissible[1] selection of $\mathbf{W}$, expressing the truth dynamics through (4.10) is equivalent to $\bar{\mathbf{x}}(k+1) = (\mathbf{I} \otimes \mathbf{\Phi})\bar{\mathbf{x}}(k)$ which eases analyses by simplifying the error dynamics

$$\mathbf{e}(k+1) = \Big(\mathbf{W} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})\Big)\mathbf{e}(k). \tag{4.11}$$

Since (4.11) is a discrete-time LTI system, it is well-understood that the origin is asymptotically stabilized if there exist weights $\mathbf{W},\ \mathbf{U}$, and gain $\mathbf{K}$ such that the spectral radius of the matrix on the right-hand side of (4.11) is less than 1. More specifically,

$$\rho\Big(\mathbf{W} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{U}\text{diagm}(\mathbf{C}^\top\mathbf{C})\Big) < 1$$

where $\rho(\cdot)$ indicates the spectral radius of the matrix, which is true if there exists weights, $\mathbf{W},\ \mathbf{U}$, and an innovation gain $\mathbf{K}$ that stabilizes (4.11); see [76, 80]. Consequently, this work seeks a formal characterization of the existence of the matrices $\mathbf{W},\ \mathbf{U},\ \mathbf{K}$ that stabilizes (4.11). We formalize the problem as follows:

Problem 1. *Given a collection of interconnected local state estimators of the form in* (4.5)*, individual output matrices* $\mathbf{C}_i$*, and communication graph* $\mathcal{G}_c$ *design fusion matrices* $\mathbf{W}_{ij}$*,* $\mathbf{U}_{ij}$ *and gain matrices* $\mathbf{L}_i$ *such that the* $\mathcal{DE}$ *meets Definition 7.*

Consider the following example: given two estimators, which we will refer to as

[1]We consider any matrix $\mathbf{W}$ to be admissible if it satisfies $\sum_{j=1}^{N} \mathbf{W}_{ij} = \mathbf{I}$ for each $i \in \{1, \ldots, N\}$. This follows previous stipulations on weighting matrices from the field of information fusion.

Estimator 1 and Estimator 2, each equipped with its own state estimator of the form

$$\begin{aligned}\hat{\mathbf{x}}_1(k+1) =& \mathbf{W}_{11}\mathbf{\Phi}_d\hat{\mathbf{x}}_1(k) + \mathbf{W}_{12}\mathbf{\Phi}_d\hat{\mathbf{x}}_2(k) + \mathbf{L}_1\mathbf{C}_1(\hat{\mathbf{x}}_1(k) - \mathbf{x}(k)) \\ \hat{\mathbf{x}}_2(k+1) =& \mathbf{W}_{21}\mathbf{\Phi}_d\hat{\mathbf{x}}_1(k) + \mathbf{W}_{22}\mathbf{\Phi}_d\hat{\mathbf{x}}_2(k) + \mathbf{L}_2\mathbf{C}_2(\hat{\mathbf{x}}_2(k) - \mathbf{x}(k))\end{aligned}$$

where $\hat{\mathbf{x}}_1(k),\ \hat{\mathbf{x}}_2(k) \in \mathbb{R}^{nN}$, $\mathbf{W}_{ij}$ admissibly fuses communicated information from j to i, and $\mathbf{L}_i \in \mathbb{R}^{nN\times p_i}$ is a tunable gain matrix. The estimator pair is tasked with independently tracking the LTI system described by (2.7) such that the distributed estimator is state omniscient. Estimator 2 is incapable of taking a measurement such that $\mathbf{C}_2 = \mathbf{0}$, and Estimator 1 may communicate its state estimate to Estimator 2. A visualization of this scenario can be seen in Figure 4.1. The blue edge represents a sensing link, and the red edge represents a communication link.

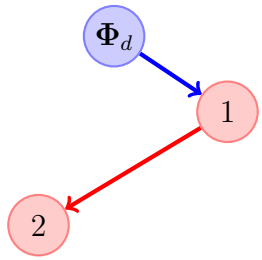


Figure 4.1 Motivating example depicting measurement relationships in blue and communication relationships in red.

Without loss of generality, we assume that $\mathbf{\Phi}_d$ is either marginally stable or unstable, thus there needs to exist a gain matrix $\mathbf{L}_i$ that stabilizes the system. It is a well-known result that there exists a gain $\mathbf{L}_i$ to stabilize the estimator at Estimator i if and only if the pair $(\mathbf{\Phi}_d, \mathbf{C}_i)$ is observable. We assume, without loss of generality, that the pair $(\mathbf{\Phi}_d, \mathbf{C}_1)$ is observable; however, because Estimator 2 is not taking any measurements then $\mathbf{C}_2 = \mathbf{0}$ the pair $(\mathbf{\Phi}_d, \mathbf{C}_2)$ is unobservable. Therefore the estimator at Estimator 2 may be written as

$$\hat{\mathbf{x}}_2(k+1) = \mathbf{W}_{21}\mathbf{\Phi}_d\hat{\mathbf{x}}_1(k) + \mathbf{W}_{22}\mathbf{\Phi}_d\hat{\mathbf{x}}_2. \tag{4.12}$$

For example, consider the case when $\mathbf{W}_{21} = \mathbf{I}$ and $\mathbf{W}_{22} = \mathbf{0}$ such that (4.12) becomes

$$\hat{\mathbf{x}}_2(k+1) = \mathbf{\Phi}_d \hat{\mathbf{x}}_1(k). \tag{4.13}$$

Since the pair $(\mathbf{\Phi}_d, \mathbf{C}_1)$ is observable we have that $\hat{\mathbf{x}}_1 \to \mathbf{x}$ and from (4.13) it follows that $\hat{\mathbf{x}}_2 \to \mathbf{x}$. However, it acts as a second order system which creates a delay in convergence. As such, while the pair $(\mathbf{\Phi}_d, \mathbf{C}_2)$ is not observable, $\hat{\mathbf{x}}_2 \to \mathbf{x}$ due to the communicated information provided by $\hat{\mathbf{x}}_1$. Since both $\hat{\mathbf{x}}_2 \to \mathbf{x}$ and $\hat{\mathbf{x}}_1 \to \mathbf{x}$, we call the networked system state omniscient. More specifically, for the case of heterogeneous sensor communication networks the goal of state omniscience is to reconstruct the state at each vertex in the network. This paper studies this phenomena and seeks to find a formal characterization of state omniscient distributed estimators. More specifically, we seek a set of necessary and sufficient conditions such that there exist a fusion/consensus matrix $\mathbf{W}$ and gain matrix $\mathbf{L}$ that yield a state omniscient distributed estimator.

Prior to discussing the existence of such matrices in the presence of multiple dynamic systems and a communication graph $\mathcal{G}_c$, let us first consider a single dynamic systems. Insights from single dynamic system case will give us a better understanding in how to explore the mult-agent case.

4.2 Ubiquitous Single-Input Controllability and Ubiquitous Single-Output Observability

This section introduces a class of linear systems that are input vector agnostic. In particular, this section introduces this class of linear systems from a transfer function and traditional controllability perspective. We use the term *ubiquitous single-input controllability* to identify this specific class of linear systems.

4.2.1 Ubiquitous Single-Input Controllability

Definition 8. ***(Ubiquitous Single-Input Controllability)*** *A linear dynamic system of the form* $\dot{\mathbf{x}} = \mathbf{\Phi}\mathbf{x} + \mathbf{B}\mathbf{u}$ *is considered ubiquitously single-input controllable (USIC) if any*

non-zero vector $\mathbf{B}$ *renders the system controllable.*

Theorem 1. *The LTI system is USIC if and only if the controllability matrix* $\mathbf{C}_c^i$ *of the pair* $(\mathbf{\Phi},\ \mathbf{B}_i)$ *is full rank for each* $i = 1, \ldots, n$*, where* $\mathbf{B}_i$ *is the* i^{th} *column of the identity matrix.*

$$\mathbf{C}_c^i = \begin{bmatrix} \mathbf{B}_i & \mathbf{\Phi}\mathbf{B}_i & \mathbf{\Phi}^2\mathbf{B}_i & \ldots & \mathbf{\Phi}^{n-1}\mathbf{B}_i \end{bmatrix} \tag{4.14}$$

Proof. Proof follows directly from the definition. □

Theorem 1 is functionally only dependent on $\mathbf{\Phi}$ which begs the question: what set of $\mathbf{\Phi}$ matrices make (4.14) full rank $\forall\ i$?

In general, $\mathbf{C}_c^i$ need not be square. However, because we are only considering the single-input case, $\mathbf{C}_c^i \in \mathbb{R}^{n\times n}$ is square which allows us to look at the columns of $\mathbf{C}_c^i$, which are more physically meaningful, for linear independence. An additional important consequence of the construction of $\mathbf{C}_c^i$ is the implication of linear independence of the i^{th} column between each matrix $\mathbf{\Phi}^l$ for each $l \in \{0, \ldots, n-1\}$.

Proposition 1. *If the* i^{th} *columns of* $\mathbf{\Phi}^l$ *for each* $l = 1, \ldots, n-1$ *are linearly independent, then* $\mathbf{C}_c^i$ *is full rank for each* i*.*

Proof. By definition, any square matrix is full rank if every column in the matrix is linearly independent. Considering $\mathbf{C}_c^i$ from (4.14), the column vector $\mathbf{B}_i$ extracts the i^{th} column of each matrix. Alternatively, (4.14) can be rewritten as:

$$\mathbf{C}_c^i = \begin{bmatrix} \mathbf{\Phi}_i^0 & \mathbf{\Phi}_i^1 & \mathbf{\Phi}_i^2 & \ldots & \mathbf{\Phi}_i^{n-1} \end{bmatrix} \tag{4.15}$$

where $(\cdot)_i$ indicates the i^{th} column of the matrix. From (4.15), it is evident that $\mathbf{C}_c^i$ is full rank if the i^{th} column of $\mathbf{\Phi}^l$ for each $l = 1, \ldots, n-1$ are linearly independent. □

Proposition 2. *If a LTI system described by* $\mathbf{\Phi}$ *is USIC, then the set* $\{\mathbf{I},\ \mathbf{\Phi},\ \mathbf{\Phi}^2\ \ldots,\ \mathbf{\Phi}^{n-1}\}$ *is linearly independent.*

Proof. By definition of linear independence of vectors we can substitute every column of $\mathbf{C}_c^i$ shown in (4.15) which leads to

$$\alpha_{i,0}\mathbf{\Phi}_i^0 + \alpha_{i,1}\mathbf{\Phi}_i^1 + \cdots + \alpha_{i,n-1}\mathbf{\Phi}_i^{n-1} = 0 \tag{4.16}$$

for each $i \in \{1, \ldots, n\}$. Since $\mathbf{\Phi}$ is USIC, $\mathbf{C}_c^i$ is full rank for each $i \in \{1, \ldots, n\}$, therefore (4.16) is true if and only if $\alpha_{i,0} = \cdots = \alpha_{i,n-1} = 0$. Now, let us consider the vector representation of each matrix where the operator $\text{vec}(\cdot) : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$ maps the matrix into a row vector, more specifically

$$\text{vec}(A) = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \ldots & \mathbf{\Phi}_n \end{bmatrix} \tag{4.17}$$

Therefore, (4.16) can be written in a matrix form.

$$\begin{bmatrix} \text{vec}(\mathbf{\Phi}^0) \\ \text{vec}(\mathbf{\Phi}^1) \\ \ldots \\ \text{vec}(\mathbf{\Phi}^{n-1}) \end{bmatrix} = \begin{bmatrix} (\mathbf{C}_c^1)^\top & (\mathbf{C}_c^2)^\top & \ldots & (\mathbf{C}_c^{n-1})^\top \end{bmatrix} \tag{4.18}$$

where $\mathbf{\Phi}_i$ is the i^{th} row of $\mathbf{\Phi}$. Note that each block matrix $(\mathbf{C}_c^i)\top$ of (4.18) is full rank down the column, the full vector form in (4.18) must also be linearly independent. Thus proving the result. □

Corollary 2 does not imply that each $\mathbf{\Phi}^l$ must be full rank for each l, but it does necessitate that every column in each $\mathbf{\Phi}^l$ must have at least one non-zero element.

Corollary 1.1. *If the LTI system with matrix $\mathbf{\Phi}$ is USIC, then the matrix $\mathbf{\Phi}$ is full rank.*

Proof. Proving the corollary is trivial and follows as a direct result of Theorem 2. Particularly, from (4.18) it is obvious that in order for the full set $\{\mathbf{I},\ \mathbf{\Phi},\ \mathbf{\Phi}^2\ \ldots,\ \mathbf{\Phi}^{n-1}\}$ to be linearly independent, each row block must also be linearly independent. □

Additionally, this also implies Corollary 1.2.

Corollary 1.2. *If a LTI system described by* $\mathbf{\Phi}$ *is USIC, then* $\mathbf{\Phi}$ *has minimum polynomial,* $\mu_{\mathbf{\Phi}}(s)$*, of degree* d*, where* $d = n$*.*

Proof. Prior to completing the proof, we first provide a definition for the minimum polynomial of a matrix. First, consider a graph $\mathcal{G}$ that is USIC and its associated adjacency matrix $\mathbf{\Phi}$ has a minimum polynomial of degree $d \leq n-1$. The minimum polynomial

$$\mathbf{\Phi}^d + \alpha_{d-1}\mathbf{\Phi}^{d-1} + \cdots + \alpha_1\mathbf{A} + \alpha_0\mathbf{I} = 0$$

with $d \leq n-1$ directly implies that there is linear dependence between the matrix powers of $\mathbf{\Phi}$ between 0 and $n-1$. This directly contradicts the implicit full rank condition of Corollary 1.1, proving that $d > n-1$.

The $\deg(\mu_{\mathbf{\Phi}}(s))$ is also upper-bounded, $d \leq n$, by the Cayley-Hamilton Theorem that states any matrix is a solution to its own characteristic polynomial, and proves the theorem. □

Furthermore, $\deg(\mu_{\mathbf{\Phi}}(s)) = n$ necessitates n distinct eigenvalues implying also that the the minimum polynomial is equal to the characteristic polynomial, $\mu_{\mathbf{\Phi}}(s) = p_{\mathbf{\Phi}}(s)$.

While Proposition 1 through Corollary 1.2 provide necessary conditions for $\mathbf{\Phi}$ that meet the USIC criteria, they all together do not sufficiently define the set. Rather, let us use the PBH Lemma [153, 154] to derive a necessary and sufficient condition.

Theorem 2. *An LTI system with the system matrix* $\mathbf{\Phi}$ *is USIC if and only if* $\mathbf{Z}_{ij}^\top \neq 0$ *for each* $i, j \in \{1, \ldots, n\}$*, where the columns of* $\mathbf{Z}$ *are the left-eigenvectors of* $\mathbf{\Phi}$*.*

Proof. Recalling that if a left-eigenvector of $\mathbf{\Phi}$ is orthogonal to a column of the input matrix $\mathbf{B}$, then that mode is uncontrollable from that column of $\mathbf{B}$, more specifically

$$\mathbf{z}^\top\mathbf{\Phi} = \lambda\mathbf{z}^\top,\ \mathbf{z}^\top\mathbf{B} = 0. \tag{4.19}$$

In other words, (4.19) says that the left eigenvector, $\mathbf{Z}$, must not be orthogonal to the columns of $\mathbf{B}$ for controllability. Recalling the definition of a USIC systems, (4.19) may be expanded to represent the effects of every combination of control input and state through

$$\mathbf{z}_j^\top \mathbf{\Phi} = \lambda_j \mathbf{z}_j^\top, \ \mathbf{z}_j^\top \mathbf{B}_i = 0. \tag{4.20}$$

If (4.20) is satisfied for each $i, j \in \{1, \ldots, n\}$, then the mode associated with left-eigenvector $\mathbf{Z}_j$ is uncontrollable from an input at node i and can be extended into an analogous matrix form as

$$\mathbf{z}^\top \mathbf{\Phi} = \lambda \mathbf{z}^\top, \ \mathbf{z}^\top \mathbf{B} = 0. \tag{4.21}$$

It has been assumed that $\mathbf{B}$ has been restricted to columns of the identity matrix, the USIC nature of a system is solely dependent on the left-eigenspace of the matrix $\mathbf{\Phi}$. Therefore, if $\mathbf{Z}_{ij}^\top = \mathbf{z}_j^\top \mathbf{B}_i = 0$, then the control input $\mathbf{B}_i$ does not have control authority of dynamic mode $\mathbf{z}_j$. Yet, the converse is also true, if $\mathbf{Z}_{ij}^\top = \mathbf{z}_j^\top \mathbf{B}_i \neq 0$ for each $i, j \in \{1, \ldots, n\}$ then the system is controllable for any combination of left eigenvectors and control inputs and, therefore, the system is USIC. □

Theorem 2 provides a necessary and sufficient condition for the set of $\mathbf{\Phi}$ matrices that meet the USIC criteria. Similarly, a system may also be ubiquitously single-output observable.

Definition 9. ***(Ubiquitous Single-Output Observability)*** *A dynamic system is considered ubiquitously single-output observable (USOO) if any single output renders the system observable. A linear dynamic system is USOO if its dual representation is USIC.*

Corollary 2.1. *A LTI system described by $\mathbf{\Phi}$ is USOO if and only if $\mathbf{V}_{ij} \neq 0$ for each $i, j = 1, \ldots, n$, where $\mathbf{V}$ is the matrix of right-eigenvectors of $\mathbf{\Phi}$.*

Proof. Follows the proof for Theorem 2 using the right-eigenvector and measurement matrix $\mathbf{C}$. □

It should be noted that it is possible for a system to be either USIC or USOO because each are dependent on either the left of the right-eigenspace. But, this result does not generalize to all systems that meet either the USIC or USOO requirement.

In terms of networked systems it is critical to evaluate the USIC condition against undirected graphs. It has previously been shown by Parlangeli [95, 155] that the controllability of graphs is highly dependent on the graph structure itself.

Theorem 3. *If $\mathbf{\Phi}$ is symmetric with $n > 2$, then the LTI system described by $\mathbf{\Phi}$ will not be USIC.*

Proof. Consider $\mathbf{\Phi}$ to be a symmetric matrix. The eigendecomposition of $\mathbf{\Phi} = \mathbf{V\Lambda V}^\top$ and according to the spectral theorem of symmetric matrices $\mathbf{\Phi}^l = \mathbf{V}\Lambda^l\mathbf{V}^\top$ and has real eigenvalues. Because $\mathbf{B}_i$ extracts the i^{th} column out of each matrix $\mathbf{\Phi}^l$ for each $l \in \{1, \ldots, n-1\}$, $\mathbf{C}_c^i$ is constructed as $\mathbf{C}_c^i = \begin{bmatrix} \mathbf{i} & \mathbf{V}\lambda_i\mathbf{v}_i & \mathbf{V}\lambda_i^2\mathbf{v}_i & \ldots & \mathbf{V}\lambda_i^{n-1}\mathbf{v}_i \end{bmatrix}$ where $\mathbf{i}$ is the i^{th} column of identity matrix. It is important to note that $(\cdot)_i$ indicates the i^{th} column of the transpose of the matrix of eigenvectors not the i^{th} eigenvector itself. This construction of $\mathbf{C}_c^i$ implies that every row is scaled up by $\mathbf{v}_i$ and therefore can not meet the condition for linear independence. This results in the matrix $\mathbf{C}_c^i$ failing to be full rank and therefore the graph $\mathcal{G}$ does not meet the controllability condition. □

A direct consequence of Theorem 3 is that $\mathbf{\Phi}$ must be asymmetric. This is a generalization of the results of Parlangeli [95, 155] that states no input to any single node of an undirected cycle graph will yield the network controllable. Furthermore, this also implies that undirected networks will never be USIC, because the $\mathbf{\Phi}$ for undirected networks will always be symmetric.

4.2.2 A Transfer Function Perspective

This next section discusses the USIC property from a transfer function or characteristic polynomial perspective. This notion was initially presented by Godsil [37] while studying

controllable subsets from a algebraic graph theoretic perspective. In the work, Godsil gives Proposition 3

Proposition 3. *A node i is controllable if and only if $p_{\mathbf{\Phi}}(s)$ and $p_{\mathbf{\Phi}\setminus i}(s)$ are coprime.*

This is to say that the characteristic polynomial of the adjacency matrix and the characteristic polynomial of the adjacency matrix with the input node removed must be coprime. Meaning the fraction

$$h(s) = \frac{p_{\mathbf{\Phi}\setminus i}(s)}{p_{\mathbf{\Phi}}(s)} \tag{4.22}$$

cannot be divided further. Proposition 3 implies Theorem 4.

Theorem 4. *An LTI system is USIC if and only if $p_{\mathbf{\Phi}}(s)$ and $p_{\mathbf{\Phi}\setminus i}(s)$ are coprime $\forall i$ and* $\deg(p_{\mathbf{\Phi}}(s)) = n$.

Proof. Recall that if any single zero does not cancel out any single pole of a system, then the system is controllable. With this in mind, we place the system in its transfer function form.

$$H(s) = \frac{Y(s)}{U(s)} = \frac{z_{\mathbf{\Phi}}(s)}{p_{\mathbf{\Phi}}(s)} \tag{4.23}$$

where $p_{\mathbf{\Phi}}(s)$ is the characteristic polynomial of the system, and $z_{\mathbf{\Phi}}(s)$ is another polynomial of degree $d \leq n-1$ which later we will show is equivalent to the characteristic polynomial of the matrix minor $M_{i,i}$ for single input linear systems. Because $z_{\mathbf{\Phi}}(s)$ has at most a degree of $n-1$ and it has previously been established that $p_{\mathbf{\Phi}}(s)$ has at least a degree of n, it is a necessary condition of uncontrollability that $p_{\mathbf{\Phi}}(s)$ is reducible. Contrarily, if $p_{\mathbf{\Phi}}(s)$ is irreducible then it follows that there does not exist any zero that will cancel out any pole of the system and therefore the theorem is proven to be sufficient. □

This fact is one alluded to by Godsil in [37] but is not explicitly stated.

4.2.3 Connections to Structural Controllability

Now that the idea of USIC systems has been introduced, this section ties the concept to a graph theory through structural controllability.

Theorem 5. *If the adjacency matrix* $\mathbf{\Phi}$ *is USIC, then it cannot be placed into* (2.16) *and is, therefore, a form of structural controllability.*

Proof. Let us assume that $\mathbf{\Phi}$ is USIC and can be placed into (2.16). The characteristic polynomial $p_{PAP^{-1}}(s) = |s\mathbf{I} - \mathbf{\Phi}_{11}||s\mathbf{I} - \mathbf{\Phi}_{22}|$ is a reducible polynomial that can be factored into the two characteristic polynomials of the block matrices $p_{\mathbf{\Phi}_{11}}(s)$ and $p_{\mathbf{\Phi}_{22}}(s)$, and therefore contradicts the property of USIC systems found in Corollary 1.2. Because any USIC system cannot be placed into (2.16), any USIC system is also of a structurally controllable form. □

Theorem 5 thus signifies that any USIC system is also a structurally controllable form, and is strongly connected. This fact directly connects the adjacency algebra, walk matrix, and diameter from algebraic graph theory [44, 115, 156] to the controllability of the graph.

It is also beneficial to discuss the graph structure of the USIC system. While a structurally controllable graph is spanned by a cactus, not all cacti are USIC. It is trivial to argue that a USIC cactus must be consist of a singular bud and stem because this work has been restricted to singular input systems where the input only has direct control over a singular state. However, the traditional concept of a bud being strictly a cycle graph is overly restrictive for USIC systems. Consider the graph formed by: It can be shown that

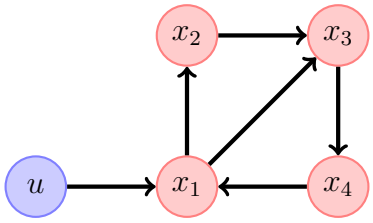


Figure 4.2 Sample USIC system that does not meet the traditional definition of a bud.

the system described in Figure 4.2 is USIC but the system does not necessarily meet the traditional definition of a bud. Therefore, a more general form of a bud is discussed herein.

The zero/nonzero patterns in eigenvectors and their association with the adjacency matrix structure of a digraph are presented in [157]. In [157], four pattern classes are

presented and are used to describe the eigenvectors for a given matrix structure: i.) Requires all strictly nonzero eigenvectors (RSN) ii.) Allows strictly nonzero eigenvectors (ASN) iii.) Allows partly zero eigenvectors (APZ) iv.) Requires partly zero eigenvectors (RPZ). It immediately follows from Theorem 2, that the matrix pattern of a USIC system must at least be ASN. Therefore, a system may be USIC if the matrix $\mathbf{\Phi}$ has a structure known to be ASN, but a system will be USIC if the matrix $\mathbf{\Phi}$ has a structure known to be RSN. Furthermore [157] makes the following proposition:

Proposition 4. *A matrix structure will require all strictly nonzero eigenvectors if and only if the digraph is a directed simple cycle on all the vertices.*

This result directly implies that any directed simple cycle structure is also USIC. This also confirms that the traditional notion of a bud (according to Lin [35]) is USIC.

However, there also exist the class of patterns that are ASN. While ASN patterns do not explicitly guarantee the USIC property of a system, there do exist a class of matrices $\mathbf{\Phi}$ that are of an ASN structure that are USIC. For example, the system from Figure 4.2 belongs to this class of ASN structures. Moreover, any USIC system is sufficient to form a bud of a structurally controllable graph. Structurally controllable graphs consists of a series of buds each with a singular input (termed "origin" in the original work) into the bud structures, but assumes that the bud itself is the cycle graph. However, if the bud is of a USIC structure then it directly follows that the bud will constitute a controllable portion of the graph. The idea of USIC systems may have additional interesting implications to the field of graph theory, input node selection for large complex networks, and the construction of controllable graphs using strictly USIC subgraphs may be interesting subjects for future work.

4.2.4 Practical Examples

Fault-tolerance Example

Consider the LTI system

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ 0 & 0 & \beta_3 \end{bmatrix}$$

where $\beta_i \neq 0$. It is trivial to see such a system is controllable in the traditional sense. Next, consider a fault in the system that renders either one or more nonzero values of $\mathbf{B}$ to become a zero. Regardless of which input(s) failed, we wish for the LTI system to maintain controllability - which is USIC. To show this system is USIC, keep in mind Theorem 2. We recall $\mathbf{Z}_i^\top \mathbf{B}_j \neq 0$ and examine the left eigenvectors of $\mathbf{\Phi}$:

$$\mathbf{z}^\top = \begin{bmatrix} 0.5774 & -0.2887 - 0.5j & -0.2887 + 0.5j \\ 0.5774 & -0.2887 + 0.5j & -0.2887 - 0.5j \\ -0.5774 & -0.5774 & -0.5774 \end{bmatrix}$$

It can be verified that $\mathbf{z}_i^\top \mathbf{B}_j \neq 0$ holds true independent of the exact values of i and j selected. Therefore the system is controllable independent of which input(s) fail.

Controller Tasking Example

For this next example, consider the same LTI system formed by the pair $(\mathbf{A}, \mathbf{B})$ where $\mathbf{B}$ is a single input vector of the form $\mathbf{B} = [0, \ldots, \beta_i, \ldots, 0]^\top$. Considering the same $\mathbf{\Phi}$ from Section 4.2.4, the exact input vector may be selected from one of the options specified:

$$\mathcal{B} = \left\{ \begin{bmatrix} \beta_1 \\ 0 \\ 0 \end{bmatrix}, \ \begin{bmatrix} 0 \\ \beta_2 \\ 0 \end{bmatrix}, \ \begin{bmatrix} 0 \\ 0 \\ \beta_3 \end{bmatrix} \right\}$$

In order for a selection of the input vector to be valid, the pair $(\mathbf{A}, \mathbf{B})$ must be controllable. Extending this idea, in order for the set of potential input vectors to be valid, then the pair $(\mathbf{A}, \mathbf{B})$ must be controllable for each possible selection of $\mathbf{B} \in \mathcal{B}$. Using the same arguments as in Section 4.2.4 it can be shown that the LTI system is controllable independent of the exact selection of $\mathbf{B} \in \mathcal{B}$.

Controlling Into USIC

Next consider an LTI system with two input matrices of the form $\mathbf{B}_1$ and $\mathbf{B}_2$. The only assumption made about $\mathbf{B}_1$ is that the pair $(\mathbf{A}, \mathbf{B}_1)$ must be completely controllable and $\mathbf{\Phi}$ is not necessarily in a USIC form. The matrix $\mathbf{B}_2$ may take any form of $\mathbf{B} = [0, \ldots, \beta_i, \ldots, 0]^\top$. Because the pair $(\mathbf{A}, \mathbf{B}_1)$ is completely controllable, then there exists some matrix $\mathbf{K}_1$ such that the closed-loop matrix $\mathbf{\Phi}_{cl} = \mathbf{A} + \mathbf{B}_1\mathbf{K}_1$ takes on a USIC form. This is directly analogous to the standard pole-placement procedure from classical control theory. Therefore $\mathbf{\Phi}_{cl}$ is controllable independent of the exact matrix $\mathbf{B}_2$.

Independent Observer Example

The independent observer example considers an LTI system with multiple estimators where each estimator is only measuring a single state but still wants to estimate the state of the entire system. It is a well-known result from estimation theory, that the estimator will converge if and only if the LTI system $(\mathbf{A}, \mathbf{C})$ forms an observable pair. Consider the example in Figure 4.3 with four independent estimators, each observing a single state of the system with no communication between the estimators. The LTI system visualized in

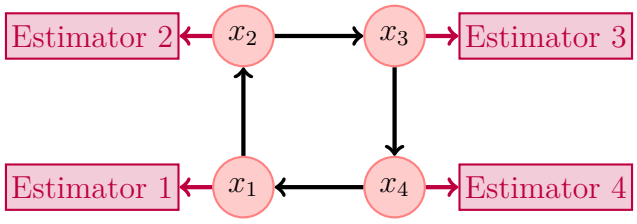


Figure 4.3 Independent Observer Example. Each estimator is only measuring the state it is linked to, but still wants to estimate the state of the entire system.

Figure 4.3 takes the matrix form

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x, \ \mathbf{y}_i = \mathbf{i}^\top \mathbf{x}$$

where $\mathbf{y}_i$ is the measurement available to Estimator i. In order for each estimator to converge the pair $(\mathbf{A}, \mathbf{i}^\top)$ must be observable for each $\mathbf{i}^\top$ - therefore such a system must be USOO.

Studying criterion for the ubiquitous observability (and controllability) has afforded us critical insight for the state omniscience of multi-agent systems. The next section will explore the theory of ubiquitous observability to multi-agent systems and generalize it to a theory of state omniscience.

4.3 State Omniscience, Ubiquitous Observability, and Ubiquitous Detectability

Ultimately, the objective of the augmented distributed estimator in (4.9) is to ensure

$$\lim_{k\to\infty} ||\hat{\mathbf{x}}_k^i - \mathbf{x}_k|| = 0, \ \text{for each } i \in \mathcal{V}$$

more specifically, the error dynamics of (4.11) are convergent to the origin for each $i \in \mathcal{V}$. If this holds true, we say the distributed estimator is state omniscient [74, 82, 83]. It follows immediately that a distributed estimator is state omniscient if there exists fusion matrices $\mathbf{W}$, $\mathbf{U}$, and gain matrix $\mathbf{K}$ that stabilizes (4.11). The remainder of this section discusses state omniscience and its variations that arise under specific design considerations, namely, the exclusion of predictor fusion. Particularly, we are interested in answering the question, "under what graph topologies will $\mathbf{W}$ and $\mathbf{U}$ exist?" The existence of $\mathbf{K}$ is well-studied in linear control-theory and, therefore, is not discussed extensively within the present work; yet, we want to explore the existence of fusion matrices $\mathbf{W}$, $\mathbf{U}$ as

well as how the matrices $\mathbf{K}$, $\mathbf{W}$, and $\mathbf{U}$ interact with one another. We consider the observable set at vertex i to be $\mathcal{O}_i \subseteq \mathcal{X}$ with $r \leq n$ basis vectors $\mathcal{O}_i := \{\mathbf{o}_1^i, \ldots, \mathbf{o}_r^i\}$. More specifically, we say for the eigenpair $(\lambda, \mathbf{v})$ that $\mathbf{v} \in \mathcal{O}_i$ if and only if $\mathbf{C}_i\mathbf{v} \neq 0$, which is the PBH-lemma [154] which implies the existence of $\mathbf{K}_i$ that stabilizes the estimate error for each of the modes $\mathbf{v}_l \in \mathcal{O}_i$. In the case that $r = n$, we say the system is observable from vertex i. Alternatively when $r < n$, we say the system is unobservable from vertex i. It is a well-known result from linear control theory that a mode may only be stabilized by $\mathbf{K}_i$ if that mode is observable; unobservable modes cannot be stabilized by any $\mathbf{K}_i$. Formally, we may define the local observable set as

$$\mathcal{O}_i := \{\mathbf{v} \mid \mathbf{C}_i\mathbf{v} \neq \mathbf{0}\}. \tag{4.24}$$

Similarly, we define the neighborhood observable set as

$$\bar{\mathcal{O}}_i := \left\{\mathbf{z} \mid \sum_{j \in \mathcal{N}_i} \mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j\mathbf{z} \neq \mathbf{0}\right\} \tag{4.25}$$

which may be interpreted as the set of modes $\mathbf{v}$ that are made observable via output fusion. Prior to presenting the primary result built on (4.24) and (4.25), we must first introduce the following lemmas.

Lemma 5.1. *There exists a collection of matrices* $\{\mathbf{U}_{ij}\}_{j \in \mathcal{N}_i}$ *such that* $\sum_{j \in \mathcal{N}_i} \mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j\mathbf{z} \neq \mathbf{0}$ *if and only if* $\mathbf{z} \in \mathcal{O}_j$ *for any* $j \in \mathcal{N}_i$.

Proof. The necessary condition follows immediately from the PBH-lemma [154] and (4.24) - if $\mathbf{C}_j\mathbf{z} = \mathbf{0}$, then there does not exist a $\mathbf{U}_{ij}$ such that $\mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j\mathbf{z} \neq \mathbf{0}$. The sufficient condition follows from the fact that $\mathbf{U}_{ij}$ is a tunable parameter and may be constructed such that $\mathbf{z} \notin \text{Null}(\sum_{j \in \mathcal{N}_i} \mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j)$. □

To shed some light on Lemma 5.1 consider the illustrative example.

Illustrative Example: Consider the sets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$ to be subspaces of vector space $\mathcal{V}$. Let

$$\mathcal{A} = \begin{bmatrix} t & -t \\ t & -t \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} -t & 2t \\ -t & 2t \end{bmatrix}, \ \mathbf{z} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

It can easily be verified that $\mathbf{z} \in \text{Im}(\mathcal{A}), \text{Im}(\mathcal{B})$ but $\mathbf{z} \in \text{Null}(\mathcal{A} + \mathcal{B})$, where

$$\mathcal{A} + \mathcal{B} := \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathcal{A}, \ \mathbf{b} \in \mathcal{B}\}$$

is the subspace sum. Next, let us scale each of the subspaces such that

$$\mathcal{A} = \begin{bmatrix} \frac{t}{3} & -\frac{t}{3} \\ \frac{t}{3} & -\frac{t}{3} \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} -\frac{2t}{3} & \frac{4t}{3} \\ -\frac{2t}{3} & \frac{4t}{3} \end{bmatrix}, \ \mathbf{z} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The eigenvector $\mathbf{z} \in \text{Im}(\mathcal{A}), \ \text{Im}(\mathcal{B})$ and is also $\mathbf{z} \in \text{Im}(\mathcal{A} + \mathcal{B})$.

Lemma 5.2. *Given vectors* $\mathbf{u}_j$ *and* $\mathbf{z}$*, and assuming Lemma 5.1 holds then* $\mathbf{u}_j \in \mathcal{O}'_j$ *if and only if* $\mathbf{z} \in \mathcal{O}_j$*.*

Proof. First, let us assume that $\mathbf{z} \in \mathcal{O}_j$ and $\mathbf{u}_j \notin \mathcal{O}'_j$. This implies that $\mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} \neq \mathbf{0}$ and $\mathbf{U}_{ij} \mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} = \mathbf{0}$ which contradicts Lemma 5.1. Next, assume that $\mathbf{z} \notin \mathcal{O}_j$ and $\mathbf{u}_j \in \mathcal{O}'_j$. This implies that $\mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} = \mathbf{0}$ and $\mathbf{U}_{ij} \mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} \neq \mathbf{0}$. However, $\mathbf{u}_j = \mathbf{U}_{ij} \mathbf{C}_j^\top \mathbf{C}_j \mathbf{z}$ and because $\mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} = \mathbf{0}$ it follows that $\mathbf{U}_{ij} \mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} = \mathbf{0}$ which contradicts the assumption that $\mathbf{u}_j \in \mathcal{O}'_j$. □

Lemma 5.3. *Given vectors* $\mathbf{u}$ *and* $\mathbf{z}$ *and assuming Lemma 5.1 holds, then* $\mathbf{u} \in \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$ *if and only if* $\mathbf{z} \in \bar{\mathcal{O}}_i$*.*

Proof. Assume that $\mathbf{z} \in \bar{\mathcal{O}}_i$ which implies that $\mathbf{U}_{ij} \mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} \neq \mathbf{0}$. While formally $\mathbf{0} \in \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$, Lemma 5.1 precludes $\mathbf{u} = \mathbf{0}$ from occurring; therefore, we consider $\mathbf{u} \in \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$ to imply $\mathbf{u} \neq \mathbf{0}$. Next, assume that $\mathbf{u} \notin \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$ and that Lemma 5.1 holds

which implies $\mathbf{u} \neq \mathbf{0}$. Expanding $\mathbf{u}$ yields $\mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j \mathbf{z} = \mathbf{0}$ which contradicts the original assumption that $\mathbf{z} \in \bar{\mathcal{O}}_i$.

Complementary, if we begin with the assumptions that $\mathbf{z} \notin \bar{\mathcal{O}}_i$ and $\mathbf{u} \in \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$ an identical contradiction arises. □

In fact, considering the preceding lemmas alongside (4.24) and (4.25) we may introduce the following theorem:

Theorem 6. *The neighborhood observable set $\bar{\mathcal{O}}_i$ is equivalent to the subspace sum of $\mathcal{O}'_j \subseteq \mathcal{O}_j$ for each $j \in \mathcal{N}_i$. More specifically,*

$$\bar{\mathcal{O}}_i \equiv \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j.$$

Proof. Assuming $\mathbf{U}_{ij}$ exists according to Lemma 5.1, let us define $\mathbf{u}_j = \mathbf{U}_{ij}\mathbf{C}_j^\top \mathbf{C}_j \mathbf{z}$, $\mathbf{u} = \sum_{j \in \mathcal{N}_i} \mathbf{u}_j$, and

$$\mathcal{O}'_j := \{\mathbf{u}_j \mid \mathbf{u}_j \neq \mathbf{0}\}. \tag{4.26}$$

We may substitute these definitions into (4.25) which yields

$$\bar{\mathcal{O}}_i := \left\{ \mathbf{z} \mid \sum_{j \in \mathcal{N}_i} \mathbf{u}_j \neq \mathbf{0} \right\}. \tag{4.27}$$

As a consequence of Lemma 5.2, each output fusion submatrix $\mathbf{U}_{ij}$ maps $\mathbf{z} \in \mathcal{O}_j \rightarrow \mathbf{u}_j \in \mathcal{O}'_j$. Using the definition of subspace sums, we make the statement

$$\sum_{j \in \mathcal{N}_i} \mathcal{O}'_j := \left\{ \sum_{j \in \mathcal{N}_i} \mathbf{u}_j \mid \mathbf{u}_j \in \mathcal{O}'_j \right\}. \tag{4.28}$$

Moreover, Lemma 5.2 may be expanded to form Lemma 5.3. The result follows as a direct consequence of Lemma 5.3 proving $\mathbf{u} \in \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j$ and $\mathbf{z} \in \bar{\mathcal{O}}_i$ are equivalent statements. □

4.3.1 Ubiquitous Observability

The case where no local state estimate is shared amongst its neighbors is a scenario more akin to USOO; but in its current form, both Definition 9 of USOO and Corollary 2.1 are only applicable to single-agent LTI systems. Therefore, we provide an extension of Definition 9.

Definition 10. ***(Ubiquitous Observability)*** *A distributed estimator is ubiquitously observable (UO) if and only if the global system state is observable at each vertex $i \in \mathcal{V}$.*

Under this definition, we introduce the following lemma that leads to Corollary 6.1:

Lemma 6.1. *UO is state omniscience without predictor-fusion.*

Proof. The no predictor-fusion case is equivalent to setting the predictor fusion matrix $\mathbf{W} = \mathbf{I}$. We allow each local estimator to only share its current measurement information amongst its neighbors. This measured information is incorporated into the local measurements through the output fusion matrix $\mathbf{U}$. However, this fusion only affects the local state estimate and is not passed along either directly or indirectly through local state estimate sharing. The fact that $\mathbf{W} = \mathbf{I}$ implies that each local estimator must be convergent for the system to be state omniscient; which only occurs if the global system states are observable from each vertex i for each $i \in \mathcal{V}$. Thus concludes the proof. □

Corollary 6.1. *A distributed estimator is UO if and only if*

$$\bar{\mathcal{O}}_i \equiv \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j = \mathcal{X}$$

more specifically, the sum of the observable subspaces within the neighbor set $\mathcal{N}_i$ spans the state space for each $i \in \mathcal{V}$.

Proof. Lemma 6.1 explicitly states that ubiquitous observability is state omniscience without predictor-fusion implying that each $i \in \mathcal{V}$ must be independently observable to

yield $\mathbf{K}$ that stabilizes 4.11. The statement

$$\bar{\mathcal{O}}_i \equiv \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j \subset \mathcal{X}$$

indicates the existence of at least one mode that is not in the convergent subspace of vertex i, implying there does not exist $\mathbf{U}$ nor $\mathbf{K}$ that may converge that mode of the system. Conversely,

$$\mathcal{O}_i \equiv \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j = \mathcal{X}$$

indicates the existence of $\mathbf{U}$ and/or $\mathbf{K}$ that may converge all modes of the system. Thus concludes the proof. □

4.3.2 Ubiquitous Detectability

In some cases, it may not be feasible for the entire distributed estimator to be observable [158], rather, the goal may then become to observe the unstable modes - a phenomena known as detectability. Similarly, there exists an analogous form of ubiquitous observability we will refer to as ubiquitous detectability.

Definition 11. ***(Ubiquitous Detectability)*** *A distributed estimator is considered to be ubiquitously detectable (UD) if at each agent the unobservable modes are asymptotically stable to the origin.*

Corollary 6.2. *A distributed estimator is ubiquitously detectable if and only if*

$$\bar{\mathcal{O}}_i \equiv \sum_{j \in \mathcal{N}_i} \mathcal{O}'_j + \mathcal{S} = \mathcal{X}$$

more specifically, the sum of the stable subspace, $\mathcal{S}$, and observable subspaces within the neighbor set $\mathcal{N}_i$ spans the state space for each $i \in \mathcal{V}$.

Proof. We define the stable subspace to be

$$\mathcal{S} := \{\mathbf{v} \mid (\lambda, \mathbf{v}) : \ \text{Real}(\lambda) < 0\} \tag{4.29}$$

and consider the eigenvector $\mathbf{z} \notin \bar{\mathcal{O}}'_j$ but $\mathbf{z} \in \mathcal{S}$. The rest of the proof follows from Corollary 6.1 and the definition of detectability. □

4.3.3 State Omniscience

When the agents choose to share their local state estimates with their neighbors, our first concern is the existence of $\mathbf{W}$ that may fuse the local state estimate at vertex i with the incoming state estimates from vertices $j \in \mathcal{N}_i$. Next we define a generalization of the observable subspace as the convergent subspace with $r \leq l \leq n$ basis vectors $\mathcal{C}_i := \{\mathbf{c}_1^i, \ldots, \mathbf{c}_l^i\}$ such that $\mathcal{O}_i \subseteq \bar{\mathcal{O}}_i \subseteq \mathcal{C}_i$. In fact, $\mathbf{z} \in \mathcal{C}_i$ if there exists a prediction-fusion matrix $\mathbf{W}$, innovation gain $\mathbf{K}$, or output-fusion gain matrix $\mathbf{U}$ that stabilizes the mode.

Theorem 7. *Given an observable state estimator in* (4.9) *then an eigenvector* $\mathbf{z} \in \mathcal{C}_i$ *for an* $i \in \mathcal{V}$ *if and only if* $\mathbf{W}_i\mathbf{K}\mathbf{C}_\mathbf{A}(\mathbf{1} \otimes \mathbf{z}) \neq \mathbf{0}$.

Proof. Note that if $\mathbf{C}_\mathbf{A}(\mathbf{1} \otimes \mathbf{z}) = \mathbf{0}$, then $\mathbf{W}_i\mathbf{K}\mathbf{C}_\mathbf{A}(\mathbf{1} \otimes \mathbf{z}) = \mathbf{0}$. Consequently, we assume that $\mathbf{C}_\mathbf{A}\mathbf{z} \neq \mathbf{0}$ for the remainder of the proof. Under this assumption, we make the observation that for an example N-agent system if $\mathbf{z} \in \mathcal{O}_1$ but $\mathbf{z} \notin \mathcal{O}_2, \ldots, \mathcal{O}_N$ then

$$\mathbf{K}\mathbf{C}_\mathbf{A}(\mathbf{1} \otimes \mathbf{z}) = \begin{bmatrix} \mathbf{K}_1\mathbf{C}_\mathbf{A}^1\mathbf{z} \\ \mathbf{K}_2\mathbf{C}_\mathbf{A}^2\mathbf{z} \\ \vdots \\ \mathbf{K}_N\mathbf{C}_\mathbf{A}^N\mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_1\mathbf{C}_\mathbf{A}^1\mathbf{z} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \tag{4.30}$$

Note that while $\mathbf{C}_\mathbf{A}\mathbf{z} \neq \mathbf{0}$, it is possible that $\mathbf{C}_\mathbf{A}^i\mathbf{z} = \mathbf{0}$ for a given i. Next, we must show that there exist at least a single walk from a nonzero partition of $\mathbf{C}\mathbf{z}$ to vertex i. Here, we define block generalizations of $\mathbf{s}$, $\mathbf{t}$, and $\mathbf{W}_i$ from (2.3) and (2.4) such that

$$\bar{\mathbf{s}} = \mathbf{K}\mathbf{C}_\mathbf{A}(\mathbf{1} \otimes \mathbf{z}) \text{ and } \bar{\mathbf{t}} = \mathbf{J}_i^\top \tag{4.31}$$

for the target vertex i and

$$\mathbf{W}_i = \begin{bmatrix} \mathbf{J}_i \\ \mathbf{J}_i(\mathbf{W} \otimes \mathbf{\Phi}) \\ \vdots \\ \mathbf{J}_i(\mathbf{W} \otimes \mathbf{\Phi})^d \end{bmatrix}. \tag{4.32}$$

At first glance, these definitions of $\mathbf{s}$ and $\mathbf{t}$ may seem significantly different then those presented in (2.3) and (2.4). However, consider the example in (4.30) for which $\bar{\mathbf{s}}$ would take on the form $\bar{\mathbf{s}} = [\bar{\mathbf{s}}_1^\top, \ldots, \bar{\mathbf{s}}_N^\top]^\top = [\times, \ldots, \times, \mathbf{0}^\top, \ldots, \mathbf{0}^\top]^\top$ where $\times$ represents an non-zero value. We may then define

$$\mathbf{s}_j = \begin{cases} 1, & \bar{\mathbf{s}}_j \neq \mathbf{0} \\ 0, & \text{otherwise} \end{cases} \qquad \mathbf{t}_j = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases}. \tag{4.33}$$

Physically, this means the form in (4.31) considers the graph interconnections at the state-level defined by the dynamics and the agent-level defined by $\mathcal{G}_c$; where, on the other hand, (4.33) only considers the graph interconnections at the agent-level. In fact, all of the walks defined by $\mathcal{G}_c$ are independent of the underlying states. The interpretation of the walk matrix, consequently, remains unchanged regardless of whether the walk matrix is constructed via (4.33) leading to (2.4) or (4.31) leading to (4.32). Therefore, $\mathbf{W}_i\mathbf{s} = \mathbf{0}$, independent of how the walk matrix is constructed, implies that there does not exist a walk from a vertex j where $\mathbf{z}_l \in \mathcal{O}_j$ to vertex i. Conversely, if $\mathbf{W}_i\mathbf{s}_j \neq \mathbf{0}$ implies that there exists at least one walk from a vertex j where $\mathbf{z} \in \mathcal{O}_j$ to vertex i. □

The convergent subspace $\mathcal{C}_i$ not only encapsulates the observable information at vertex i, but also the convergent information provided by its neighbors. Consequently, we define this set at vertex i to be

$$\mathcal{C}_i := \{\mathbf{z} \mid \mathbf{W}_i\mathbf{K}\mathbf{C}_\mathbf{A}\mathbf{z} \neq \mathbf{0}\}. \tag{4.34}$$

It is apparent this also implies the existence of $\mathbf{W}$ that will converge the modes $\mathcal{C}_i \backslash \mathcal{O}_i$. In

fact, we may also define the set

$$\mathcal{C}_i' := \left\{ \mathbf{z} \mid \sum_{s=0}^{k-1} \mathbf{J}_i (\mathbf{W} \otimes \mathbf{\Phi})^s \mathbf{K} \mathbf{C}_\mathbf{A} (\mathbf{1} \otimes \mathbf{z}) \neq \mathbf{0} \right\}.$$

Prior to directly considering the consequences of $\mathcal{C}_i'$, we introduce an analogous form of Lemma 5.1:

Lemma 7.1. *There exists a collection of matrices* $\{\mathbf{W}_{ij}\}$ *such that* $\sum_{s=0}^{k-1} \mathbf{W}_{is} \mathbf{K} \mathbf{C}_\mathbf{A} (\mathbf{1} \otimes \mathbf{z}) \neq \mathbf{0}$ *if and only if* $\mathbf{z} \in \mathcal{O}_j$ *for any* $j \in \mathcal{W}_i$.

Proof. We begin by setting $k - 1 = d$ which allows us to encapsulate all of the walks that end at agent i. The remainder of the proof follows from the proof of Lemma 5.1. □

Lemma 7.2. *An eigenvector* $\mathbf{z} \in \mathcal{C}_i$ *if and only if* $\mathbf{z} \in \mathcal{C}_i'$.

Proof. From the definition of the walk matrix, $\mathcal{C}_i'$ may be rewritten as

$$\mathcal{C}_i' := \left\{ \mathbf{z} \mid \sum_{s=0}^{k-1} \mathbf{W}_{is} \mathbf{K} \mathbf{C}_\mathbf{A} (\mathbf{1} \otimes \mathbf{z}) \neq \mathbf{0} \right\}$$

where $\mathbf{W}_{is}$ is the s^{th} row block of $\mathbf{W}_i$. Let us first assume that $\mathbf{z} \in \mathcal{C}_i'$, it follows that from the proof of Theorem 7 that there exists a walk from an agent j with $\mathbf{z} \in \mathcal{O}_j$ of length less than $k - 1$ implying that $\mathbf{z} \in \mathcal{C}_i$. Next assume that $\mathbf{z} \in \mathcal{C}_i$, it also follows that from the proof of Lemma 7.1 that there exists a walk from agent j with $\mathbf{z} \in \mathcal{O}_j$ of length less than $k - 1$ implying that $\mathbf{z} \in \mathcal{C}_i'$. □

Corollary 7.1. *A distributed estimator is state omniscient if and only if*

$$\mathcal{C}_i \equiv \sum_{j \in \mathcal{W}_i} \mathcal{O}_j' = \mathcal{X} \tag{4.35}$$

more specifically, the sum of the observable subspaces within the walk set $\mathcal{W}_i$ *spans the state space for each* $i \in \mathcal{V}$. *If there exists a walk* $j \to i$, *then we say* $j \in \mathcal{W}_i$. *We also assume* i *to be in its own extended walk set.*

Proof. It follows from Lemma 7.2 and from the definition of vector subspace sums that

$$\mathcal{C}_i \equiv \sum_{j \in \mathcal{W}_i} \mathcal{O}'_j = \mathcal{X}.$$

Therefore,

$$\mathcal{C}_i \equiv \sum_{j \in \mathcal{W}_i} \mathcal{O}'_j \subset \mathcal{X}$$

indicates the existence of at least one mode that is not in the convergent subspace of vertex i, implying there does not exist $\mathbf{W}$ nor $\mathbf{K}$ that may converge that mode of the system. Conversely,

$$\mathcal{C}_i \equiv \sum_{j \in \mathcal{W}_i} \mathcal{O}'_j = \mathcal{X}$$

indicates the existence of $\mathbf{W}$ and/or $\mathbf{K}$ that may converge all modes of the system. Thus concludes the proof. □

Remark 1. *It should be noted that Corollary 7.1 does not necessarily place any restrictions on the connectivity of the communication network. From the definition of strong connectivity and Corollary 7.1, a strongly connected network is sufficient to satisfy the conditions of the theorem, but is not necessary. A weakly connected network will also satisfy the conditions given appropriate edge placement. Additionally, an unconnected network can be state omniscient provided the connected components satisfy the conditions of Corollary 7.1. For example, consider the networked system shown in Figure 2.2 and the communication graph and sensing graph depicted in Figure 2.3a and Figure 2.3b, respectively. It can be shown that the networked system is state omniscient even though $\mathcal{G}_c$ is unconnected. Agent 3 directly measures all other agents within the network; while Agent 1 does not take any measurements directly, Agent 2 communicates its estimates to Agent 1; agents 4 and agents 5 work together to gather measurements of the remaining agents and then communicate their estimates to one another.*

Moreover, the convergent subspace condition in (4.35) may be generalized to a

detectable form as

$$\mathcal{C}_i := \sum_{j \in \mathcal{W}_i} \mathcal{O}'_j + \mathcal{S}.$$

or $\mathcal{S}$ may be implicitly included as a subset of $\mathcal{C}_i$. Such a system may also be considered state omniscient even if $\sum_{j \in \mathcal{W}_i} \mathcal{O}'_j - \mathcal{S} = \emptyset$.

4.4 Main Results

In practice, however, to evaluate the omniscient properties of distributed estimators, a method of quantifying state omniscience beyond set-theoretic approaches is necessary. Consequently, this section discusses state omniscient in a quantifiable and computable way; this method is analogous to classical observability analysis tools.

4.4.1 The Omniscient Matrix

Consider a discrete-time Luenberger-style observer as a linear difference equation

$$\hat{\mathbf{x}}(k) = \bar{\mathbf{\Phi}}(k, k-1)\hat{\mathbf{x}}(k-1) + \mathbf{LCe}(k-1)$$

where $\mathbf{L} = \text{diag}([\mathbf{L}_1, \ldots, \mathbf{L}_N])$, $\hat{\mathbf{x}}(k) \in \mathbb{R}^{nN^2}$ represents all of the state estimates across the network, and $\mathbf{e}(k) = \bar{\mathbf{x}}(k) - \hat{\mathbf{x}}(k)$ with $\bar{\mathbf{x}}(k) = \mathbf{1}_N \otimes \mathbf{x}(k)$ is the estimate error. The linear difference equation has the solution

$$\hat{\mathbf{x}}(k) = \bar{\mathbf{\Phi}}(k, 0)\hat{\mathbf{x}}(0) + \sum_{s=0}^{k-1} \bar{\mathbf{\Phi}}(s, 0)\mathbf{LCe}(s) \tag{4.36}$$

where

$$\bar{\mathbf{\Phi}}(k, 0) = (\mathbf{W} \otimes \mathbf{\Phi})^k. \tag{4.37}$$

The matrix $\mathbf{W} \in \mathbb{R}^{nN \times nN}$ is a tunable matrix constrained to the structure of the adjacency matrix $\mathbf{A}$ such that

$$(\mathbf{W})_{ij} = \begin{cases} \mathbf{0}_n, \ a_{ij} = 0 \\ \mathbf{W}_{ij}, \ a_{ij} = 1 \end{cases}$$

where $\mathbf{W}_{ij} \in \mathbb{R}^{n\times n}$. The operator $\otimes$ is a block-matrix generalization of the standard Kronecker product denoted $\mathbf{W} \otimes \mathbf{\Phi}$ that may be applied to block matrices of $\mathbf{W}$, namely $\mathbf{W}_{ij}$, such that $(\mathbf{W} \otimes \mathbf{\Phi})_{ij} = \mathbf{W}_{ij}\mathbf{\Phi} \in \mathbb{R}^{n\times n}$ given $\mathbf{W} \in \mathbb{R}^{nN\times nN}$, $\mathbf{\Phi} \in \mathbb{R}^{n\times n}$, and $\mathbf{W} \otimes \mathbf{\Phi} \in \mathbb{R}^{nN\times nN}$. Substituting (4.37) into (4.36) yields

$$\hat{\mathbf{x}}(k) = (\mathbf{W} \otimes \mathbf{\Phi})^k \hat{\mathbf{x}}(0) + \sum_{s=0}^{k-1} (\mathbf{W} \otimes \mathbf{\Phi})^{k-1-s} \mathbf{L}\mathbf{C}\mathbf{e}(s). \tag{4.38}$$

Expanding (4.38) with the understanding that $\mathbf{e}(s) = \hat{\mathbf{x}}(s) - \mathbf{x}(s)$ with $\mathbf{x}(s) = (\mathbf{I}_N \otimes \mathbf{\Phi})^s (\mathbf{1}_N \otimes \mathbf{x}(0))$ where $\mathbf{1}_N$ is the N-dimensional vector of all ones such that $\mathbf{e}(s) = (\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{L}\mathbf{C})^s \mathbf{e}(0)$. Placing all state estimate dependent terms on the right-hand side yields

$$\begin{aligned}\sum_{s=0}^{k-1} (\mathbf{W} \otimes \mathbf{\Phi})^{k-1-s} \mathbf{L}\mathbf{C} (\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{L}\mathbf{C})^s \mathbf{x}(0) \\ = (\mathbf{W} \otimes \mathbf{\Phi})^k \hat{\mathbf{x}}(0) - \hat{\mathbf{x}}(k) \\ + \sum_{s=0}^{k-1} (\mathbf{W} \otimes \mathbf{\Phi})^{k-1-s} \mathbf{L}\mathbf{C}\hat{\mathbf{x}}(s).\end{aligned} \tag{4.39}$$

We will refer to the right-hand side of (4.39) as $\mathbf{p}(k)$, therefore (4.39) may be rewritten as

$$\mathbf{p}(k) = \sum_{s=0}^{k-1} (\mathbf{W} \otimes \mathbf{\Phi})^{k-1-s} \mathbf{L}\mathbf{C} (\mathbf{I} \otimes \mathbf{\Phi})^s \mathbf{x}(0). \tag{4.40}$$

Stacking the time history of $\mathbf{p}$ up to timestep k results in

$$\begin{bmatrix} \mathbf{p}(1) \\ \mathbf{p}(2) \\ \vdots \\ \mathbf{p}(k) \end{bmatrix} = \mathbf{M}\mathbf{x}(0) \tag{4.41}$$

where

$$\mathbf{M} = \begin{bmatrix} \mathbf{LC} \\ (\mathbf{W} \otimes \mathbf{\Phi})\mathbf{LC} + \mathbf{LC}(\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{LC}) \\ \vdots \\ \sum_{s=0}^{k-1} (\mathbf{W} \otimes \mathbf{\Phi})^{k-1-s} \mathbf{LC} (\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{LC})^s \end{bmatrix} \in \mathbb{R}^{nN^2k \times nN}.$$

Coincidentally, $\mathbf{M}$ resembles a coordinate-change of the observability matrix utilized for time-delay systems [159]. Our considered case, generalizes the results by restricting the coordinate-change matrix to the structure of the communication graph. We utilize (4.41) and Definition 7 to construct a mathematical characterization of state omniscience. Furthermore, we may apply a permutation matrix such that

$$\mathbf{P} \begin{bmatrix} \mathbf{p}(1) \\ \mathbf{p}(2) \\ \vdots \\ \mathbf{p}(k) \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1(1) \\ \vdots \\ \mathbf{p}_1(k) \\ \vdots \\ \mathbf{p}_N(1) \\ \vdots \\ \mathbf{p}_N(k) \end{bmatrix} = \mathbf{PMx}(0);$$

which results in subvectors $\mathbf{p}_i = [\mathbf{p}_i^\top(1), \ldots, \mathbf{p}_i^\top(k)]^\top$ that group all of the information available to observer i for each timestep $0 < s \leq k$.

Theorem 8. *A distributed estimator is state omniscient if and only if* $\mathbf{x}(0)$ *may be reconstructed from* $\mathbf{p}_i$ *for each* $i \in \mathcal{V}$ *such that*

$$\mathbf{x}(0) = (\mathbf{PM})_i^\dagger \mathbf{p}_i.$$

where $(\mathbf{PM})_i$ *indicates the* i^{th} *row-block of* $\mathbf{PM}$*. For brevity we will from now on refer to* $(\mathbf{PM})_i$ *as* $\mathbf{M}_i$*. In other words,* $\mathbf{M}_i$ *must be left-unimodular for each* $i \in \mathcal{V}$*.*

Proof. Let us treat $\mathbf{M}_i\mathbf{x}(0) = \mathbf{p}_i$ as a least-squares problem such that

$$\arg\min_{\mathbf{x}(0)} ||\mathbf{M}_i\mathbf{x}(0) - \mathbf{p}_i||_2$$

is sought. It is known that the general solution to the least squares problem takes the form

$$\mathbf{x}(0) = \begin{bmatrix} \mathbf{V}_r & \mathbf{V}_{n-r} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_r^{-1}\mathbf{D}^\top\mathbf{p}_i \\ \mathbf{b}_{n-r} \end{bmatrix} = \mathbf{M}_i^\dagger\mathbf{p}_i + \mathbf{V}_{n-r}\mathbf{b}_{n-r}$$

implying $\mathbf{x}(0)$ may take any vector along the line define by $\mathbf{M}_i^\dagger\mathbf{p}_i + \mathbf{V}_{n-r}\mathbf{b}_{n-r}$ for any $\mathbf{b}_{n-r} \in \mathbb{R}^{n-r}$. Consequently for any $r < n$, the solution to the least-sqares problem is nonunique. However if $r = n$, $\mathbf{V}_{n-r} = \mathbf{0}$, and therefore $\mathbf{x}(0) = \mathbf{M}_i\mathbf{p}_\mathbf{p}$ and may be uniquely determined. The matrix $\mathbf{M}_i$ is full rank when $r = n$ indicating there exists a simple algebraic expression $\mathbf{M}_i^\dagger(\mathbf{M}_i^\top\mathbf{M}_i)^{-1}\mathbf{M}_i^\top$ and $\mathbf{M}_i^\dagger\mathbf{M}_i = \mathbf{I}_N$. □

But under what conditions will $\mathbf{M}_i$ be left-unimodular for each $i \in \mathcal{V}$? To further this discussion we introduce the following lemma:

Lemma 8.1. *Given an eigenpair $(\lambda, \mathbf{v}) = \text{eig}(\mathbf{W} \otimes \boldsymbol{\Phi})$ where λ is the eigenvalue and $\mathbf{v}$ is its associated eigenvector, the mode $\mathbf{v} \in \mathcal{O}_{cl}$ if and only if $\mathbf{v} \in \mathcal{O}$ where $\mathcal{O}_{cl}$ and $\mathcal{O}$ are the closed-loop and open-loop observable subspaces, respectively.*

Proof. Suppose $(\mathbf{W} \otimes \boldsymbol{\Phi})\mathbf{v} = \lambda\mathbf{v},\ \mathbf{C}\mathbf{v} = \mathbf{0}$ corresponds to an unobservable mode of the open-loop system and implies $\mathbf{v} \notin \mathcal{O}$, then $(\mathbf{W} \otimes \boldsymbol{\Phi} + \mathbf{L}\mathbf{C})\mathbf{v} = (\mathbf{W} \otimes \boldsymbol{\Phi})\mathbf{v} + \mathbf{L}\mathbf{C}\mathbf{v} = \lambda\mathbf{v}$ so $\mathbf{v} \notin \mathcal{O}_{cl}$. Consequently, the state feedback cannot manipulate unobservable modes of the system. The converse may be proven by initially supposing $(\mathbf{W} \otimes \boldsymbol{\Phi})\mathbf{v} = \lambda\mathbf{v},\ \mathbf{C}\mathbf{v} \neq \mathbf{0}$. □

From Theorem 8 and making use of Lemma 8.1 we may induce the following corollary:

Corollary 8.1. *If a distributed estimator is state omniscient, then it is observable.*

Proof. It can be shown that $\mathbf{M}$ may be separated into three separate matrices such that

$$\mathbf{M} = \mathbf{F}\bar{\mathbf{L}}\mathbf{O}_{cl} \tag{4.42}$$

where $\bar{\mathbf{L}} = (\mathbf{I}_k \otimes \mathbf{L})$, $\mathbf{O}_{cl} \in \mathbb{R}^{pEk \times nN}$ is a generalized closed-loop observability matrix

$$\mathbf{O}_{cl} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}(\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{LC}) \\ \vdots \\ \mathbf{C}(\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{LC})^{k-1} \end{bmatrix},$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_{nN^2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi}) & \mathbf{I}_{nN^2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi})^2 & (\mathbf{W} \otimes \mathbf{\Phi}) & \mathbf{I}_{nN^2} & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi})^{k-1} & \dots & \dots & (\mathbf{W} \otimes \mathbf{\Phi}) & \mathbf{I}_{nN^2} \end{bmatrix}$$

with $\mathbf{F} \in \mathbb{R}^{nN^2k \times nN^2k}$. Considering (4.42) and rank properties of matrix products implies that $\text{rank}(\mathbf{M}) \leq \min(\text{rank}(\mathbf{F}), \text{rank}(\mathbf{I} \otimes \mathbf{L}), \text{rank}(\mathbf{O}_{cl}))$. By construction, $\text{rank}(\mathbf{F}) = nN^2k$ and because $\text{rank}(\mathbf{I} \otimes \mathbf{L})$ is tunable indicates that $\text{rank}(\mathbf{M}) \leq \text{rank}(\mathbf{O}_{cl})$. Therefore, it necessitates that $\text{rank}(\mathbf{O}_{cl}) = nN$ in order for $\text{rank}(\mathbf{M}) \leq nN$. In fact, we may replace $\mathbf{O}_{cl}$ with $\mathbf{O}$ where $\mathbf{O}^\top = [\mathbf{C}^\top, (\mathbf{W} \otimes \mathbf{\Phi})^\top, \dots, (\mathbf{W}^\top \otimes \mathbf{\Phi}^\top)^{k-1}]$ because of Lemma 8.1. Ultimately, Lemma 8.1 implies $\mathbf{O}_{cl}$ and $\mathbf{O}$ share the same subspace, and therefore, the same basis. Therefore, the notation used initially by (4.42) may be replaced such that

$$\mathbf{M} = \mathbf{F}\bar{\mathbf{L}}\mathbf{O}. \tag{4.43}$$

To illustrate the "not sufficient" condition, we consider the case when there is no predictor fusion amongst the agents such that $\mathbf{W} = \mathbf{0}$; this induces $\mathbf{F} = \mathbf{I}$ and $\mathbf{p}_i = \mathbf{y}_i$. Applying the

permutation matrix yields $\mathbf{PM} = \mathbf{PF}(\mathbf{I} \otimes \mathbf{L})\mathbf{O}$ such that

$$\mathbf{PF} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_1 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{J}_N & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_N & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_N \end{bmatrix}$$

where $\mathbf{J}_i = [\bar{\mathbf{J}}_1, \dots, \bar{\mathbf{J}}_N] \in \mathbb{R}^{nN \times nN^2}$ and

$$\bar{\mathbf{J}}_j = \begin{cases} \mathbf{0}, \ j \neq i \\ \mathbf{I}, \ j = i \end{cases} .$$

We then expand $\mathbf{M}_1$ as an illustrative example:

$$\mathbf{M}_1 = (\mathbf{PF})_1\mathbf{O} = \begin{bmatrix} \mathbf{J}_1\mathbf{L} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_1\mathbf{L} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_1\mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{C} \\ \mathbf{C}(\mathbf{I} \otimes \mathbf{\Phi}) \\ \vdots \\ \mathbf{C}(\mathbf{I} \otimes \mathbf{\Phi})^{k-1} \end{bmatrix}$$

similarly, for brevity, we will from now on refer to $(\mathbf{PF})_i$ as $\mathbf{F}_i$. By its construction, $\mathbf{J}_1$

extracts the measurements associated with agent 1 such that

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{L}_1\mathbf{C}_1 \\ \mathbf{L}_1\mathbf{C}_1(\mathbf{I}\otimes\mathbf{\Phi}) \\ \vdots \\ \mathbf{L}_1\mathbf{C}_1(\mathbf{I}\otimes\mathbf{\Phi})^{k-1} \end{bmatrix} = (\mathbf{I}\otimes\mathbf{L}_1)\mathbf{O}_1$$

which, by definition, is the observability matrix of agent 1. Extrapolating this result to each $i \in \mathcal{V}$ for all $\mathbf{PM}$ induces

$$(\mathbf{PM}) = (\mathbf{L}\otimes\mathbf{I}) \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_N \end{bmatrix} = \mathbf{PO}.$$

Because it is possible for $\text{rank}(\mathbf{O}) = nN$ and for $\text{rank}(\mathbf{O}_1) < nN$, then $\mathbf{x}(0)$ may not be reconstructed from $\mathbf{y}_1$ alone and the result follows. □

For a general case,

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{I}_{nN} & \mathbf{0} & \dots & \mathbf{0} \\ (\mathbf{W}\otimes\mathbf{\Phi})_i & \mathbf{I}_{nN} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ (\mathbf{W}\otimes\mathbf{\Phi})_i^{k-1} & \dots & (\mathbf{W}\otimes\mathbf{\Phi})_i & \mathbf{I}_{nN} \end{bmatrix} \tag{4.44}$$

with $\mathbf{F}_i \in \mathbb{R}^{nNk\times nNk}$. Consequently, the structure of $\mathbf{W}$ plays a non-trivial role in the construction of $\mathbf{M}$ and the state omniscience of the distributed estimator. So, for what structures of $\mathbf{W}$ will $\mathbf{M}$ meet the state omniscience condition? Prior to presenting the result, we must first introduce the following lemma found in [160].

Lemma 8.2. *(Theorem 6.11.2, [160]) The images of linearly independent vectors by a non-singular linear operator are linearly independent vectors.*

Proof. We make use of the lemma without proof, but direct the interested reader to the original reference. □

Theorem 9. *Each* $\mathbf{M}_i$ *is left-unimodular if and only if the observable subspaces within* $\mathcal{W}_i$ *span the state-space.*

Proof. In general, we consider each partition of the omniscient matrix as $\mathbf{M}_i = \mathbf{F}_i\mathbf{O}$. We assume each $\mathbf{M}_i$ may be constructed from a set of basis vectors such that $\mathbf{M}_i := \{\mathbf{m}_1^i, \dots, \mathbf{m}_{q_i}^i\}$ where q_i is the number of basis vectors for $\mathbf{M}_i$ and is equivalent to $\text{rank}(\mathbf{M}_i)$. From Theorem 8 we know that $q_i = nN$ for each $i \in \mathcal{V}$ is necessary and sufficient for state omniscience. Without loss of generality, we assume the similarity transform matrix $\mathbf{T} = \text{diag}([\mathbf{T}_1, \dots, \mathbf{T}_N])$ which yields the Kalman canonical form $\mathbf{C}\mathbf{T}^{-1} = \text{diag}([\mathbf{C}_1\mathbf{T}_1^{-1}, \dots, \mathbf{C}_N\mathbf{T}_N^{-1}])$ and $(\mathbf{I} \otimes \mathbf{T})(\mathbf{W} \otimes \mathbf{\Phi})(\mathbf{I} \otimes \mathbf{T}^{-1}) = (\mathbf{W} \otimes \mathbf{T}\mathbf{\Phi}\mathbf{T}^{-1})$ where $\mathbf{T}\mathbf{\Phi}\mathbf{T}^{-1} = \text{diag}([\mathbf{T}_1\mathbf{\Phi}_1\mathbf{T}_1^{-1}, \dots, \mathbf{T}_N\mathbf{\Phi}_N\mathbf{T}_N^{-1}])$ and where

$$\mathbf{C}_i\mathbf{T}_i^{-1} = \begin{bmatrix} \mathbf{C}_i^o & \mathbf{0} \end{bmatrix}$$

$$\mathbf{T}_i\mathbf{\Phi}_i\mathbf{T}_i^{-1} = \begin{bmatrix} \mathbf{\Phi}_i^o & \mathbf{0} \\ \mathbf{\Phi}_i^\times & \mathbf{\Phi}_i^{uo} \end{bmatrix}.$$

Consequently, the generalized observability matrix takes the form $(\mathbf{O}^o)^\top = [(\mathbf{C}^o)^\top, (\mathbf{C}^o\mathbf{\Phi}^o)^\top, \dots, (\mathbf{C}^o(\mathbf{\Phi}^o)^{k-1})^\top]$ where $\mathbf{C}^o = \text{diag}([\mathbf{C}_1^o, \dots, \mathbf{C}_N^o])$ and $\mathbf{\Phi}^o = \text{diag}([\mathbf{\Phi}_1^o, \dots, \mathbf{\Phi}_N^o])$. It follows that $\mathbf{M}_i = \mathbf{F}_i\mathbf{O}$ may take a Kalman canonical form $\mathbf{M}_i = \mathbf{F}_i\mathbf{O}^o$ where $\mathbf{\Phi}$ is replaced by $\mathbf{\Phi}^o$ in (4.44) to construct $\mathbf{F}_i$. Next, we choose a permutation matrix $\mathbf{Q}_1$ such that

$$\mathbf{F}_i\bar{\mathbf{L}}\mathbf{Q}_1^\top = \begin{bmatrix} \bar{\mathbf{F}}_i & \mathbf{0} \end{bmatrix}$$

where

$$\bar{\mathbf{F}}_i = \begin{bmatrix} \bar{\mathbf{F}}_{ii} & \bar{\mathbf{F}}_{ij} & \dots \end{bmatrix} \; \forall \; j \in \mathcal{W}_i$$

and each $\bar{\mathbf{F}}_{ii}$ and $\bar{\mathbf{F}}_{ij}$ takes the structure

$$\bar{\mathbf{F}}_{ij} = \begin{bmatrix} \mathbf{L}_j & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{W}_{ij}\mathbf{\Phi}_j^o\mathbf{L}_j & \mathbf{L}_j & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{W}_{ij}^{k-1}(\mathbf{\Phi}^o)_j^{k-1}\mathbf{L}_j & \dots & \mathbf{W}_{ij}\mathbf{\Phi}_j^o\mathbf{L}_j & \mathbf{L}_j \end{bmatrix}.$$

with $\bar{\mathbf{F}}_{ij} \in \mathbb{R}^{nNk \times pE_j k}$. In effect, $\mathbf{Q}_1^\top$ organizes $\mathbf{F}_i\bar{\mathbf{L}}$ into partitions comprised of the extended walk matrix for each $j \in \mathcal{W}_i$. We can then partition

$$\mathbf{Q}_1\mathbf{O}^o = \begin{bmatrix} \mathbf{O}^o_{j\in\mathcal{W}_i} \\ \mathbf{O}^o_{j\notin\mathcal{W}_i} \end{bmatrix}$$

where $\mathbf{O}^o_{j\in\mathcal{W}_i} = [\mathbf{O}^o_i, \dots, \mathbf{O}^o_j]\ \forall\ j \in \mathcal{W}_i$ and $\mathbf{O}^o_{j\notin\mathcal{W}_i}$ is comprised of all other observability matrices. Similarly, $\mathbf{Q}_1$ organized $\mathbf{O}^o$ into partitions comprised of the observability matrices for each $j \in \mathcal{W}_i$. Then

$$\mathbf{F}_i\mathbf{Q}_1^\top\mathbf{Q}_1\mathbf{O}^o = \begin{bmatrix} \bar{\mathbf{F}}_i & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{O}^o_{j\in\mathcal{W}_i} \\ \mathbf{O}^o_{j\notin\mathcal{W}_i} \end{bmatrix} = \bar{\mathbf{F}}_i\mathbf{O}^o_{j\in\mathcal{W}_i} \tag{4.45}$$

therefore, $\mathbf{M}_i$ is only dependent on the observable subset $\mathbf{O}^o_j$ for each $j \in \mathcal{W}_i$. We then apply QR-decomposition [161] using an orthogonal matrix $\mathbf{Q}_2$ such that

$$\mathbf{Q}_2^\top\mathbf{O}^o_{j\in\mathcal{W}_i} = \begin{bmatrix} \mathbf{Q}_{21}^\top \\ \mathbf{Q}_{22}^\top \end{bmatrix} \mathbf{O}^o_{j\in\mathcal{W}_i} = \begin{bmatrix} \mathbf{O}^{o'}_{j\in\mathcal{W}_i} \\ \mathbf{0} \end{bmatrix}$$

where $\mathbf{O}^{o'}_{j\in\mathcal{W}_i} \in \mathbb{R}^{nN\times nN}$ is upper-triangular. The same matrix $\mathbf{Q}_2$ also partitions

$$\bar{\mathbf{F}}_i\mathbf{Q}_2 = \begin{bmatrix} \bar{\mathbf{F}}_{i1} & \bar{\mathbf{F}}_{i2} \end{bmatrix}$$

such that

$$\bar{\mathbf{F}}_i \mathbf{Q}_2 \mathbf{Q}_2^\top \mathbf{O}^o_{j\in\mathcal{W}_i} = \bar{\mathbf{F}}_i \begin{bmatrix} \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{O}^{o'}_{j\in\mathcal{W}_i} \\ \mathbf{0} \end{bmatrix} = \bar{\mathbf{F}}_i \mathbf{Q}_{21} \mathbf{O}^{o'}_{j\in\mathcal{W}_i}$$

Because $\mathbf{Q}_2$ forms an orthogonal basis for $\mathbf{O}^o_{j\in\mathcal{W}_i}$, it is necessary that $\mathbf{Q}_2\mathbf{Q}_2^\top = \mathbf{I}_{nN}$ for state omniscience since

$$\text{rank}(\mathbf{M}_i) = \text{rank}(\bar{\mathbf{F}}_i \mathbf{Q}_2 \mathbf{Q}_2^\top \mathbf{O}^o_{j\in\mathcal{W}_i}).$$

For a moment, consider that $\mathbf{M}_i = \bar{\mathbf{F}}_i \mathbf{Q}_{21} \mathbf{O}^{o'}_{j\in\mathcal{W}_i}$ where $\text{rank}(\bar{\mathbf{F}}_i) = nNk$, by construction, and $\mathbf{Q}_{21}$ forms an orthogonal basis for $\mathbf{O}^o_{j\in\mathcal{W}_i}$ comprised of nN linear independent vectors. If we consider $\bar{\mathbf{F}}_i$ to be the nonsingular linear operator and $\mathbf{Q}_{21}$ to be the linear independent vectors, it follows that the vectors of $\mathbf{M}_i$ are also linearly independent. Or, $\text{rank}(\mathbf{M}_i) = nN$. Thus proves the sufficiency condition. □

The matrix $\mathbf{M}$ serves as a distributed estimator analogy to the observability matrix $\mathbf{O}$ for single estimator systems and can be utilized to set necessary and sufficient conditions on the distributed estimator to meet the desired convergent properties. It also has the added benefit of encapsulting both the observability information and graph walk information; each of which may be extrapolated and used for a variety of purposes as needed.

Corollary 9.1. *The omniscient matrix* $\mathbf{M}$ *may be written as* $\sum_{l=0}^{k-1}(\mathbf{P}\mathbf{H}_l)_i(\mathbf{I}\otimes\mathbf{L})\mathbf{O}$ *representing the observability matrices for all agents* $j \in \mathcal{W}_i$ *when* $k = nN + d$.

Proof. To begin, we take advantage of the lower triangular Toeplitz structure of $\mathbf{F}$ and decompose it such that

$$\mathbf{F} = \sum_{l=0}^{k-1} \mathbf{H}_l$$

where

$$(\mathbf{H}_l)_{ij} = \begin{cases} (\mathbf{W}\otimes\mathbf{\Phi})^l, \ j = i - l \\ \mathbf{0}, \ \text{otherwise} \end{cases}.$$

For example, consider a 3-agent system with $k = 3$,

$$\mathbf{F} = \sum_{l=0}^{2} \mathbf{H}_l = \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}}_{\mathbf{H}_0} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{W} \otimes \mathbf{\Phi}) & \mathbf{0} \end{bmatrix}}_{\mathbf{H}_1} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi})^2 & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{H}_2}.$$

Consequently applying the permutation yields

$$(\mathbf{PF})_i = \sum_{l=0}^{2} (\mathbf{PH}_l)_i = \underbrace{\begin{bmatrix} \mathbf{J}_i & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{J}_i \end{bmatrix}}_{\mathbf{H}_0} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi})_i & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{W} \otimes \mathbf{\Phi})_i & \mathbf{0} \end{bmatrix}}_{\mathbf{H}_1} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W} \otimes \mathbf{\Phi})_i^2 & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\mathbf{H}_2}$$

for each $i \in \mathcal{V}$ maintains the lower triangular Toeplitz structure of $\mathbf{F}$ for each $(\mathbf{PM})_i$ block. We observe that the illustrative example in the proof of Corollary 8.1 is equivalent to $\mathbf{H}_0$, therefore

$$(\mathbf{PM})_i = (\mathbf{I} \otimes \mathbf{L}_i)\mathbf{O}_i + \begin{bmatrix} \mathbf{0}_{nNl \times nN(k-1)} \\ \sum_{l=1}^{k-1} (\mathbf{PH}_l)_i (\mathbf{I} \otimes \mathbf{L})\mathbf{O} \end{bmatrix}$$

It is apparent that $\mathbf{O}_i$ is the observability matrix of the measurements taken at Agent i and we use the following lemma to interpret the second term. Let us first consider $l = 1$,

which yields,

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W}\otimes\mathbf{\Phi})_i & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{W}\otimes\mathbf{\Phi})_i & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{LC} \\ \mathbf{LC\Phi} \\ \mathbf{LC\Phi}^2 \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ (\mathbf{W}\otimes\mathbf{\Phi})_i\mathbf{LC} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{W}\otimes\mathbf{\Phi})_i\mathbf{LC\Phi} & \mathbf{0} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0}_{nN\times nN^2} \\ \sum_{j\in\mathcal{N}_i}\mathbf{W}_{ij}\mathbf{\Phi}\mathbf{L}_j\mathbf{C}_j \\ \sum_{j\in\mathcal{N}_i}\mathbf{W}_{ij}\mathbf{\Phi}\mathbf{L}_j\mathbf{C}_j\mathbf{\Phi} \end{bmatrix} = \sum_{j\in\mathcal{N}_i}\begin{bmatrix} \mathbf{0}_{nN\times nN^2} \\ (\mathbf{I}\otimes\mathbf{W}_{ij}\mathbf{\Phi}\mathbf{L}_j)\bar{\mathbf{O}}_j \end{bmatrix}$$

with $\bar{\mathbf{O}}_j^\top = [\mathbf{C}_j^\top, \mathbf{\Phi}^\top\mathbf{C}_j^\top, \ldots, (\mathbf{\Phi}^\top)^{k-1-l}\mathbf{C}_j^\top]$. By induction, the result for each $l \in \{1, \ldots, k-1\}$ follows

$$\sum_{l=1}^{k-1}(\mathbf{PH}_l)_i(\mathbf{I}\otimes\mathbf{L})\mathbf{O} = \sum_{l=1}^{k-1}\sum_{j\in\mathcal{N}_i^l}\begin{bmatrix} \mathbf{0}_{lnN\times nN^2} \\ (\mathbf{I}\otimes(\mathbf{W}^l)_{ij}\mathbf{\Phi}^l\mathbf{L}_j)\bar{\mathbf{O}}_j \end{bmatrix}.$$

This presents a critical issue: there is an l-step delay for the information from each agent $j \in \mathcal{N}_i^l$ to reach agent i. It is a well-known result that the standard observability matrix $(\mathbf{O}_j^s)^\top = [\mathbf{C}^\top, \mathbf{C}^\top\mathbf{\Phi}^\top, \ldots, \mathbf{C}^\top(\mathbf{\Phi}^{nN-1})^\top]$ provides the necessary and sufficient amount of information to characterize the observability of the system from agent j. Consequently, k must be sufficiently large enough to allow for each block row of $\mathbf{O}_j^s$ for each $j \in \mathcal{V}$ to transverse the communication graph. Which is to say $\bar{\mathbf{O}}_j = \mathbf{O}_j^s$ which occurs when $k-1-l = nN-1$ or $k = nN+l$, where l is now the length of the shortest walk $j \to i$. The result follows by setting $l = d$ as the diameter of the communication graph representing the longest shortest path between two estimators. Because each walk follows a path from neighbor to neighbor $j \in \mathcal{N}_i^l \ \forall l \in \{0, \ldots, k-1\}$ indicates that $j \in \mathcal{W}_i$. □

4.4.2 The Omniscient Gramian

While the rank condition of the omniscient matrix $\mathbf{M}_i$ is necessary and sufficient for state omniscience, there is no direct tie Lyapunov function. In this section, we draw explicit connections between the Lyapunov function and the omniscient matrix through what we refer to as the omniscient gramian.

Theorem 10. *A distributed estimator is state omniscient if and only if* $\mathbf{M}_{ik} = \mathbf{M}_i^\top \mathbf{M}_i = \sum_{s=0}^{h} (\mathbf{\Phi}_{cl}^s)^\top \mathbf{C}^\top \mathbf{L}^\top (\mathbf{W}^\top \otimes \mathbf{\Phi}^\top)_i^{h-s} (\mathbf{W} \otimes \mathbf{\Phi})_i^{h-s} \mathbf{L}\mathbf{C}\mathbf{\Phi}_{cl}^s$ *is positive definite for each* $i \in \mathcal{V}$ *and* $\forall k \geq d$ *with* $\mathbf{\Phi}_{cl} := (\mathbf{W} \otimes \mathbf{\Phi} + \mathbf{L}\mathbf{C})$ *is the closed-loop system matrix, and where* $h = k - 1$*. Because the error dynamics are dependent on the closed-loop system, we use* $\mathbf{M}$ *from its definition in* (4.42)*.*

Proof. To prove sufficiency, we make the assumption that $\mathbf{\Phi}_{cl}^\top \mathbf{M}_{ik} \mathbf{\Phi}_{cl} - \mathbf{M}_{ik} = -\mathbf{\Psi}_k$ with $\mathbf{\Psi}_k = \mathbf{\Psi}_k^\top$ is positive definite. We now introduce the Lyapunov candidate function

$$V_i(k) = \mathbf{e}_i^\top(k) \mathbf{M}_{ik} \mathbf{e}_i(k)$$

such that

$$\Delta V_i(k) = \mathbf{e}_i^\top(k) [\mathbf{\Phi}_{cl}^\top \mathbf{M}_{ik} \mathbf{\Phi}_{cl} - \mathbf{M}_{ik}] \mathbf{e}_i(k) = -\mathbf{e}_i^\top(k) \mathbf{\Psi}_k \mathbf{e}_i(k) < 0$$

Thus, sufficiency is proven. To prove necessity, substitute $\mathbf{M}_{ik} = \sum_{s=0}^{k-1} (\mathbf{\Phi}_{cl}^s)^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^s$ into the Lyapunov candidate function such that

$$\mathbf{\Phi}_{cl}^\top \Big(\sum_{s=0}^{k-1} (\mathbf{\Phi}_{cl}^s)^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^s \Big) \mathbf{\Phi}_{cl} - \sum_{s=0}^{k-1} (\mathbf{\Phi}_{cl}^s)^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^s = -\mathbf{\Psi}_k$$

yielding

$$\sum_{s=0}^{k-1} \Big[(\mathbf{\Phi}_{cl}^{s+1})^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^{s+1} - (\mathbf{\Phi}_{cl}^s)^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^s \Big] = -\mathbf{\Psi}_k. \tag{4.46}$$

Solving (4.46) results in

$$(\mathbf{\Phi}_{cl}^k)^\top \mathbf{\Psi}_k \mathbf{\Phi}_{cl}^k - \mathbf{\Psi}_k = -\mathbf{\Psi}_k$$

and understanding that $(\mathbf{\Phi}_{cl}^k)\mathbf{\Psi}_k\mathbf{\Phi}_{cl}^k \to 0$ if and only if $\mathbf{\Phi}_{cl}$ is stable the result follows by setting $\mathbf{\Psi}_k = \sum_{s=0}^{h} \mathbf{C}^\top\mathbf{L}^\top(\mathbf{W}^\top \otimes \mathbf{\Phi}^\top)_i^{h-s}(\mathbf{W} \otimes \mathbf{\Phi})_i^{h-s}\mathbf{LC}$. □

By Theorem 10, $\mathbf{M}_i^\top\mathbf{M}_i$ solves the Lyapunov equation from Definition 7 when $\mathbf{M}_i$ is full rank.

4.4.3 An Alternative Characterization

We next introduce an alternative characterization for state omnscience akin to the PBH Lemma. The PBH Lemma [154] states that if a right-eigenvector $\mathbf{v}$ of $\mathbf{\Phi}$ with associated eigenvalue λ is orthogonal to $\mathbf{C}$, then that mode of $\mathbf{\Phi}$ is unobservable from $\mathbf{C}$, more specifically

$$\mathbf{\Phi v} = \lambda \mathbf{v},\ \mathbf{Cv} = 0. \tag{4.47}$$

We take the liberty to multiply each term in (4.47) by the walk matrix $\mathbf{W}$ such that

$$\mathbf{W\Phi v} = \mathbf{W}\lambda\mathbf{v},\ \mathbf{WCv} = \mathbf{0}$$

and introduce Theorem 11.

Theorem 11. *A distributed estimator is state omniscient if and only if* $\mathbf{W}_i\mathbf{LCv}_l \neq \mathbf{0}\ \forall\ i \in \mathcal{V}$ *and for each* $\mathbf{v}_l \in \{\mathbf{v}_1, \ldots, \mathbf{v}_{nN}\}$ *where* $\mathbf{W}_i$ *is the walk matrix with* $\mathcal{T} = i$.

Proof. Corollary 8.1 follows immediately from (4.47) as $\mathbf{W}_i\mathbf{LCv}_l = \mathbf{0}$ if $\mathbf{Cv}_l = \mathbf{0}$ indicating $\mathbf{v}_l$ is not observable anywhere in the system. Consequently, we assume that $\mathbf{Cv}_l \neq \mathbf{0}$ for the remainder of the proof. Under this assumption, we make the observation that for a 3 agent system

$$\mathbf{LCv}_l = \begin{bmatrix} \mathbf{L}_1\mathbf{C}_1\mathbf{v}_l \\ \mathbf{L}_2\mathbf{C}_2\mathbf{v}_l \\ \mathbf{L}_3\mathbf{C}_3\mathbf{v}_l \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{L}_2\mathbf{C}_2\mathbf{v}_l \\ \mathbf{0} \end{bmatrix}$$

if $\mathbf{v}_l \in \mathcal{O}_2$ but $\mathbf{v}_l \notin \mathcal{O}_1, \mathcal{O}_3$. By induction, this result extends to a N agent system. Note that while $\mathbf{C}\mathbf{v}_l \neq \mathbf{0}$, it is possible that $\mathbf{C}_i\mathbf{v}_l = \mathbf{0}$ for a given i. Next we must show that there exist at least a single walk from a nonzero partition of $\mathbf{C}\mathbf{v}_l$ to vertex i. Here, we define block generalizations of $\mathbf{s}$, $\mathbf{t}$, and $\mathbf{W}_i$ from (2.3) and (2.4) such that

$$\mathbf{s}_j = \begin{cases} \mathbf{L}_j\mathbf{C}_j\mathbf{v}_l, \ \mathbf{C}_j\mathbf{v}_l \neq \mathbf{0} \\ \mathbf{0}, \ \text{otherwise} \end{cases} \qquad \text{and } \mathbf{t}_i = \mathbf{J}_i^\top \tag{4.48}$$

for the target vertex i and

$$\mathbf{W}_i = \begin{bmatrix} \mathbf{J}_i \\ \mathbf{J}_i(\mathbf{W} \otimes \mathbf{\Phi}) \\ \vdots \\ \mathbf{J}_i(\mathbf{W} \otimes \mathbf{\Phi})^d \end{bmatrix}.$$

Therefore, $\mathbf{W}_i\mathbf{s} = \mathbf{0}$ implies that there does not exist a walk from a vertex j where $\mathbf{v}_l \in \mathcal{O}_j$ to vertex i. Conversely, if $\mathbf{W}_i\mathbf{s}_j \neq \mathbf{0}$ implies that there exists at least one walk from a vertex j where $\mathbf{v}_l \in \mathcal{O}_j$ to vertex i. □

Theorem 11 clearly defines the convergent subspace for each agent i. If $\mathbf{v}_l$ lies within the nullspace of $\mathbf{W}_i\mathbf{L}\mathbf{C}$ then $\mathbf{v}_l \notin \mathcal{C}_i$. It also provides an alternative characterization of state omniscience that resembles the PBH lemma. Moreover, this result naturally extends to asymmetric $\mathbf{\Phi}$ as long as there exists some $\mathbf{v}_l$ selected from the eigenspace of λ for which the result holds.

4.5 Considerations for Graph Topology Design

Now that a fundamental understanding of state omniscient system analysis has been laid, how does it relate to system dynamics, sensing, communication, and information fusion? This section translates the state omniscient system analysis to a set of heuristics for various design criteria.

The state omniscience and observability of the full networked system is consequently

dependent on four components - the system dynamics, the measurement function, the sensing graph, and the communication graph. This section focuses on the existence and type of potential solutions given design constraints on the networked system. Particularly, this section focuses on sensing capabilities and fusion types employed. More specifically, Table 4.1 details all of the potential combinations considered such as whether or not each agent is a capable of measuring the relative states of other agents i.e. no relative sensing (NRS) or relative sensing (RS); and specific types of fusion i.e. prediction fusion (P), output fusion (O), or both (P&O). This influences whether or not the associated matrices are a tunable parameter (D) or set as the identity matrix (**I**).

Table 4.1 Potential sensing and fusion cases.

Case #	Sensing	Fusion	W	U	$\mathbf{E}_S$
Case 1	NRS	N	**I**	**I**	**I**
Case 2	NRS	P	D	**I**	**I**
Case 3	NRS	O	**I**	D	**I**
Case 4	NRS	P&O	D	D	**I**
Case 5	RS	N	**I**	**I**	D
Case 6	RS	P	D	**I**	D
Case 7	RS	O	**I**	D	D
Case 8	RS	P&O	D	D	D

When a column of Table 4.1 is **I** indicates that the associated variable is the identity matrix and is fixed; contrarily, when a column is D indicates the associated variable is a design variable for the corresponding case. Prior to considering each case in depth, we make the following remark:

Remark 2. *Lemma 6.1 allows us to analyze each of the no predictor-fusion cases (*$\mathbf{W} = \mathbf{I}$*) from a ubiquitous observable perspective.*

4.5.1 Case 1: No Relative Sensing - No Fusion

It is trivial to show that the no relative sensing - no fusion scenario as described by Case 1 of Table 4.1 will never be observable. This result is intuitive as this case directly implies that each agent only has information about itself and does not share that information with

the other agents. While trivial and intuitive, we wish to formally prove the result as the process will also be repeated for the other scenarios in Cases 2-8 of Table 4.1.

We begin the proof by substituting the values for $\mathbf{W},\ \mathbf{U},\ \mathbf{E}_S$ of Case 1 into the error dynamics described by (4.11). It can be shown that the equation then takes the form

$$\mathbf{e}_{k+1} = \Big(\mathbf{I}_{N^2} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{I}_N(\mathbf{I}_N \otimes \mathbf{H})^\top(\mathbf{I}_N \otimes \mathbf{H})\Big)\mathbf{e}_k \tag{4.49}$$

with

$$(\mathbf{I}_N \otimes \mathbf{H})^\top = \begin{bmatrix} \mathbf{C}_1^\top & \dots & \dots & 0 \\ 0 & \mathbf{C}_2^\top & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{C}_N^\top \end{bmatrix} \in \mathbb{R}^{nN^2 \times pE} \tag{4.50}$$

where each $\mathbf{C}_i = \mathbf{i}_i \otimes \mathbf{H}^\top$ with $\mathbf{i}_i$ being the i^{th} column of $\mathbf{I}$. Using the PBH Lemma, we know that the dynamics of (4.49) is unobservable if and only if there exists some $\tilde{\mathbf{v}}_l$ such that

$$(\mathbf{I}_{N^2} \otimes \mathbf{\Phi})\tilde{\mathbf{v}}_l = \lambda_l \tilde{\mathbf{v}}_l,\ (\mathbf{I}_N \otimes \mathbf{H})^\top(\mathbf{I}_N \otimes \mathbf{H})\tilde{\mathbf{v}}_l = \mathbf{0}.$$

Therefore each $\mathcal{O}_i$ may be parameterized by all of the vectors $\tilde{\mathbf{v}}_l$ such that $\mathbf{C}_i\tilde{\mathbf{v}}_l \neq \mathbf{0}$. Because neither local state estimates nor measurements are shared $\mathcal{C}_i = \mathcal{O}_i$. Taking $\bar{\mathbf{Z}}$ to be the matrix of eigenvectors of $\mathbf{I}_{N^2} \otimes \mathbf{\Phi}$ and using properties of eigenvectors through Kronecker products yields $\bar{\mathbf{Z}} = \mathbf{I}_{N^2} \otimes \mathbf{Z}$ where $\mathbf{Z}$ is the matrix of eigenvectors of $\mathbf{\Phi}$. Since $\mathbf{\Phi}$ is itself diagonal matrix from (2.7), $\mathbf{Z} = \text{diag}([\mathbf{Z}_1, \dots, \mathbf{Z}_N])$ where $\mathbf{Z}_i$ is the matrix of eigenvectors for $\mathbf{\Phi}_i$ for each $i \in \mathcal{V}$, and since $\mathbf{C}_i$ only has the i^{th} diagonal block occupied yields a matrix form of the PBH-lemma $\mathbf{C}_\mathbf{A}\bar{\mathbf{Z}}$ where each i^{th} diagonal subblock takes the form $\mathbf{C}_\mathbf{A}^i\mathbf{Z} = (\mathbf{i}_i\mathbf{i}_i^\top \otimes \mathbf{H}^\top\mathbf{H})\mathbf{Z} = (\mathbf{i}_i\mathbf{i}_i^\top \otimes \mathbf{H}^\top\mathbf{H}\mathbf{Z}_i)$. Therefore, $\mathcal{C}_i = \mathcal{O}_i$ is only parameterized by the measurements agent i it takes of itself. Therefore, there exists no form of Case 1 such that the system will be observable from any agent let alone all agents.

4.5.2 Case 2: No Relative Sensing - Output Fusion

For the no relative sensing - output fusion case, we have the equation

$$\mathbf{e}_{k+1} = \Big((\mathbf{I}_{N^2} \otimes \mathbf{\Phi}) - \mathbf{K}\mathbf{U}(\mathbf{I}_N \otimes \mathbf{H})^\top (\mathbf{I}_N \otimes \mathbf{H})\Big)\mathbf{e}_k. \tag{4.51}$$

as each agent j does not communicate its state prediction with its neighbors which forces $\mathbf{W} = \mathbf{I}_N$. However, each agent j does communicate its output information to its neighbors. In this case, we are only interested in $\mathbf{C}_\mathbf{A}^i\mathbf{Z}$ and constructing a matrix $\mathbf{C}_\mathbf{A}$ such that the global system state as represented by $\mathbf{Z}$ may be observable from each $\mathbf{C}_\mathbf{A}^i$. The no relative sensing case with measurement sharing yields output matrices $\mathbf{C}_\mathbf{A}^i = \sum_{j\in\mathcal{N}_i} u_{ij}^2 \mathbf{C}_j^\top \mathbf{C}_j$ where each $\mathbf{C}_j^\top \mathbf{C}_j = \mathbf{i}_j\mathbf{i}_j^\top \otimes \mathbf{H}^\top\mathbf{H}$ so each agent is only taking measurements of itself, but then may share those measurements with its neighbors. Because there is no relative sensing, this mean there is no measurement overlap between agents, so output fusion only provides new state information and does not provide additional information to improve the observability of states currently measured from agent i. It follows directly, that the pair $(\mathbf{\Phi}_i, \mathbf{H})$ must be observable for each $i \in \mathcal{V}$. Consequently, the columns of $\mathbf{C}_\mathbf{A}^i$ associated with agent j are only nonzero if $j \in \mathcal{N}_i$ agent j is a neighbor of agent i in the graph. It follows that each agent has an observable form of the system - the system is ubiquitously observable - if each agent is a neighbor of every other agent, or the graph is fully connected. Therefore, in order for the no relative sensing - output fusion case to yield a UO system, each $(\mathbf{\Phi}_i, \mathbf{H})$ must be observable and $\mathcal{G}_c$ must be fully connected.

4.5.3 Case 3: No Relative Sensing - Predictor Fusion

Similarly, we begin analyzing the no relative sensing with predictor fusion scenario from Case 2 of Table 4.1 by substituting the values of $\mathbf{W}$, $\mathbf{U}$, and $\mathbf{E}_S$ into (4.11) yielding the equation

$$\mathbf{e}_{k+1} = \Big(\mathbf{W} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{I}_N(\mathbf{I}_N \otimes \mathbf{H})^\top (\mathbf{I}_N \otimes \mathbf{H})\Big)\mathbf{e}_k. \tag{4.52}$$

Because there is no relative sensing and no output fusion, $\mathbf{C}_\mathbf{A}^i = \mathbf{\Pi}\mathbf{C}_j^\top \mathbf{C}_j$ so $\mathcal{O}_i$ is only

dependent on the measurements of agent i. Similarly to the no relative sensing - output fusion case, this means each $(\mathbf{\Phi}_i, \mathbf{H})$ must be observable. Because $\mathcal{O}_i$ only contains information about agent i, $\mathcal{C}_i$ from (4.34) must provide the information about every other agent j, therefore, $j \in \mathcal{W}_i$ which for state omniscience must be true for each $i \in \mathcal{V}$. It follows that there must exist a walk from any i to any j so $\mathcal{G}_c$ must be strongly connected. Therefore, in order for the no relative sensing - predictor fusion case, to be state omniscient, each pair $(\mathbf{\Phi}_i, \mathbf{H})$ must be observable and $\mathcal{G}_c$ must be strongly connected.

4.5.4 Case 4: No Relative Sensing - Predictor & Output Fusion

Case 4 of Table 4.1 is the first to consider the interplay between predictor and output fusion. Substituting the values for the no relative sensing - predictor and output fusion in Table 4.1 yields the equation

$$\mathbf{e}_{k+1} = \Big((\mathbf{W} \otimes \mathbf{\Phi}) - \mathbf{K}\mathbf{U}(\mathbf{I}_N \otimes \mathbf{H})^\top (\mathbf{I}_N \otimes \mathbf{H})\Big)\mathbf{e}_k. \tag{4.53}$$

Because there is no relative sensing, this mean there is no measurement overlap between agents, so output fusion only provides new state information and does not provide additional information to improve the observability of states currently measured from agent i. The no relative sensing case with measurement sharing yields output matrices $\mathbf{C}_{\mathbf{A}}^i = \sum_{j \in \mathcal{N}_i} u_{ij}^2 \mathbf{C}_j^\top \mathbf{C}_j$ where each $\mathbf{C}_j^\top \mathbf{C}_j = \mathbf{i}_j \mathbf{i}_j^\top \otimes \mathbf{H}^\top \mathbf{H}$ so each agent is only taking measurements of itself, but then may share those measurements with its neighbors. In this case both the local state estimate and measurements are being shared amongst neighbors, the measurement information is redundant. Therefore, the no relative relative sensing - predictor and output fusion case performs similarly to the no relative sensing - predictor fusion case which states each pair $(\mathbf{\Phi}_i, \mathbf{H})$ must be observable and $\mathcal{G}_c$ must be strongly connected.

4.5.5 Case 5: Relative Sensing - No Fusion

Case 5 of Table 4.1 is the first case to consider relative sensing meaning $\mathbf{E}_S \neq \mathbf{I}_N$ and each agent is capable of measuring the relative states of other states to itself. Considering this case yields the equation

$$\mathbf{e}_{k+1} = \Big((\mathbf{I}_{N^2} \otimes \mathbf{\Phi}) - \mathbf{K}\mathbf{I}_N(\mathbf{E}_s^\top \otimes \mathbf{H})^\top(\mathbf{E}_s^\top \otimes \mathbf{H})\Big)\mathbf{e}_k. \tag{4.54}$$

Because there is no output fusion, $\mathbf{C}_{\mathbf{A}}^i = \mathbf{C}_i^\top \mathbf{C}_i$. The goal remains to ensure each column of $\mathbf{C}_{\mathbf{A}}^i \mathbf{Z} \neq \mathbf{0}$. This formulation is reminiscient of the results from Section 4.5.2, however the sensing capabilities of agent i through $\mathbf{C}_i^\top \mathbf{C}_i$ takes the full burden of forming an observable system. Consequently, the columns of $\mathbf{C}_{\mathbf{A}}^i$ associated with agent j are only nonzero if $j \in \mathcal{N}_i^S$ agent j is a neighbor of agent i in $\mathcal{G}_s$ agent i is taking a measurement of agent j. It follows that the system is ubiquitously observable if each agent is a neighbor of every other agent, or $\mathcal{G}_s$ is fully connected. It is also necessary that the pair $(\mathbf{\Phi}_i, \mathbf{H})$ is observable for each $i \in \mathcal{V}$.

4.5.6 Case 6: Relative Sensing - Predictor Fusion

Case 6 of Table 4.1 is the relative sensing - predictor fusion case which yields the equation

$$\mathbf{e}_{k+1} = \Big(\mathbf{W} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{I}_N(\mathbf{E}_s^\top \otimes \mathbf{H})^\top(\mathbf{E}_s^\top \otimes \mathbf{H})\Big)\mathbf{e}_k. \tag{4.55}$$

Using a similar argument to the one presented in Section 4.5.3, one may naïvely say a sufficient condition is to ensure the $\mathcal{G}_s \cup \mathcal{G}_c$ is strongly connected for any structures of $\mathcal{G}_s$ and $\mathcal{G}_c$. It is here that the problem discussed in Section 2.3 becomes apparent. In order to maintain a continuous flow of information from one agent to the next, each sensing link must be followed by a communication link ultimately placing a constraint on the structure of $\mathcal{G}_c$ given $\mathcal{G}_s$. This indicates that $\mathcal{G}_c \supseteq \partial_e^{in}\mathcal{G}_s$ where ∂_e^{in} is taken to mean the in-edge boundary of $\mathcal{G}_s$. We consider the in-edge boundary of $\mathcal{G}_s$ to be the set of edges exiting a vertex in $\mathcal{G}_s$. Therefore, the edges in $\mathcal{G}_c$ must at least be a superset of the in-edge boundary

on $\mathcal{G}_s$. Consequently, it is true that the relative sensing - predictor fusion case requires that $\mathcal{G}_s \cup \mathcal{G}_c$ be strongly connected given that $\mathcal{G}_c \supseteq \partial_e^{in}\mathcal{G}_s$; because there is no output fusion, the pair $(\mathbf{\Phi}_j, \mathbf{H})$ must be observable for each $i \in \mathcal{V}$ where $j \in \mathcal{N}_i^s$.

4.5.7 Case 7: Relative Sensing - Output Fusion

Next, we consider Case 7 of Table 4.1. Because no prediction information is shared across the network, yields the pair

$$\mathbf{e}_{k+1} = \Big((\mathbf{I}_{N^2} \otimes \mathbf{\Phi}) - \mathbf{K}\mathbf{U}(\mathbf{E}_s^\top \otimes \mathbf{H})^\top(\mathbf{E}_s^\top \otimes \mathbf{H})\Big)\mathbf{e}_k, \tag{4.56}$$

so we are only concerned with each $\mathbf{C}_{\mathbf{A}}^i\mathbf{Z}$. Unlike the no relative sensing - output fusion case each $\mathbf{C}_j^\top\mathbf{C}_j = {}^j\mathbf{E}_S\,{}^j\mathbf{E}_S^\top \otimes \mathbf{H}^\top\mathbf{H}$ indicates agent j is not only capable of taking measurements of itself, but also its neighbors in $\mathcal{G}_s$. Furthermore, Case 3 in Section 4.5.2 and Case 5 in Section 4.5.5 are specific cases of the relative sensing - output fusion case, with either $\mathbf{E}_S$ or $\mathbf{U}$ set to $\mathbf{I}_N$, each requiring a fully connected graph to yield an ubiquitously observable system. It follows that the relative sensing - output fusion cases also requires a fully connected network, yet only necessitates the $\mathcal{G}_s \cup \mathcal{G}_c$ is fully connected. The communication graph supplements the measurements missed by the sensing graph. Contrary to each of the previous cases, the pair $(\mathbf{\Phi}_j, \mathbf{H})$ does not necessarily need to be observable, but the pair $(\mathbf{\Phi}, \mathbf{C}_{\mathbf{A}}^i)$ must be observable.

4.5.8 Case 8: Relative Sensing - Predictor & Output Fusion

Finally, we consider Case 8 of Table 4.1, the relative sensing - predictor & output fusion case. This is the most general, and most complex case that may be constructed. Similarly to all of the previous cases, our goal is to define potential constructions of $\mathbf{W},\ \mathbf{U},\ \mathbf{E}_S$ such that the equation

$$\mathbf{e}_{k+1} = \Big(\mathbf{W} \otimes \mathbf{\Phi} - \mathbf{K}\mathbf{U}(\mathbf{E}_s^\top \otimes \mathbf{H})^\top(\mathbf{E}_s^\top \otimes \mathbf{H})\Big)\mathbf{e}_k \tag{4.57}$$

is state omniscient. Contrarily to the no relative sensing - predictor & output fusion case in Section 4.5.4. There exists relative measurements for the output fusion scheme to merge that may improve the overall observability at each agent. Therefore, it is not necessary for each pair $(\mathbf{\Phi}_j, \mathbf{H})$ to be observable, rather the augmented measurement form $(\mathbf{\Phi}, \mathbf{C_A})$ should be centrally observable. Notice, because of the inclusion of predictor fusion, each pair $(\mathbf{\Phi}_j, \mathbf{C}^i_\mathbf{A})$ need not be observable either; each $\mathcal{O}_i$ may be complemented by the converging information provided by $\mathcal{N}_i$ through predictor fusion. Because the shared prediction and output values utilize the same communication links the overall structure of the graphs take the form of relative sensing - predictor fusion which is to say $\mathcal{G}_s \cup \mathcal{G}_c$ must be strongly connected given that $\mathcal{G}_c \supseteq \partial^{in}_e \mathcal{G}_s$. If it is the case, that output fusion and predictor fusion utilize separate communication graphs, $\mathcal{G}^p_c$ and $\mathcal{G}^o_c$ for predictor and output fusion, respectively. Using the same argument from Section 4.5.6 it can be shown that $\mathcal{G}_s \cup \mathcal{G}_c$ must be strongly connected given that $\mathcal{G}_c := \mathcal{G}^p_c \cup \mathcal{G}^o_c$ and $\mathcal{G}_c \supseteq \partial^{in}_e \mathcal{G}_s \cup \partial^{in}_e \mathcal{G}^o_c$. Therefore, the relative sensing - predictor & output fusion case may be rendered state omniscient when the pair $(\mathbf{\Phi}, \mathbf{C_A})$ is centrally observable and $\mathcal{G}_s \cup \mathcal{G}_c$ is strongly connected given the conditions specified.

4.6 Illustrative Example

To illustrate the state omniscience property discussed in Section 4.4, let us revisit the motivating example from Section 4.1 and consider three distinct cases. In each case, we consider two estimators each tasked with tracking an LTI system, but with slight variations in the measurements and communicated information. In our example we consider the standard Clohessy-Wiltshire dynamics [128] from relative orbital mechanics 2.17 with $\eta = 0.001027$.

4.6.1 Case 1

In the first case, we model the motivating example exactly. Estimator 1 is taking direct measurements of the full LTI system state $\mathbf{C}_1 = \mathbf{I}$, but Estimator 2 is not taking any measurements of the system and thus is unable to estimate the system state $\mathbf{C}_2 = \mathbf{0}$.

Rather, we allow Estimator 1 to communicate its local state estimate to Estimator 2 such that the local state estimator at Estimator 2 takes the form of (4.13). In Figure 4.4, we see that each estimator converges to the true value with a slight delay in the convergence of Estimator 2 that is proportional to the state update that maps the local state estimate of Estimator 1 to the local state estimate of Estimator 2.

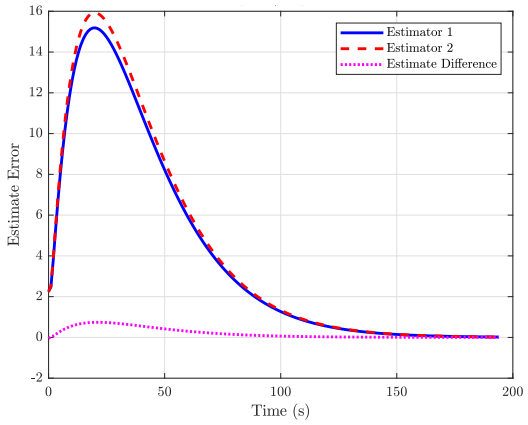


Figure 4.4 Motivating example results with $\mathbf{C}_1 = \mathbf{I}$ and $\mathbf{C}_2 = \mathbf{0}$ and Estimator 1 can communicate to Estimator 2.

4.6.2 Case 2

Next, we consider a modified case of the motivating example such that each estimator may only measure an unobservable portion of the system. We allow Estimator 1 to measure the first and second states along with their rates of change; we allow Estimator 2 to measure the third state along with its rate of change - the linear measurement matrices are given by

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Neither are observable on their own - Figure 4.5 illustrates this point because it does not allow communication between the estimator, so neither are able to converge to the true value.

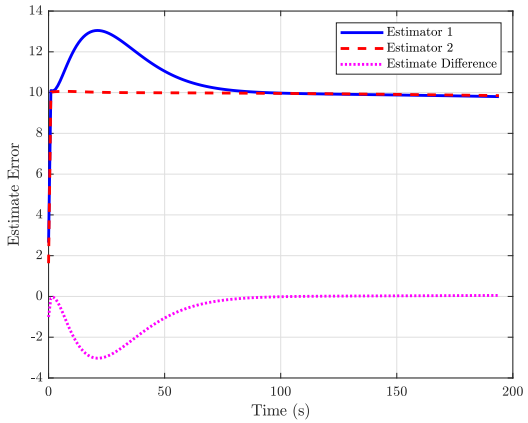


Figure 4.5 Motivating example with neither estimator able to fully observe the system and unable to communicate with each other.

4.6.3 Case 3

Contrarily to Case 2, in Case 3, we allow communication between the estimators such that

$$\mathbf{W}_{11} = \mathbf{W}_{12} = \mathbf{W}_{21} = \mathbf{W}_{22} = \frac{1}{2}\mathbf{I}.$$

Consequently, as seen in Figure 4.6, each estimator is able to converge because of the shared information.

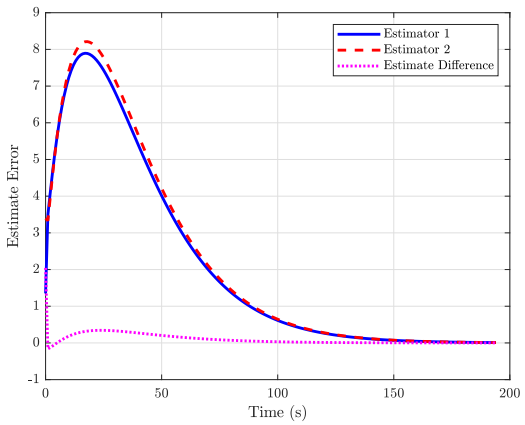


Figure 4.6 Motivating example with neither estimator able to fully observe the system and they are able to communicate with each other.

5 Information Fusion via Matrix Decomposition

Information flow thorughout the network creates an influx of information at each agent. Each agent must then be capable of fusing the incoming information to yield an updated state estimate; and it must do it in a way that maintains confidence levels without becoming over or under confident in its estimate.

5.1 Fused Matrix Bounds

Recalling the joint matrix bounds from Equations 2.35, it can be shown that the joint bounds are analogous to bounds on the fused solution of the form of Equation 2.32. The joint bounds of Equation 2.35 and joint covariance of Equation 2.33 with its delimiters may be substituted into Equation 2.32 and then into $\mathbf{\Pi}_u - \mathbf{P}_j \geq 0$:

$$\begin{bmatrix} k_a \mu \mathbf{P}_a & -\sqrt{k_a k_b}\mathbf{L}_a \mathbf{\Omega} \mathbf{L}_b^\top \\ -\sqrt{k_a k_b}\mathbf{L}_b \mathbf{\Omega}^\top \mathbf{L}_a^\top & \frac{k_b}{\mu}\mathbf{P}_b \end{bmatrix} \geq 0 \tag{5.1}$$

The left-hand side of which may be expanded into

$$\begin{bmatrix} \sqrt{\mu}\mathbf{A} & 0 \\ 0 & \frac{1}{\sqrt{\mu}}\mathbf{B} \end{bmatrix} \begin{bmatrix} k_a \mathbf{I} & -\sqrt{k_a k_b}\mathbf{\Omega} \\ -\sqrt{k_a k_b}\mathbf{\Omega}^\top & k_b \mathbf{I} \end{bmatrix} \begin{bmatrix} \sqrt{\mu}\mathbf{A} & 0 \\ 0 & \frac{1}{\sqrt{\mu}}\mathbf{B} \end{bmatrix}^\top \tag{5.2}$$

Equation 5.2 holds true when the middle term is positive semi-definite which requires $\mathbf{I} \geq \mathbf{\Omega}\mathbf{\Omega}^\top$ and is independent on the exact fraction of the square roots. Next, we consider partial explicit knowledge of the correlation of the form in Equation 5.3

$$\mathbf{P}_{ab} \in \{\mathbf{M} + \mathbf{A}\mathbf{\Omega}\mathbf{B}^\top \mid \begin{bmatrix} \rho\mathbf{P}_a & \mathbf{P}_{ab} \\ \mathbf{P}_{ab}^\top & \rho\mathbf{P}_b \end{bmatrix} \geq 0\} \tag{5.3}$$

where $0 \leq \rho \leq 1$ is the known explicit knowledge correlation. This is a generalization of the condition given in Equation 2.34 to explicit partial knowledge correlations [112]. It can be shown via the process above that under this partial explicit correlation knowledge that

$\rho^2\mathbf{I} \geq \mathbf{\Omega}\mathbf{\Omega}^\top$ must be true to maintain the postive semi-definite condition. The joint matrix bounds expressed in Equation 5.2 can be shown to be Equation 5.4 in its fused form.

$$\mu\mathbf{A}\mathbf{A}^\top + \frac{1}{\mu}\mathbf{B}\mathbf{B}^\top \leq (\mathbf{A}\mathbf{\Omega}\mathbf{B}^\top + \mathbf{B}\mathbf{\Omega}\mathbf{A}^\top) \tag{5.4}$$

It is assumed that $\mathbf{\Omega}$, the right-hand side of Equation 5.4 is symmetric, so by cancelling out applicable terms and solving for $\mathbf{\Omega}$ yields:

$$\mathbf{\Omega} \geq \frac{1}{2}\mathbf{L}_a^{-1}(\mu\mathbf{P}_a + \frac{1}{\mu}\mathbf{P}_b)\mathbf{L}_b^{-T} \tag{5.5}$$

Therefore, an upper-bound exists of the form presented by Equation 5.5 for any $0 \leq \mu < \infty$.

Similarly, an analogous lower-bound may be found using $\mathbf{P}_j - \mathbf{\Pi}_l \geq 0$:

$$\mathbf{\Omega} \leq \frac{1}{2}\mathbf{L}_a^{-1}(\nu\mathbf{P}_a + \frac{1}{\nu}\mathbf{P}_b)\mathbf{L}_b^{-T} \tag{5.6}$$

for $0 \leq \nu < \infty$. Notice that each of the presented bounds are independent of the fraction k of the square root and only dependent on the square root of $\mathbf{P}$ itself.

5.2 Matrix Decomposition Approaches

This section derives approximations to Equation 2.32 using the bound tool presented in Section 5.1. In particular, two square root decomposition approaches are presented to approximate $(\mathbf{P}_{ab} + \mathbf{P}_{ba})$, and an approximation for the mutual mean $\boldsymbol{\gamma}$ is derived following an analogous derivation to Ellipsoidal Intersection [108].

5.2.1 Cholesky Decomposition

Because a general covariance matrix is positive semi-definite, and in many cases positive definite, the Cholesky decomposition may be used to factor $\mathbf{P}$ into a single lower triangular matrix $\mathbf{L}$ with real and positive values along the diagonal.

$$\mathbf{P} = \mathbf{L}\mathbf{L}^\top \tag{5.7}$$

Using this fact within Equation 2.32, both matrices $\mathbf{P}_a$ and $\mathbf{P}_b$ may be factored into $\mathbf{L}_a$ and $\mathbf{L}_b$. As a result, Equation 2.32 can be rewritten in a factored form as,

$$\mathbf{P}_c^{-1} = (\mathbf{L}_a\mathbf{L}_a^\top)^{-1} + (\mathbf{L}_b\mathbf{L}_b^\top)^{-1} - (\mathbf{P}_{ab} + \mathbf{P}_{ba})^{-1} \tag{5.8}$$

Equation 5.8 effectively provides an analytical means of optimally fusing measurements and uncertainties between two sensors. This formulation assumes a known correlation between **a** and **b**. Which, in most cases, is neither known nor measurable. If it was known, Equation 5.8 would be extraneous. However, this equation become relevant if there exists an approximation to $\mathbf{P}_{ab} + \mathbf{L}_{ba}\mathbf{L}_{ba}^\top$ that can be constructed from known information, and the approximation maintains consistency.

Using Equation 2.34, consider the candidate approximation,

$$\mathbf{P}_{ab} \approx \mathbf{M} + \mathbf{L}_a\boldsymbol{\Omega}\mathbf{L}_b^\top \tag{5.9}$$

such that Equation 5.8 is approximated as

$$\mathbf{P}_c^{-1} \approx (\mathbf{L}_a\mathbf{L}_a^\top)^{-1} + (\mathbf{L}_b\mathbf{L}_b^\top)^{-1} - (\mathbf{M} + \mathbf{L}_a\boldsymbol{\Omega}\mathbf{L}_b^\top + \mathbf{L}_b\boldsymbol{\Omega}^\top\mathbf{L}_a^\top + \mathbf{M}^\top)^{-1} \tag{5.10}$$

Theorem 12. *Equation 5.10 provides a consistent and tight approximation of the optimal fusion solution.*

Proof. Prior to proving Theorem 12 directly, we introduce Corollary 12.1 that allows for the elimination of $\mathbf{M}$

Lemma 12.1. $\boldsymbol{M} = 0$ *provides the ultimate upper-bound for the family of covariance matrices with given* μ *and admissible values of* $\boldsymbol{\Omega}$.

Proof. The matrix $\mathbf{M} = 0$ provides an upper-bound for this family of covariances matrices

if the inequality in Equation 5.11 holds true.

$$\mathbf{P}_c^{\mathbf{0}} - \mathbf{P}_c^{\mathbf{M}} \geq 0 \tag{5.11}$$

where the superscript $\mathbf{0}$ indicates $\mathbf{M} = \mathbf{0}$. Substituting the reciprocal of Equation 5.10 into Equation 5.11 and cancelling out applicable terms yields

$$-(\mathbf{L}_a\boldsymbol{\Omega}\mathbf{L}_b^\top + \mathbf{L}_b\boldsymbol{\Omega}^\top\mathbf{L}_a^\top)^{-1} + (\mathbf{M} + \mathbf{L}_a\boldsymbol{\Omega}\mathbf{L}_b^\top + \mathbf{L}_b\boldsymbol{\Omega}^\top\mathbf{L}_a^\top + \mathbf{M}^\top)^{-1} \leq 0$$

because both $\mathbf{P}_c^{\mathbf{0}}$ and $\mathbf{P}_c^{\mathbf{M}}$ are positive semi-definite the reciprocal flips the inequality. Taking the reciprocal and cancelling out applicable terms yields:

$$\mathbf{M} + \mathbf{M}^\top \geq 0 \tag{5.12}$$

Because it is assumed that $\mathbf{M}$ represents any known elements of $\mathbf{P}_{ab}$ it follows that $\mathbf{M}$ is positive definite thus the above inequality holds and proves Lemma 12.1. □

This result implies that $\mathbf{M} = 0$ is the ultimate upper-bound in the family of upper-bounds for a given μ and $\boldsymbol{\Omega}$. Consequently, any partial knowledge of the state uncertainty only aids in decreasing this upper-bound. Fortunately, this leads to the result that $\mathbf{M} = 0$ provides an upper-limit when the exact structure of the cross-correlation terms may be unknown, as is the case in Equation 5.10.

This is particularly useful for Equation 5.10 because it limits the number of required known elements and allows for a singular tuning parameter in $\boldsymbol{\Omega} \in \mathbb{R}^{n\times n}$. Proving this limit maintains *consistency* is a direct result of Lemma 12.2.

Lemma 12.2. *Equation 5.10 is consistent because* $\boldsymbol{P}_{ab} \geq \boldsymbol{L}_a\boldsymbol{\Omega}\boldsymbol{L}_b^\top$*, where* $\boldsymbol{\Omega} = \rho\boldsymbol{I}$ *is the delimiting case.*

Proof. By the definition of a consistency, a fused estimate provides an upper-bound for the

optimal solution:

$$\mathbf{P}_c - \mathbf{P}_c^o \geq 0$$

where $(\cdot)^o$ denotes the optimal solution. It is trivial to see that by substituting Equation 5.8 and Equation 5.10 into the definition of consistency and using the fact that $\mathbf{M} = 0$ provides an ultimate upper-bound, yields:

$$\mathbf{P}_{ab} \geq \mathbf{L}_a \mathbf{\Omega} \mathbf{L}_b^\top$$

Now, recall the matrix Cauchy-Schwarz Equation with partial knowledge correlation

$$\rho \mathbf{P}_a \geq \mathbf{P}_{ab} (\rho \mathbf{P}_b)^{-1} \mathbf{P}_{ab}^\top \tag{5.13}$$

This is reminiscent of the explicit correlation knowledge condition from Equation 5.3, so by substituting the $\mathbf{P}_{ab}$ approximation from Equation 5.9, $\mathbf{M} = 0$, and performing some simple algebraic manipulation yields

$$\rho^2 \mathbf{I} \geq \mathbf{\Omega} \mathbf{\Omega}^\top \tag{5.14}$$

Where $\rho \mathbf{I}$ is an upper-bound in the solution space and thus proves the Lemma. □

While the optimal solution provides a lower bound of the covariance matrix, it remains possible for the true covariances to be significantly over-approximated. As a solution, an upper-bound to the known covariance matrix is considered feasible if it maintains consistency and tightness.

Lemma 12.3. *Equation 5.10 is tight because the ellipse formed by* $\mathbf{\Gamma}^{-1} \leq (\boldsymbol{L}_a \mathbf{\Omega} \boldsymbol{L}_b^\top + \boldsymbol{L}_b \mathbf{\Omega}^\top \boldsymbol{L}_a^\top)^{-1} \leq (\omega \boldsymbol{P}_a + (1-\omega) \boldsymbol{P}_b)^{-1}$ *using the Cholesky Decomposition is upper-bounded by the ellipse formed using ICI.*

Proof. From Definition 6, [162] shows that this concept may be parameterized in the form

$$\mathbf{\Lambda} = \alpha_1 \mathbf{P}_a + \alpha_2 \mathbf{P}_b$$

where $\sum \alpha_i = 1$. Noack in [109] uses this form to prove *tightness* of ICI when $\omega \in [0,\ 1]$.

$$\mathbf{\Gamma}^{-1} \leq (\omega \mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1} \tag{5.15}$$

Inherently, the Cholesky Decomposition approach is not in a form to directly evaluate *tightness* from Equation 5.15. Rather, it is possible to indirectly prove *tightness* by showing Lemma 12.3 holds true.

$$\mathbf{P}_{ICI} - \mathbf{P}_{chol} \geq 0 \tag{5.16}$$

substituting Equation 2.39 and Equation 5.10 with $\mathbf{M} = 0$ into Equation 5.16 yields,

$$[\mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\omega \mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1}]^{-1} \geq [\mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{L}_a \mathbf{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \mathbf{\Omega}^\top \mathbf{L}_a)^{-1}]^{-1}$$

cancelling the applicable terms,

$$(\omega \mathbf{P}_a + (1-\omega)\mathbf{P}_b) \leq (\mathbf{L}_a \mathbf{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \mathbf{\Omega}^\top \mathbf{L}_a^\top)$$

which can also be written on the information space as,

$$(\omega \mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1} \geq (\mathbf{L}_a \mathbf{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \mathbf{\Omega}^\top \mathbf{L}_a^\top)^{-1} \tag{5.17}$$

Thus Lemma 12.3 is proven. □

In this form, the above equation has no analytical solution. However, the equation may be solved numerically to find a solution for $\mathbf{\Omega}$. Because only admissible values for $\mathbf{\Omega}$ are

assumed, the term $\mathbf{L}_a\mathbf{\Omega}\mathbf{L}_b^\top$ is symmetric and may therefore be rewritten as,

$$(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b) \leq 2(\mathbf{L}_a\mathbf{\Omega}\mathbf{L}_b^\top)$$

Solving for $\mathbf{\Omega}$ results in an upper-bound on the Cholesky Decomposition solution being the ICI solution.

$$\mathbf{\Omega} \geq \frac{1}{2}\mathbf{L}_a^{-1}(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)\mathbf{L}_b^{-T} \tag{5.18}$$

Because the ICI solution has previously been proven to uphold Definition 6, the Cholesky Decomposition approach is implied to uphold *tightness* because the ICI solution serves as the upper-bound of the solution given $\mathbf{\Omega}$ is selected to maintain the inequality in Equation 5.18.

The bound specified by Equation 5.18 may be generalized to include the explicit partial correlation knowledge as

$$\mathbf{\Omega} \geq \frac{1}{2}\mathbf{L}_a^{-1}(\omega_a\mathbf{P}_a + \omega_b\mathbf{P}_b)\mathbf{L}_b^{-T} \tag{5.19}$$

where ω_a and ω_b are defined in Equation 5.20.

$$\begin{aligned} \omega_a &= \omega(\omega + \rho(1-\omega))^{-1} \\ \omega_b &= (1-\omega)(\rho\omega + (1-\omega))^{-1} \end{aligned} \tag{5.20}$$

Furthermore, because Lemma 12.2 and Lemma 12.3 are proven to be true, Theorem 12 is also upheld. □

Next, we discuss how to calculate the fused mean by implementing the mutual information approximation using the mutual mean $\boldsymbol{\gamma}$. The mutual mean represents a unique balance between the random variable **a** and **b** and can be considered "optimal" in

the sense it minimizes a weighted average cost function between the two random variables.

$$J(\boldsymbol{\gamma}) = (\boldsymbol{\gamma} - \mathbf{a})^\top \mathbf{W}_a (\boldsymbol{\gamma} - \mathbf{a}) + (\boldsymbol{\gamma} - \mathbf{b})^\top \mathbf{W}_b (\boldsymbol{\gamma} - \mathbf{b})$$

which yields the gradient

$$\frac{\partial J(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = 2(\boldsymbol{\gamma} - \mathbf{a})^\top \mathbf{W}_a + 2(\boldsymbol{\gamma} - \mathbf{b})^\top \mathbf{W}_b$$

with the stationary points

$$\boldsymbol{\gamma} = (\mathbf{W}_a + \mathbf{W}_b)^{-1}(\mathbf{W}_a \mathbf{a} + \mathbf{W}_b \mathbf{b})$$

Now, let us define $\mathbf{W}_a = \mathbf{P}_a^{-1} - \frac{1}{2}(\mathbf{L}_a \boldsymbol{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \boldsymbol{\Omega} \mathbf{L}_a^\top)^{-1}$ and $\mathbf{W}_b = \mathbf{P}_b^{-1} - \frac{1}{2}(\mathbf{L}_a \boldsymbol{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \boldsymbol{\Omega} \mathbf{L}_a^\top)^{-1}$. Substituting these definitions for $\mathbf{W}_a$ and $\mathbf{W}_b$ into the stationary points equations yields

$$\boldsymbol{\gamma} = \mathbf{P}_c(\mathbf{W}_a \mathbf{a} + \mathbf{W}_b \mathbf{b}) \tag{5.21}$$

In summary, the Cholesky Decomposition Approach is formalized in Equation 5.22,

$$\begin{aligned} \mathbf{P}_c^{-1} &= \mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{L}_a \boldsymbol{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \boldsymbol{\Omega}^\top \mathbf{L}_a^\top)^{-1} \\ \bar{\mathbf{c}} &= \mathbf{P}_c\Big(\mathbf{P}_a^{-1}\mathbf{a} + \mathbf{P}_b^{-1}\mathbf{b} - (\mathbf{L}_a \boldsymbol{\Omega} \mathbf{L}_b^\top + \mathbf{L}_b \boldsymbol{\Omega}^\top \mathbf{L}_a^\top)^{-1}\boldsymbol{\gamma}\Big) \end{aligned} \tag{5.22}$$

provides a consistent and tight approximation of the optimal solution when $\boldsymbol{\Omega}$ meets the criteria

$$\frac{1}{2}\mathbf{L}_a^{-1}(\omega_a \mathbf{P}_a + \omega_b \mathbf{P}_b)\mathbf{L}_b^{-T} \leq \boldsymbol{\Omega} \leq \rho \mathbf{I}$$

where ω_a and ω_b are defined in Equation 5.20.

A sample Cholesky Decomposition fusion result is found in Figure 5.1. A selection of bounds set between the identity and Equation 5.18. Graphically, tighter approximations of

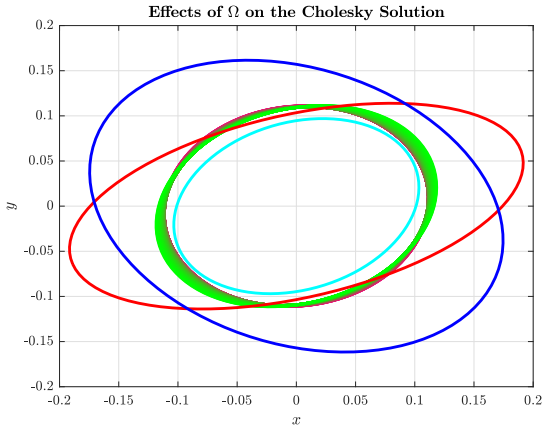


Figure 5.1 Demonstration of how $\mathbf{\Omega}$ influences the Cholesky Approximation solution by fusing Covariance Ellipse 1 with Covariance Ellipse 2. The **ICI Solution** is used as an upper-bound on the solution and the Optimal Solution provides the lower-bound. Results are presented for bounds on $\mathbf{\Omega}$ as Lower-Bound from Equation 5.18 and the Upper-Bound is the **Identity Matrix** because we consider $\rho = 1$ in this example.

the optimal solution are achieved when compared to both the CI and ICI solutions. Because of the dependence of the upper-bound on the ICI solution, the values for $\mathbf{\Omega}$ can be designed to always be tighter than the ICI solution.

5.2.2 Schur Decomposition

Alternatively to the Cholesky Decomposition defined in Equation 5.7, the cross-correlation terms may be approximated using a matrix square root found via the Schur Decomposition ([163, 164]).

$$(\mathbf{Q}_{ab})^2 \approx \mathbf{M} + \mathbf{Q}_a \mathbf{\Omega} \mathbf{Q}_b \tag{5.23}$$

where $\mathbf{Q} = \mathbf{P}^{1/2}$. This implies a fundamental change to Equation 5.8 where each term is decomposed into a square root rather than its Cholesky Decomposition.

$$\mathbf{P}_c^{-1} = \mathbf{Q}_a^{-2} + \mathbf{Q}_b^{-2} - (\mathbf{Q}_{ab}^2 + \mathbf{Q}_{ba}^2)^{-1} \tag{5.24}$$

Substituting Equation 5.23 into Equation 5.24 yields

$$\mathbf{P}_c^{-1} \approx \mathbf{Q}_a^{-2} + \mathbf{Q}_b^{-2} - (\mathbf{M} + \mathbf{Q}_a \boldsymbol{\Omega} \mathbf{Q}_b + \mathbf{M}^\top + \mathbf{Q}_b^\top \boldsymbol{\Omega}^\top \mathbf{Q}_a^\top)^{-1} \tag{5.25}$$

which is an analogous form of Equation 5.7 (*consistency* and *tightness* may be proven through analogous forms of Theorem 12, Lemma 12.2, and Lemma 12.3). While the transformation is trivial, for completeness the proofs of each are provided as Theorem 13, Lemma 13.1, and Lemma 13.2. In fact, a more thorough discussion between Cholesky and Schur decomposition approaches for matrix square roots is presented in [165].

Theorem 13. *Equation 5.25 provides a consistent and tight approximation of the optimal fusion solution.*

Proof. It is trivial to show that Lemma 12.1 still applies, allowing for $\mathbf{M} = 0$ to be treated as the ultimate upper-bound for the family of covariance matrices represented in Equation 5.25. Additionally, $\boldsymbol{\Omega}^\top \boldsymbol{\Omega} \leq \mathbf{I}$ which is an implied result from Lemma 13.1.

Lemma 13.1. *Equation 5.25 is consistent because* $\boldsymbol{Q}_{ab}^2 \geq \boldsymbol{Q}_a \boldsymbol{\Omega} \boldsymbol{Q}_b$*, where* $\boldsymbol{\Omega}$ *is delimited by* $\rho \boldsymbol{I}$*.*

Proof. Similar to the proof of Lemma 12.2, the consistency definition is applied

$$\mathbf{P}_c - \mathbf{P}_c^o \geq 0$$

It can be shown that by substituting Equation 5.24 and Equation 5.25 into the definition of consistency and using $\mathbf{M} = 0$ providing the ultimate upper-bound results in:

$$\mathbf{Q}_{ab}^2 \geq \mathbf{Q}_a \boldsymbol{\Omega} \mathbf{Q}_b$$

It follows that because $Cov(\mathbf{a}, \mathbf{b}^\top) \geq 0$ then $\mathbb{E}[\mathbf{a}\mathbf{b}^\top] \geq \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}^\top]$. Now, let us state that:

$$\begin{aligned}\mathbb{E}[\mathbf{a}\mathbf{b}^\top] = \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}^\top] = \mathbf{P}_{ab} = \mathbf{Q}_{ab}^2 \\ \mathbb{E}[\mathbf{a}\mathbf{a}^\top] = \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{a}^\top] = \mathbf{P}_{a} = \mathbf{Q}_{a}^2 \\ \mathbb{E}[\mathbf{b}\mathbf{b}^\top] = \mathbb{E}[\mathbf{b}]\mathbb{E}[\mathbf{b}^\top] = \mathbf{P}_{b} = \mathbf{Q}_{b}^2\end{aligned} \tag{5.26}$$

Which allows $\mathbf{Q}_a$ and $\mathbf{Q}_b$ to be considered as analogous matrix forms $\mathbb{E}[\mathbf{a}]$ and $\mathbb{E}[\mathbf{b}]$, respectively. Next, recall the Cauchy-Schwarz inequality with partial knowledge correlation

$$\rho\mathbf{P}_a \geq \mathbf{Q}_{ab}^2(\rho\mathbf{P}_b)^{-1}\mathbf{Q}_{ab}^{2T} \tag{5.27}$$

This is reminiscent of the explicit correlation knowledge condition from Equation 5.3, so by substituting the $\mathbf{Q}_{ab}^2$ approximation from Equation 5.23, $\mathbf{M} = 0$, the definitions in Equation 5.26, and performing some simple algebraic manipulation yields

$$\rho^2\mathbf{I} \geq \mathbf{\Omega}\mathbf{\Omega}^\top \tag{5.28}$$

Where $\rho\mathbf{I}$ is an upper-bound in the solution space and thus proves the Lemma. □

Lemma 13.2. *Equation 5.25 is tight because the ellipse formed by* $\mathbf{\Gamma}^{-1} \leq (\boldsymbol{Q}_a\boldsymbol{\Omega}\boldsymbol{Q}_b + \boldsymbol{Q}_b^\top\boldsymbol{\Omega}^\top\boldsymbol{Q}_a^\top)^{-1} \leq (\omega\boldsymbol{P}_a + (1-\omega)\boldsymbol{P}_b)^{-1}$ *using the Matrix Square Root Decomposition is upper-bounded by the ellipse formed using ICI.*

Proof. Similarly to Lemma 12.3, the Schur Decomposition solution is not in a form to directly allow evaluation against Definition 6. Rather, it is induced that if the ICI solution provides an upper-bound, which has previously been shown to be *tight*, then the Schur Decomposition itself must be *tight*.

$$\mathbf{P}_{ICI} - \mathbf{P}_{schur} \geq 0 \tag{5.29}$$

substituting Equation 2.39 and Equation 5.25 with $\mathbf{M} = 0$ into Equation 5.29 yields,

$$[\mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)^{-1}]^{-1} \geq [\mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b^\top\boldsymbol{\Omega}^\top\mathbf{Q}_a^\top)^{-1}]^{-1}$$

cancelling the applicable terms,

$$(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b) \leq (\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b^\top\boldsymbol{\Omega}^\top\mathbf{Q}_a^\top)$$

In this form, the above equation has no analytical solution. However, the equation may be solved numerically to find a solution for $\boldsymbol{\Omega}$. Because only admissible values for $\boldsymbol{\Omega}$ are assumed, the term $\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b$ is symmetric and may therefore be rewritten as,

$$(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b) \leq 2(\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b)$$

Solving for $\boldsymbol{\Omega}$ results in an upper-bound on the Cholesky Decomposition solution being the ICI solution.

$$\boldsymbol{\Omega} \geq \frac{1}{2}\mathbf{Q}_a^{-1}(\omega\mathbf{P}_a + (1-\omega)\mathbf{P}_b)\mathbf{Q}_b^{-1} \tag{5.30}$$

The bound specified by Equation 5.30 may be generalized to include the explicit partial correlation knowledge as

$$\boldsymbol{\Omega} \geq \frac{1}{2}\mathbf{Q}_a^{-1}(\omega_a\mathbf{P}_a + \omega_b\mathbf{P}_b)\mathbf{Q}_b^{-1} \tag{5.31}$$

where ω_a and ω_b are defined in Equation 5.20. □

Furthermore, because Lemma 13.1 and Lemma 13.2 are proven to be true, Theorem 13 is also upheld. □

Calculating the mutual mean of the Schur Decomposition approach follows analogously

to the method presented in Section 5.2.1. Resulting in

$$\boldsymbol{\gamma} = \mathbf{P}_c(\mathbf{W}_a\mathbf{a} + \mathbf{W}_b\mathbf{b}) \tag{5.32}$$

where $\mathbf{W}_a = \mathbf{P}_a^{-1} - \frac{1}{2}(\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b\boldsymbol{\Omega}\mathbf{Q}_a)^{-1}$ and $\mathbf{W}_b = \mathbf{P}_b^{-1} - \frac{1}{2}(\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b^\top\boldsymbol{\Omega}\mathbf{Q}_a^\top)^{-1}$.

In summary, the Schur Decomposition Approach is formalized in Equation 5.33

$$\begin{aligned}\mathbf{P}_c^{-1} &= \mathbf{P}_a^{-1} + \mathbf{P}_b^{-1} - (\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b^\top\boldsymbol{\Omega}^\top\mathbf{Q}_a^\top)^{-1} \\ \bar{\mathbf{c}} &= \mathbf{P}_c\Big(\mathbf{P}_a^{-1}\mathbf{a} + \mathbf{P}_b^{-1}\mathbf{b} - (\mathbf{Q}_a\boldsymbol{\Omega}\mathbf{Q}_b + \mathbf{Q}_b^\top\boldsymbol{\Omega}^\top\mathbf{Q}_a^\top)^{-1}\boldsymbol{\gamma}\Big)\end{aligned} \tag{5.33}$$

provides a consistent and tight approximation of the optimal solution when $\boldsymbol{\Omega}$ meets the criteria

$$\frac{1}{2}\mathbf{Q}_a^{-1}(\omega_a\mathbf{P}_a + \omega_b)\mathbf{P}_b)\mathbf{Q}_b^{-1} \leq \boldsymbol{\Omega} \leq \rho\mathbf{I}$$

A sample Schur Decomposition Matrix Square Root fusion result is found in Figure 5.2. Much like the Cholesky Decomposition approach, the Schur approach provides tighter approximations of the optimal solution when compared to both the ICI and CI solutions.

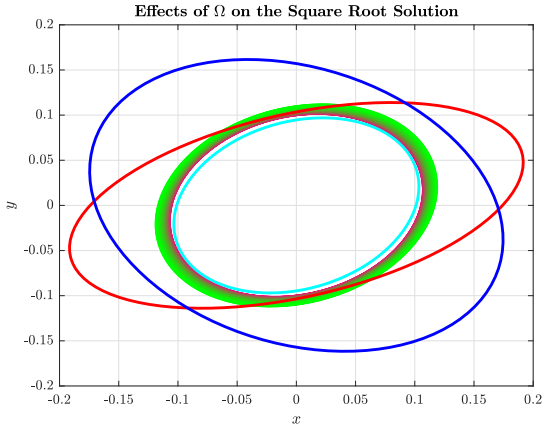


Figure 5.2 Demonstration of how $\boldsymbol{\Omega}$ influences the Square Root Approximation solution by fusing Covariance Ellipse 1 with Covariance Ellipse 2. The **ICI Solution** is used as an upper-bound on the solution and the Optimal Solution provides the lower-bound. Results are presented for bounds on $\boldsymbol{\Omega}$ as Lower-Bound from Equation 5.18 and the Upper-Bound is the **Identity Matrix** because we consider $\rho = 1$ in this example.

However, a few graphical differences are noted between the results Schur results presented in Figure 5.2 and the Cholesky results presented in Figure 5.1. One particular difference is the shape and size of the fusion ellipse as the $\mathbf{\Omega} \to \rho\mathbf{I}$. Within the Schur decomposition approach, as $\mathbf{\Omega} \to \rho\mathbf{I}$, the fused ellipse clearly approaches the optimal solution. On the other hand, for the Cholesky decomposition, this is not necessarily the case.

The cause of these difference is, as of yet, unknown but is a subject of further investigation. It is suspected the explanation lies within the differences in the numerical construction of the matrix square root provided by the Schur decomposition and the Cholesky decomposition.

6 Cooperative Local Catalog Maintenance of Close Proximity Satellite Systems

We now consider the case of multiple attitude-controllable satellites equipped with sensors and on-board estimation capabilities that must track and maintain the states of the other satellites which lay on a closed natural motion trajectory [138] about the local vertical, local horizontal (LVLH) frame of a virtual chief. The problem which considers a single satellite (chief) which maintains a catalog of multiple other satellites (deputies) has been studied in Chapter 3 and [26], where it is referred to as the *local catalog maintenance problem*. This work adopts the same moniker and generalizes the problem to include multiple satellites that may not necessarily be at the origin of the LVLH frame of the virtual chief. To avoid confusion, we refer to the satellite under consideration at a specific time as "agent" and the other satellites as "deputies". This distinction is necessary because the attitude control laws and target deputy decisions are locally computed at each capable satellite; that is, each satellite considers itself to be the "agent" and every other satellite to be "deputies". We refer to the set of cooperative satellites as "agents" and the set of non-cooperative objects as "objects". The local catalog we wish to maintain is the concatenated list of all the agents' and objects' positional and velocity states relative to a virtual chief. However, a satellite's state is inherently uncertain. Generally, these state uncertainties are represented through stochastic parameters, such as a mean and state covariance matrix pair for a Gaussian probability distribution, also known as the "belief state." Through the use of a Bayesian filter [149], the fusion between the agent's prior belief state and an observation of some subset of deputies results in a posterior belief state with reduced uncertainty. Crucially, the agent cannot observe the states of all the targets simultaneously and is only able to make measurements of a specific target when it lies within the agent's field-of-view. Thus, the uncertainties of unobserved deputies will grow accordingly with the amount of time elapsed without measurement. This engenders the core of the local catalog maintenance problem: how does each agent know when and where to look in order to provide a catalog of target states that are closest to their true values?

Moreover, we assume each agent is capable of communicating with the other satellites in the environment. Thus, each agent may fuse the information provided by its neighbors into its own belief state; which introduces a secondary problem: what communication links may be established, if any exist, that increase the quantity and quality of information flowing into each satellite? If a system is capable of converging the estimate error and uncertainty under a desired threshold, we refer to that system as being state omniscient.

6.1 Problem Statement

Consider a set of N_a agents, denoted $\mathcal{V}_a = \{1, \ldots, N_a\}$, where each agent is tasked with maintaining a catalog of objects $\mathcal{V}_u = \{1, \ldots, N_u\}$, every other agent, and itself. Each agent $i \in \mathcal{V}_a$ has a state space $\mathcal{X}_i \subseteq \mathbb{R}^n$ and its state evolves according to the CWH equations given in (2.17) and its attitude evolves according to (2.19) and (2.23) and has a torque input space $\mathcal{U}_i \subseteq \mathbb{R}^q$. Each object $i \in \mathcal{V}_u$ has a state space $\mathcal{X}_i \subseteq \mathbb{R}^n$ and its state evolves according to the CWH equations given in (2.17). The global system state is a concatenation of every agent and object state $\mathbf{x}^\top = [\mathbf{x}_1^\top, \ldots, \mathbf{x}_N^\top]^\top$. Each agent for each $i \in \mathcal{V}_a$ must maintain a local estimate of the global system state $\hat{\mathbf{x}}_i \in \mathbb{R}^{nN}$ with $N = N_a + N_u$. Each estimator must maintain its certainty in the global system state below a certain threshold $v(\mathbf{P}_i(k)) \leq \epsilon$. The agents can communicate their estimates and measurements through a communication graph $\mathcal{G}_c$. The agents may accomplish this by slewing their attitude towards a selected target and taking a measurement. A measurement will be taken as long as a body is within the FOV of the agent. Furthermore, we consider a constraint function $c : \prod_{i \in \mathcal{V}_a} \mathcal{X}_i \times \prod_{i \in \mathcal{V}_a} \mathcal{U} \to \mathbb{R}^r$ where global state $\mathbf{x}$ and torque inputs $\{\boldsymbol{\tau}_i\}_{i \in \mathcal{V}_a}$ satisfy the constraints when $c_l(\mathbf{x}, \boldsymbol{\tau}) \leq 0$ for each row $l \in \{1, \ldots, r\}$ in the constraint vector. The problem is then to specify communication graph $\mathcal{G}_c$ and control inputs $\boldsymbol{\tau}_i(k)$ for each agent $i \in \mathcal{V}_a$ at each time $k \geq 0$ such that $v(\mathbf{P}_i(k)) \leq \epsilon \; \forall \; k^* \leq k < \infty$ where k^* is when $v(\mathbf{P}_i(k^*)) = \epsilon$ for each $i \in \mathcal{V}$ and $c_l(\mathbf{x}(k), \boldsymbol{\tau}(k)) \leq 0$ for each row $i \in \{1, \ldots, r\}$. That is, the distributed estimator is state omniscient. See Chapter 4 for a further discussion on state omniscience.

6.2 Approach

To solve the cooperative local catalog maintenance problem, we propose a multi-layered closed-loop control strategy (see Fig.6.1) that incorporates continuous dynamics with impulsive zero-order hold controls applied at discrete time instances. The blue dashed line indicates the recursive nature of the Bayesian filter, where updated belief states are supplied to the propagation block for the next time step. Both the supervisor and controller are fed from these state estimates to make high-level target selection decisions and low-level control decisions. More specifically, at a given time instant k, the state estimator of agent i fuses predicted deputy belief states $(\hat{\mathbf{x}}_i^-(k), \mathbf{P}_i^-(k))$ with available observation data $\mathbf{y}_i(k)$ and the information provided by its neighbors in $\mathcal{G}_c:\ (\hat{\mathbf{x}}_j^-(k), \mathbf{P}_j^-(k)), \mathbf{y}_j(k))$ where $j \in \mathcal{N}_i$ to yield posterior deputy belief states $(\hat{\mathbf{x}}^+(k), \mathbf{P}^+(k))$ through a Bayesian update scheme such as the commonly-used extended Kalman filter or unscented Kalman filter, this work uses the Networked Distributed Kalman Estimator as described in (6.7). A centralized supervisor (possibly a ground station or an on-orbit orchestrator) dictates the target for each agent, but each agent runs its own controller and its own estimator. Preliminary versions of the supervisor algorithm and model predictive controller executed at each agent are presented in [26]. A separate version of the supervisor algorithm along with a preliminary communication graph construction algorithm for full state measurements are presented in [26, 28]. The remainder of this section updates the problem formulation and the preliminary algorithms to a more general class of scenarios and utilizes new knowledge about the interactions of system dynamics, sensed information, and communicated information inspired by the necessary and sufficient conditions for state omniscience.

6.2.1 Distributed Bayesian Filter

Each agent is equipped with a local state estimator to track the N bodies of the form (6.7) where a communication graph $\mathcal{G}_c$ dictates the structure of $\mathbf{W}$ and, therefore, the flow of information sharing. We treat this distributed estimator as a distributed Extended Kalman Filter similar to the one presented in [76, 77] and parameterized by augmented

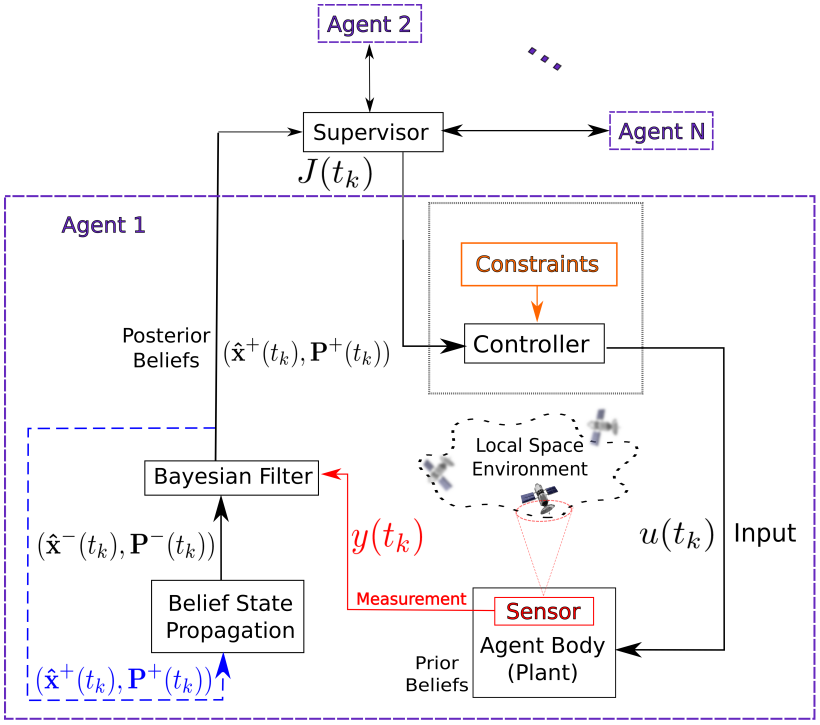


Figure 6.1 Block diagram of the proposed cooperative catalog maintenance system architecture.

mean $\hat{\mathbf{x}} := [\hat{\mathbf{x}}_1^\top \ \dots \ \hat{\mathbf{x}}_{N_a}^\top]^\top \in \mathbb{R}^{nNN_a}$ and block covariance matrix

$\mathbf{P} := \text{diag}([\mathbf{P}_1, \dots, \mathbf{P}_{N_a}]) \in \mathbb{R}^{nNN_a \times nNN_a}$ with each $\mathbf{P}_i \in \mathbb{R}^{nN}$ and $\mathbf{P}$ evolves according to

$$\mathbf{P}(k+1) = \mathbf{\Phi}\mathbf{P}(k)\mathbf{\Phi}^\top + \mathbf{Q}^{proc} \tag{6.1}$$

where, $\mathbf{Q}^{proc} := \text{diag}([\mathbf{Q}_1^{proc}, \dots, \mathbf{Q}_{N_a}^{proc}]) \in \mathbb{R}^{nNN_a \times nNN_a}$ is the process noise matrix.

Each agent is equipped with a sensor parameterized by its pointing vector in the agent's body frame $\mathbf{p}^{\mathscr{B}}$ and its angle of view α. The sensor can observe the states of the deputies within some sensing field-of-view (FOV) defined by the pair $(\mathbf{p}^{\mathscr{B}}, \alpha)$, the agent's current state $\mathbf{x}_i(k)$, and the agent's current orientation $\mathbf{R}_i(k)$. If agent j is within the FOV of agent i, then we apply a nonlinear angles-only measurement model such that

$$\mathbf{y}_{ij}(k) = g(\mathbf{x}_j - \mathbf{x}_i) = \frac{\boldsymbol{\rho}}{||\boldsymbol{\rho}||} \tag{6.2}$$

where in this case $\boldsymbol{\rho} = \mathbf{x}_j - \mathbf{x}_i$ and $\rho := [\rho_1 \ \rho_2 \ \rho_3]^\top$. The linearized measurement model of the sensor is given by

$$\mathbf{C}_i = \mathbf{E}_i^\top \otimes \mathbf{H} \tag{6.3}$$

where $(\mathbf{E}_i^\top)_{ij}$ the incidence matrix and is given as

$$(\mathbf{E}_i^\top)_{ij} = \begin{cases} 1 & \text{if agent } j \text{ is observed by agent } i \\ -1 & j = i \\ 0 & \text{else.} \end{cases} \tag{6.4}$$

and $\mathbf{H}$ is the local measurement matrix and is given by

$$\mathbf{H} = \frac{1}{||\boldsymbol{\rho}||}\left(\mathbf{I} - \frac{\boldsymbol{\rho}^\top \boldsymbol{\rho}}{||\boldsymbol{\rho}||^2}\right). \tag{6.5}$$

Almagamating all of the individual systems into a single LTI system yields

$$\begin{aligned} \dot{\bar{\mathbf{x}}} &= \big(\mathbf{I}_\otimes \boldsymbol{\Phi}\big)\bar{\mathbf{x}} \\ \mathbf{y} &= (\mathbf{E}_s^\top \otimes \mathbf{H})\bar{\mathbf{x}} = \mathbf{C}_{\mathbf{A}}\bar{\mathbf{x}} \end{aligned}. \tag{6.6}$$

It is from this form that we may consider the distributed estimator. Whenever the agent makes an observation of another body $j \in \mathcal{V}$ at a time instant k, the measurement data is incorporated by the on-board estimator to impulsively update the local state estimate such that

$$(\hat{\mathbf{x}}_i(k^-), \mathbf{P}_i(k^-)) \to (\hat{\mathbf{x}}_i(k^+), \mathbf{P}_i(k^+)).$$

As a result, the magnitude of the covariance matrices of observed deputies will necessarily decrease $v(\mathbf{P}_i(k^+)) \leq v(\mathbf{P}_i(k^-))$ and the updated belief state will be recursively supplied to the propagator until the next available measurement update. Here $g[\cdot]$ is any metric quantifying the magnitude of the covariance matrix or its representative hyperellipsoid.

However, because the sensor has a limited FOV, it may not be possible for agent i to observe every other body j at once. Therefore, agent i must maneuver to view each other body j to maintain a level of certainty in the state estimate, that is $v(\mathbf{P}_i(k)) \leq \epsilon$ for all $k \geq 0$, where ϵ is some positive value determined by the sensor FOV. Or, agent i may leverage communication amongst the other agents to maintain $v(\mathbf{P}_i(k)) \leq \epsilon$ for all $k \geq 0$. The communicated information $(\hat{\mathbf{x}}_j(k), \mathbf{y}_j(k))$ for each $j \in \mathcal{N}_i$ may be fused to the current state estimate $\hat{\mathbf{x}}_i(k)$ or current measurement $\mathbf{y}_i(k)$. This communicated information may be fused to the local scheme via any fusion method, we choose to use Inverse Covariance Intersection (ICI) [109]. This particular work utilizes the filter known as the Extended Networked Distributed Kalman Estimator [77] because it has the potential to fuse both the measurement and state estimate at each agent so that its local state estimate converges to that of the true system state.

$$\begin{aligned}
&\text{Predictor Fusion} \begin{cases} \hat{\mathbf{x}}_i^-(k+1) &= \sum_{j\in\mathcal{N}_i} \mathbf{W}_{ij}\mathbf{\Phi}\hat{\mathbf{x}}_j(k) \\ \mathbf{P}_i^-(k+1) &= \sum_{j\in\mathcal{N}_i} \mathbf{W}_{ij}\mathbf{\Phi}\mathbf{P}_j(k)\mathbf{\Phi}^\top\mathbf{W}_{ij}^\top + \mathbf{Q}_j^{proc} \end{cases} \\
&\text{Output Fusion} \begin{cases} \hat{\mathbf{x}}_i(k+1) = \hat{\mathbf{x}}_i^-(k+1) + \sum_{j\in\mathcal{N}_i} \mathbf{L}_i\mathbf{U}_{ij}\mathbf{C}_j^\top(\mathbf{y}_j(k) - \mathbf{C}_j\hat{\mathbf{x}}_i(k)) \\ \mathbf{S}_i = \mathbf{U}_i\mathbf{C}_\mathbf{A}\mathbf{P}_i^-(k+1)\mathbf{C}_\mathbf{A}^\top\mathbf{U}_i^\top + \mathbf{Q}_{\mathcal{N}_i}^{meas} \\ \mathbf{L}_i = \mathbf{P}_i^-(k+1)\mathbf{C}_\mathbf{A}^\top\mathbf{U}_i^\top\mathbf{S}_i^{-1} \end{cases}
\end{aligned} \tag{6.7}$$

where $\mathbf{W}_{ij}$ and $\mathbf{U}_{ij}$ are fusion weight terms that are dependent on the fusion rule selected [106, 109, 114]; in this case, we use inverse covariance intersection (ICI) [109] when multiple pieces of information need to be fused. We use a matrix decomposition approach when only two pieces of information need to be fused [114]. The output fusion matrix $\mathbf{U}_i := [\mathbf{U}_{i1}, \ldots, \mathbf{U}_{iN}]$ is a block row vector comprised of all of the output fusion matrices that come from a predecessor of agent i in $\mathcal{G}_c$. It is noted in Section 4.3.3 that $\mathbf{W}$, $\mathbf{U}$, $\mathbf{L}$ exist that may converge each local estimator if and only if all of the observable modes are

within the walk set of agent i. It should be noted that as a consequence of relative measurements, a coordinate transform is required to place the measurements into the same frame prior to fusion. This work elects to avoid this step altogether by solely considering state estimate predictor fusion by setting $\mathbf{U} = \mathbf{I}$.

6.2.2 Supervisor

These updated beliefs are fed into a higher-level autonomous "supervisor" which accounts for (i) belief state information gaps (ii) prespecified constraints and (iii) the current angular states of the agents. The supervisor amalgamates this information and provides decision-making capabilities to a lower-level controller at each agent, which drives the agent's orientation trajectory by administering torque inputs at discrete timesteps. Recall the local catalog maintenance problem where an agent, with a sensor parametrized by $\mathbf{p}^{\mathscr{B}}$ and α, must maintain state estimates of N bodies in its local environement.

The objective of the supervisor algorithm is to select the "optimal" target body for each agent to track based on some cost function notated by $\mathcal{J}$. A detailed breakdown of the supervisor algorithm can be seen in Algorithm 2. Given the state estimate-covariance pairs of the bodies $\{(\hat{\mathbf{x}}_j, \mathbf{P}_j)\}_{j=1}^{N}$ and the current target body j^*. Additionally, we introduce a new variable Δt^* which "remembers" the time that the supervisor switched to deputy j^*. The output is the body j' to be viewed as well as Δt^*. Then, the supervisor algorithm is the following. If some Δt time has not passed since the last switching time, or the entropy of the current deputy has not fallen below some threshold ϵ, the current body j^* and current Δt^* is returned. If not, then the entropy of all the deputies is computed. The target to be viewed is chosen to be the one with maximum entropy, and Δt^* is updated to the current time.

The supervisor cost function in this work is considered to be the Shannon entropy of the covariance matrix which is assumed to be Gaussian and is given by

$$\mathcal{J} = \frac{n}{2}(1 + \log(2\pi)) + \log(|\mathbf{P}_j|);$$

other metrics such as measures of uncertainty relating to the deputies' belief states, such as the det or trace() of the covariance matrix, or a measure of information gain post-fusion, such as the posterior Fisher information [150] may also be utilized. The cost value for each target is calculated, and then the target with the highest cost value is selected to be the agent's target j'. However, to prevent the possibility of instantaneous switching, a form of hysteresis on the dwell time is applied. A target switch is only triggered when the cost of the current target deputy falls below some threshold, the potential other target deputy crosses over the same threshold, and two target switches cannot occur within a time threshold, $\delta \geq \mu$, of each other, where δ is the time differential between switches, and μ is the desired time hysteresis threshold.

Algorithm 2: Supervisor algorithm

Input: $\{(\hat{\mathbf{x}}_j, \mathbf{P}_j)\}_{j=1}^{N}$, j^* - current target, Δt^* - previous switch time
Output: j', Δt^*_{new}

$t \leftarrow \mathsf{CurrentTime}$
if $(t - \Delta t^* \leq \Delta) \ \vee \ (\frac{n}{2}(1 + \log(2\pi)) + \log(|\mathbf{P}_j|) > \epsilon)$ **then**
    $j' \leftarrow j$
    $\Delta t^*_{new} \leftarrow \Delta t^*$
else
    for $i = 1, \ldots, N$ **do**
        $\mathcal{J}_i \leftarrow \frac{n}{2}(1 + \log(2\pi)) + \log(|\mathbf{P}_i|)$
    $j' \leftarrow \arg\max_{j \in \{1,\ldots,N\}} \mathcal{J}_j$
    $\Delta t^*_{new} \leftarrow t$
return j', Δt^*_{new}

6.2.3 Communication Graph Constructor

At a given time instant, assume $\mathcal{G}_s$ is available. The goal is to then construct $\mathcal{G}_c$ to complement $\mathcal{G}_s$ such that $\mathcal{G} \leftarrow \mathcal{G}_s \cup \mathcal{G}_c$ is state omniscient. Alternatively, it is sufficient to show that the generic observability matrix in

$$\mathbf{G}_i = \sum_{j \in \mathcal{W}_i} \mathbf{C}_j^T \mathbf{C}_j \tag{6.8}$$

is full rank for each $i \in \mathcal{V}$. This approach to constructing state omniscient distributed estimators is first considered in [27], albeit under different terminology. We extend the construction algorithm presented in [27] with the more recent results on state omniscience - there must exist a walk from each observable mode to each agent i.

Naïvely, it could be said that through set algebraic properties that $\mathcal{G} \leftarrow \mathcal{G}_s \cup \mathcal{G}_c$ must be strongly connected and therefore there exists a walk from every vertex to every other vertex. As a result, $(\mathbf{I}_N + \mathbf{A})^{N-1} \succ 0$, where $\succ$ is the element-wise greater-than [166]. $\mathbf{A}$ may then be constructed to be a strongly connected graph. However, this approach neglects the original problem of information being unable to pass through a vertex without a communication edge to pass along information. With this in mind, we consider $\partial^{in}\mathcal{G}_c := \mathcal{G}_s$ so the boundary of the communication graph is the sensing graph, which transforms $\mathcal{G} \leftarrow \mathcal{G}_s \cup \mathcal{G}_c$ to $\mathcal{G} \leftarrow \partial^{in}\mathcal{G}_c \cup \mathcal{G}_c$ allowing us to use $\mathcal{G}_s$ as the foundation to build $\mathcal{G}_c$. The communication graph is constructed by adding a directed edge between each in-vertex of $\mathcal{G}_s$ and every other agent j. A sample $\mathcal{G}_c$ construction can be seen in Figure 6.2.

(a) $\mathcal{G}_s$ *(b)* $\mathcal{G} := \mathcal{G}_c \cup \partial^{in}\mathcal{G}_c$

Figure 6.2 Graph models illustrating the construction of $\mathcal{G}_c$ from $\mathcal{G}_s := \partial^{in}\mathcal{G}_c$

This method of construction is susceptible to redundant edges that could be pruned. Rather, this work employs a brute force algorithm to "clean-up" $\mathcal{G}_c$ or remove redundant edges that do not improve the state omniscience objective. By doing nothing redundant edges remain and could potentially take up necessary bandwidth and computational load on system, but the state omniscient condition is met. Future work may address this issue by exploring performance-based costs for edge pruning. To alleviate potential and unnecessary computational load, these redundant edges

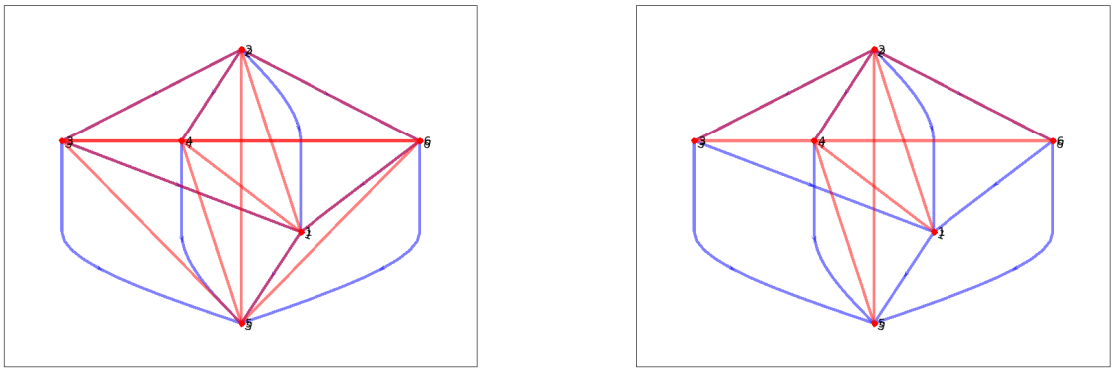
are pruned by looping through each edge in $\mathcal{G}_c$ and checking the state omniscient condition and calculating the associated cost $\mathcal{J}$. If the state omniscient conditions is met, then the cut-edge is a redundant edge and may be removed. This process repeats until the state omniscient condition is violated. In a way, this approach can be considered a brute force approach to the optimization problem where $\mathcal{J}$ is the cost function and the state omniscient condition is a constraint. This work employs the edge cardinality $\mathcal{J} = E_c$ as the cost. Algorithm 3 shows each of these steps in more detail.

Algorithm 3: Communication Graph Construction Algorithm

Input: $\mathcal{G}_s$
Output: $\mathcal{G}_c$
for $i \in \mathcal{V}$ **do**
  if $\Delta_s(i) > 0$ **then**
    $\mathcal{G}_c^i \leftarrow \{e_{ij} \in \mathcal{E}_c | j \in \mathcal{V}\}$
$\mathcal{G}_c \leftarrow \bigcup_{i \in \mathcal{V}} \mathcal{G}_c^i$
$\mathcal{G}_c^* = \emptyset$
while $\mathcal{G}_c^* \neq \mathcal{G}_c$ **do**
  $\mathcal{G}_c^* \leftarrow \mathcal{G}_c$
  $\mathcal{J}_0 \leftarrow E_c^*$
  for $e \in \mathcal{G}_c$ **do**
    $\mathcal{G}_e \leftarrow \mathsf{RemoveEdge}(\mathcal{G}_c^*, e)$
    if *StateOmniscient*$(\mathcal{G}_e)$ **then**
      $\mathcal{J}_e \leftarrow E_e$
  $\mathcal{G}_c \leftarrow \arg\min_{\mathcal{G}_e} \mathcal{J}_e$

Theorem 14. *If a networked LTI system described according to* (6.6) *is observable, then a communication graph constructed according to Algorithm 3 will be state omniscient.*

Proof. To begin the proof we must first recall that a system will be state omniscient if the global system is observable and there exists a walk from the source of each observable mode to every other agent within the network. By this definition, Algorithm 2 must ensure a walk exists from the source of every observable mode to every other agent within the network. Assuming the global system is observable, Algorithm 3 begins by adding an edge



(a) Initial Graph *(b)* Final Graph

Figure 6.3 Graph topology after pruning

to the communication graph from each observable mode source to every other agent in the network. By construction this iteration of $\mathcal{G}_c$ meets the state omniscience condition. In practice, we ensure that $\mathbf{W}_i\mathbf{s} \neq \mathbf{0}$ for each observable mode and each agent $i \in \mathcal{V}_a$ and where $\mathbf{s} \in \mathbb{R}^N$ and is populated by

$$\mathbf{s}_j = \begin{cases} 1, \text{ mode is observable from agent } j \\ 0, \text{ otherwise} \end{cases}.$$

Consequently, this allows for the algorithm to check against the state omniscience condition everytime an edge is removed from $\mathcal{G}_c$ and therefore the state omniscient condition will never be violated. The result follows. □

It must also be noted that Algorithm 2 will only guarantee the existence of a walk from an observable for a particular mode only if that mode is observable somewhere in the network. If a specific mode is not observable anywhere in the network, then Algorithm 2 will only ensure the state omniscient condition for observable modes. Figure 6.3 demonstrates the communication graph pruning process at the initial and final steps. In this particular example, seven edges were pruned from the communication graph while the state omniscient condition is maintained. This sample assumed all of the vertices are

agents capable of communication.

7 Results

This section contains numerical results to demonstrate the efficacy of the proposed algorithms. For these sample scenarios, the target for each agent is selected locally via the Supervisor Algorithm detailed in Algorithm 1. The MPC algorithm, also computed locally, then uses the azimuth-elevation tracks of agent j' as the reference trajectory. The CasADi optimization toolbox is used to solve the nonlinear optimal control problem [151] using a direct single shooting method. The MPC cost function weighting matrices were set as $\mathbf{W}_1 = \mathbf{I}_2$ and $\mathbf{W}_2 = \mathbf{I}_3$. The torque constraint was set to be $u_{max} = 2\pi$ Nm, the angular velocity constraint was set as $\omega_{max} = \pi$ rad/s. Each agent's initial attitude and angular velocity were randomly generated with the initial angular velocity being restricted to a bounded box of $\pi/2$ m/s about the origin. Each agent was placed on an elliptical NMT with an initial position randomly selected uniformly from a 100 m bounding box about the origin. The bodies' initial position was randomly selected from the same 100 m bounding box about the origin with an initial velocity that allowed it to slowly escape from the local neighborhood - see Figure 7.1 that has 4 agents and 2 objects. Because the focus of this work is on target selection and control, the initial estimates were set to within 10^1 m and 10^{-3} m/s of truth values for position and velocitiy, respectively and the initial covariances for each target were initialized accordingly. The process and noise covariance matrices were set to $\mathbf{Q}^{proc} = 10^{-3}$ and $\mathbf{Q}^{meas} = 10^{-4}$, respectively. For the supervisor algorithm, the threshold was set to $\epsilon = -45$ with a hysterisis value of $\mu = 20$ s.

The efficacy of the proposed algorithms is demonstrated through numerical results for a total of 120 simulations of 500 s each. The simulation trials are broken down into 12 categories by sensor FOV, agent-to-body ratio, and communication graph construction type. More specifically, we consider both 10° and 45° sensor FOV; we consider 5 agents with agent-to-body ratios of 1:1, 1:2, and 1:5 for 5 bodies, 10 bodies, and 25 bodies, respectively, and we consider both pruned and unpruned communication graph construction types.

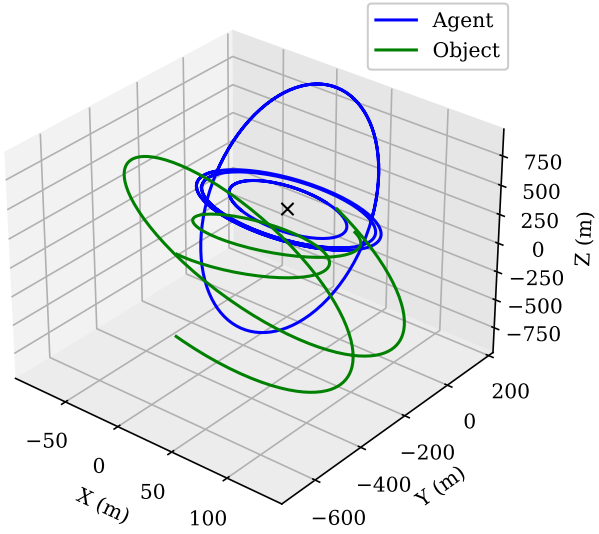
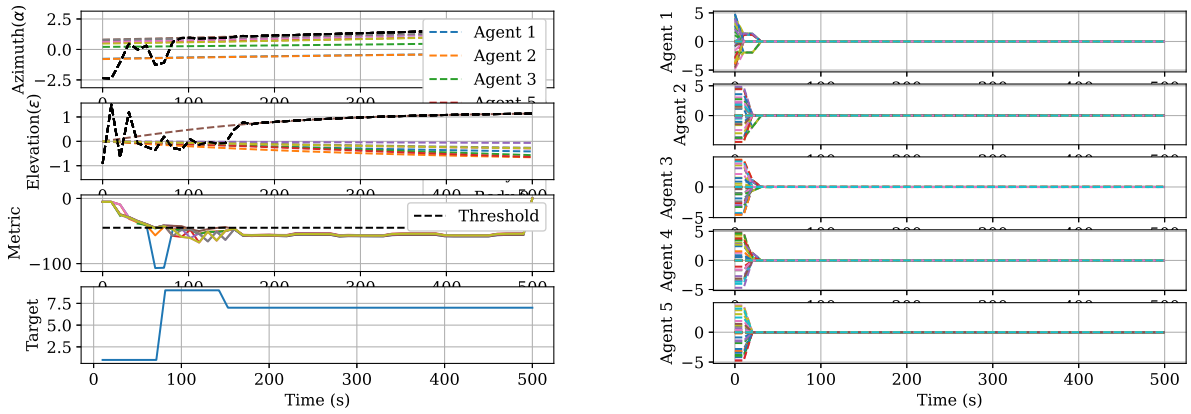


Figure 7.1 Scenario Example with 4 Agents and 2 Bodies

7.1 Relative Position Measurements Numerical Results

This section considers results when relative position measurement model is employed. In Figure 7.2a, the supervisor is able to select an appropriate target for each agent and the local MPC is able to maintain the reference trajectory of the target. From Figure 7.2b, it is also seen that the estimated states are converging to the correct state for every agent for the chosen sample scenario.

Moreover, Figures 7.3-7.7 considers the results from all 120 simulation runs, holistically. Specifically, Figure 7.3 presents the root-mean-square-error (RMSE) of the state estimate across all of the agents within the network. Visually, the RMSE of the pruned cases oscillate as a result of the consistent switching and varying information availability at each agent over time. The RMSE for some of the trials diverge which may be attributed to overconfidence in the estimate. The Kalman filter trusts the prediction state over the measurements more than it should be which results in a divergence. Evidence of this overconfidence may be seen in Figure 7.5 with the outlying entropy values near -125. In



(a) Simulation Example *(b)* Estimate Error

Figure 7.2 5 agents tracking 5 bodies with communication enabled for relative position measurements.

fact, Figure 7.4, which plots the number of edges in the communication graph, suggests that for the relative position case edge pruning siginificantly reduces the number of edges required to maintain a state omniscient system. While the unpruned cases are primarily fully connected graphs, the pruning removes 5 edges, at a minimum. On average, pruning leaves the communication graph with 10 edges for a network of agents with 10° FOV sensors, and 6 edges for a network of agents with 45° FOV sensors.

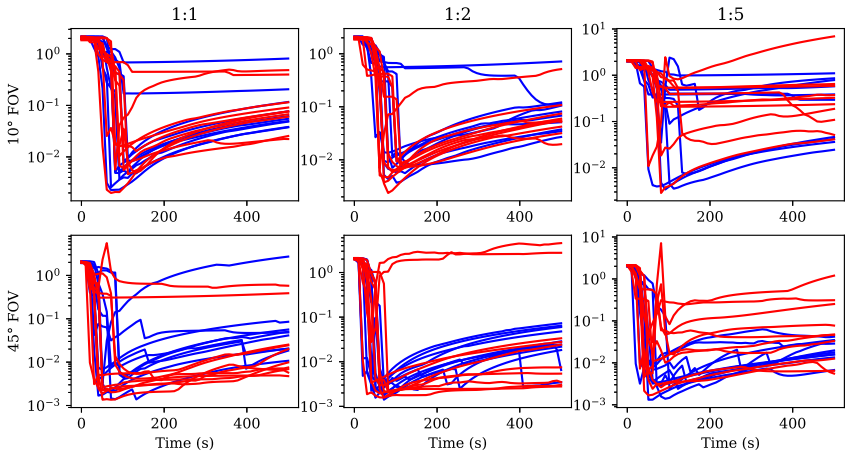


Figure 7.3 RMSE across network for pruned and unpruned communication graphs for relative position measurements.

Additionally, wider FOVs seem to be able to improve estimator performance through lower RMSE across the network due to the increased information availability. Wider FOVs imply more sensor can track a greater number of objects at any given time which improves fusion performance. It is the increased information availability provided by wider FOVs that causes the Shannon entropy to reach and remain below the desired threshold. Fortunately, the agent-to-body ratio has little effect on estimator performance as the estimator is able to maintain similar levels of performance across all agent-to-body ratios for a given FOV.

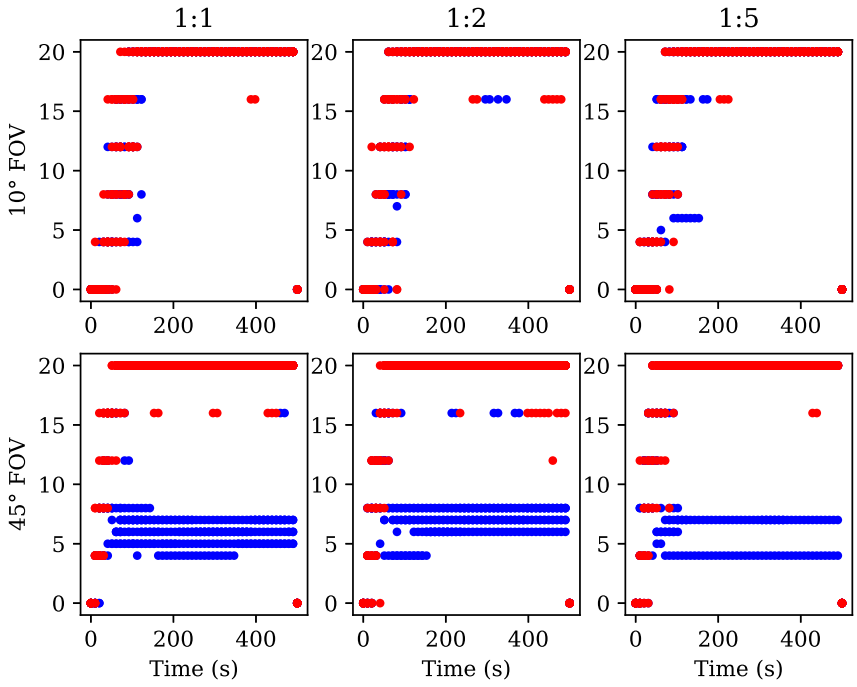


Figure 7.4 Number of edges in the communication graph for pruned and unpruned cases with relative position measurements.

Figure 7.5 presents the Shannon entropy for all 120 runs and shows the entropy is able to converge under and remain under the desired threshold for all of them. The numerical results presented demonstrate the capability of the algorithm to achieve and maintain state omniscience. Oddly though, for both the unpruned and pruned cases a ”yo-yoing” affect appears where the entropy will yo-yo between a lower-value and an upper-value. As of yet,

the cause is unknown, and only appears in the relative position results and does not appear in the angles-only results.

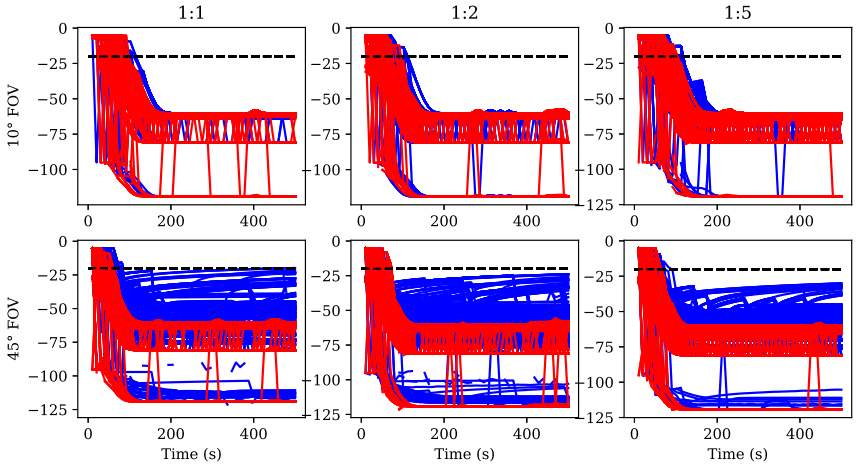


Figure 7.5 Shannon entropy for every agent's certainty of every other tracked object for pruned and unpruned communication graphs under the desired threshold for relative position measurements.

In fact for pruned cases, a 45° FOV will reach the threshold between 60-80 seconds where a 10° FOV will reach the threshold between 100-110 seconds, on average. On the other hand for unpruned cases a 45° FOV will reach the threshold between 90-100 seconds where a 10° FOV will reach the threshold 150-200 seconds, on average but with a larger variance. In turn, reaching the desired threshold faster means less overall switching for the communication graph indicated by both Figure 7.6 and Figure 7.7.

Moreover, a plot of the switch time history in Figure 7.7 indicates the pruned cases will consistently switch to maintain the uncertainty threshold. This also may be an effect of the hysterisis being shorter than the time it takes information to flow from one end of the network to the other according to the communication graph's diameter. If a switch occurs before all information makes it to an agent, that agent will not have received all information the state omniscient condition says it should have.

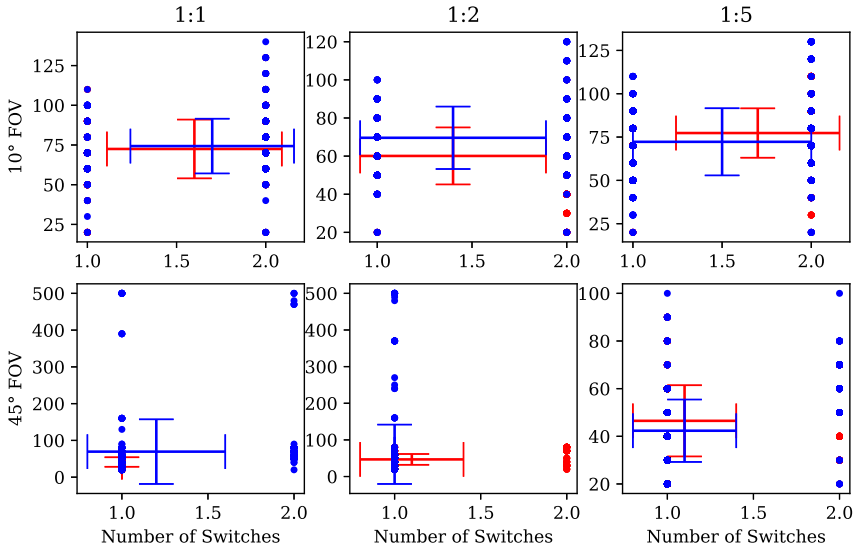


Figure 7.6 Metric time to desired threshold plotted against number of communication graph switches for pruned and unpruned cases with relative position measurements.

7.2 Angles-Only Measurements Numerical Results

This section considers results when an angles-only measurement model is employed. In Figure 7.8a, the supervisor is able to select an appropriate target for each agent and the local MPC is able to maintain the reference trajectory of the target. From Figure 7.8b, it is also seen that the estimated states are converging to the correct state for every agent for the chosen sample scenario.

Moreover, Figures 7.9-7.13 considers the results from all 120 simulation runs, holistically. Specifically, Figure 7.9 presents the root-mean-square-error (RMSE) of the state estimate across all of the agents within the network. Visually, the RMSE of the pruned cases oscillate as a result of target switching, objects coming in and out of view, and varying information availability at each agent over time. In fact, Figure 7.10, which plots the number of edges in the communication graph, suggests that for the angles-only case edge pruning reduces the number of edges required to maintain a state omniscient system. The unpruned cases are primarily fully connected graphs, while pruning does occur for the 10° FOV cases, it is more pronounced in the 45° case because each agent has

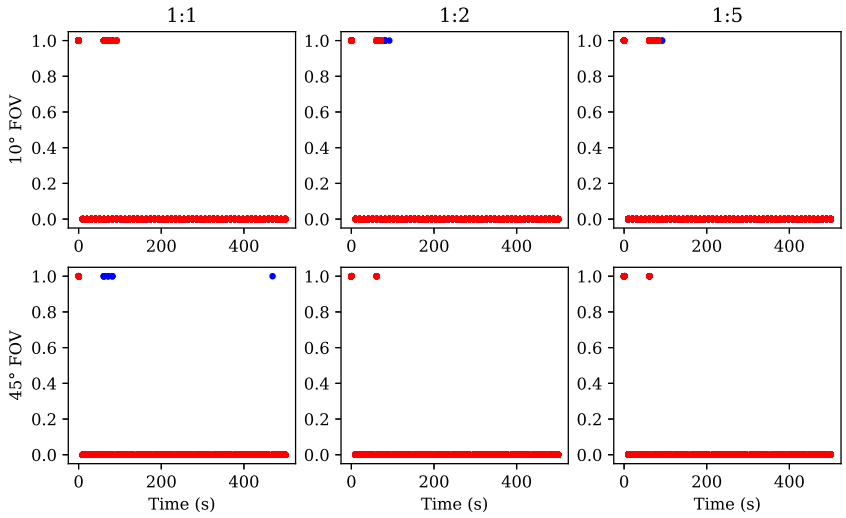
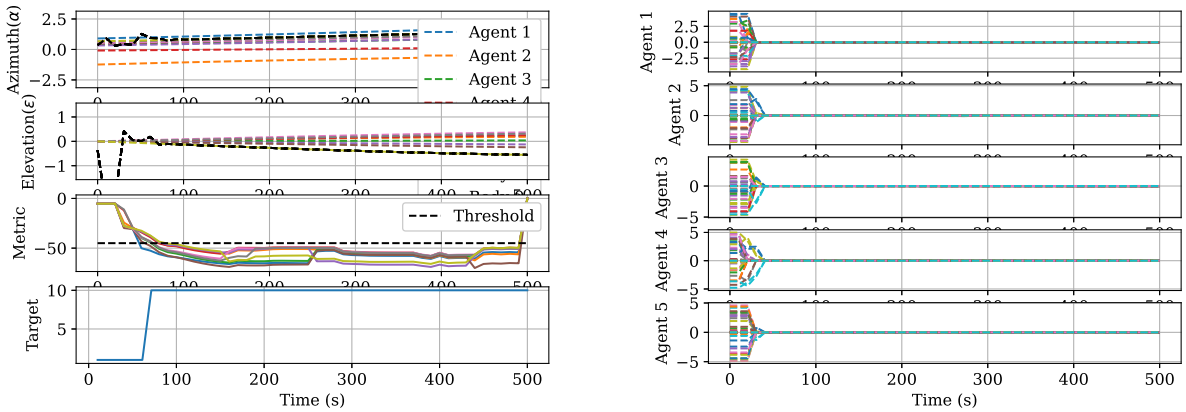


Figure 7.7 Switch history for pruned and unpruned communication graphs for relative position measurements.

access to more measurements with a wider FOV and is not as reliant on communication. On average, pruning leaves the communication graph with 6 edges for a network of agents with 45° FOV sensors.

Additionally, wider FOVs seem to be able to improve estimator performance through lower RMSE across the network due to the increased information availability. Fortunately, the agent-to-body ratio has little effect on estimator performance as the estimator is able to maintain similar levels of performance across all agent-to-body ratios for a given FOV.

Figure 7.11 presents the Shannon entropy for all 120 runs and shows the entropy is able to converge to and stay near the desired uncertainty threshold. For the 10° FOV cases the existence of more edges within the communication graph induces a greater number of fusion elements at each agent which drastically reduces the associated uncertainty of a given estimate. Because the 45° FOV has less edges, each agent has access to less information to collapse the associated uncertainty. In the absence of sufficient information to combat the range ambiguity, the associated range uncertainty will tend to grow inducing an overall (albeit slower) growth in the Shannon entropy. However, the estimator is still able to maintain the desired uncertainty under the threshold. The numerical results



(a) Simulation Example

(b) Estimate Error

Figure 7.8 5 agents tracking 5 bodies with communication enabled and for angles-only measurements

presented demonstrate the capability of the algorithm to achieve and maintain state omniscience. The distributed estimator is even capable of reaching and maintaining the desired uncertainty threshold.

In fact, a 45° FOV will reach the threshold between 60-80 seconds where a 10° FOV will also reach the threshold between 60-80 seconds, on average. Coincidentally for unpruned cases a 45° FOV will reach the threshold between 60-80 seconds where a 10° FOV will also reach the threshold 60-80 seconds, on average but with a larger variance. In turn, wider FOVs imply less overall switching for the communication graph indicated by both Figure 7.12 and Figure 7.13.

Moreover, a plot of the switch time history in Figure 7.13 indicates all of the switching occurs early in the simulation while the estimator is converging to the true state. Once converged, less switching is required because the uncertainy stays below the desired threshold.

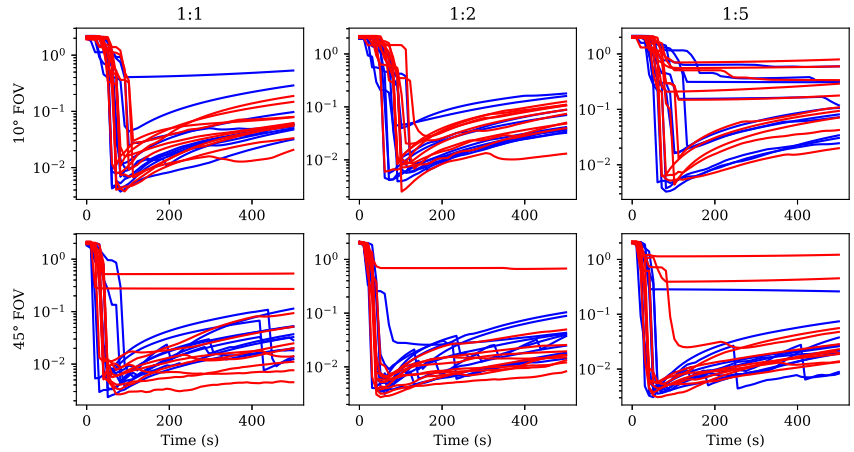


Figure 7.9 RMSE across network for pruned and unpruned communication graphs for angles-only measurements.

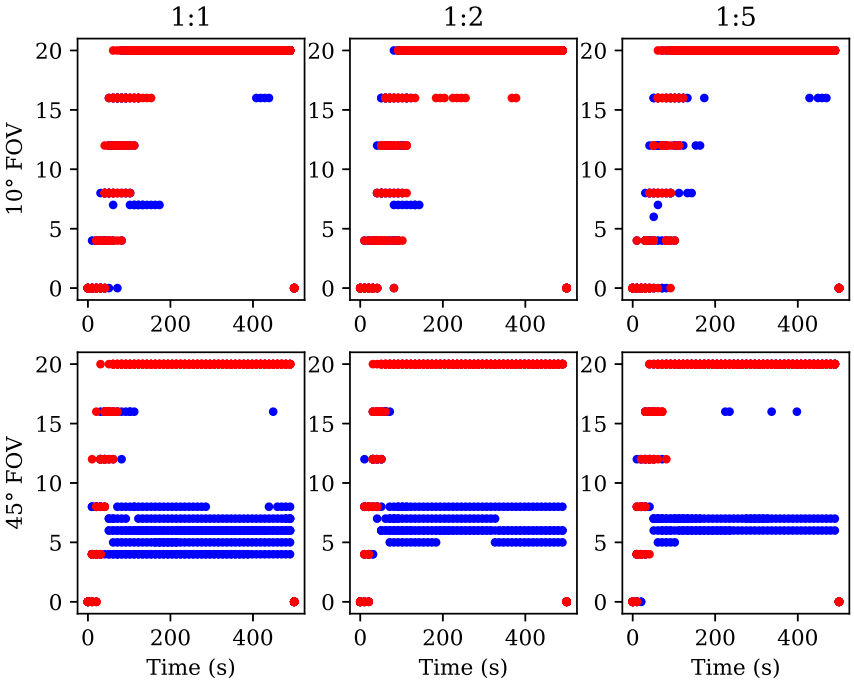


Figure 7.10 Number of edges in the communication graph for pruned and unpruned cases with angles-only measurements.

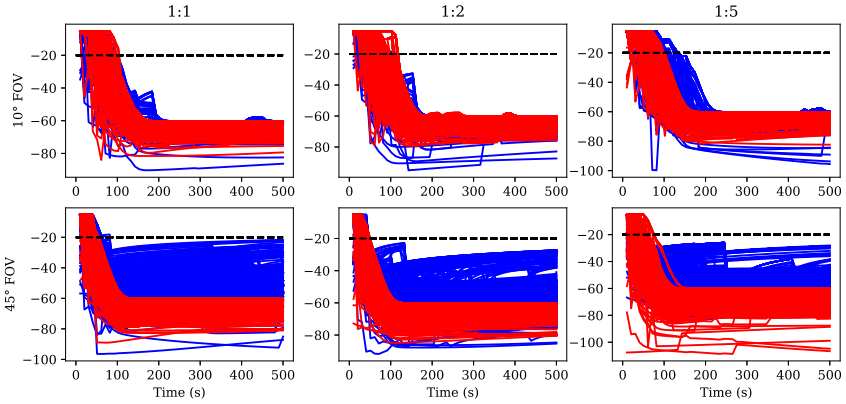


Figure 7.11 Shannon entropy for every agent's certainty of every other tracked object for pruned and unpruned communication graphs under the desired threshold for angles-only measurements.

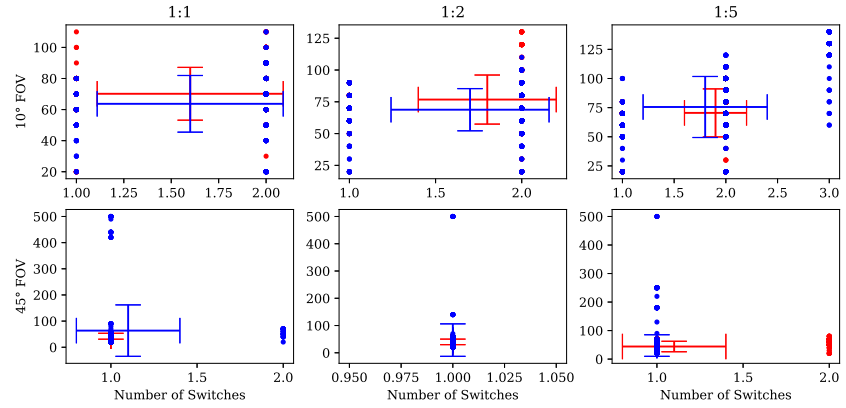


Figure 7.12 Metric time to desired threshold plotted against number of communication graph switches for pruned and unpruned cases with angles-only measurements.

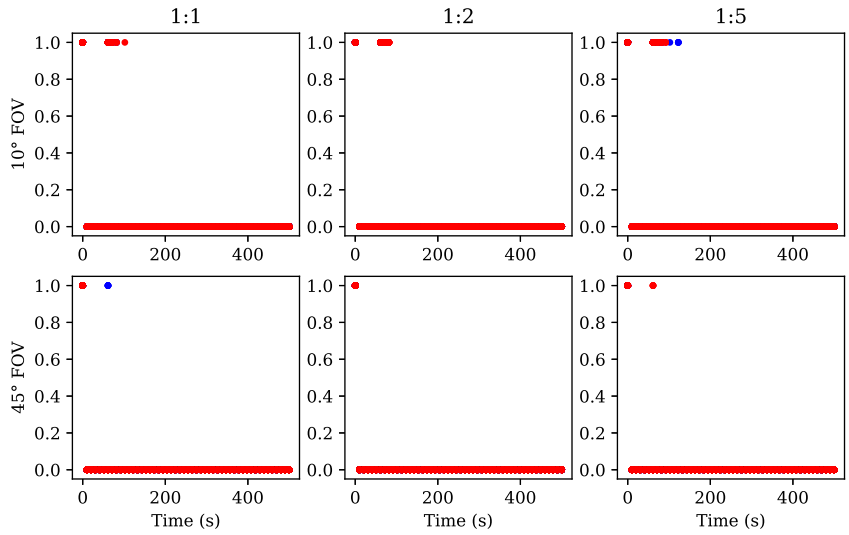


Figure 7.13 Switch history for pruned and unpruned communication graphs for angles-only measurements.

8 Conclusions and Recommendations

8.1 State Omniscience

Designing state omniscient distributed estimators has recently been of increasing interest has the applicability of multi-agent systems to practical problems becomes realizable. While the design of state omniscient distributed estimators has been of interest, there is a notable lack of generalizable results for the characterization of state omniscient systems. Most results place a primary focus on identifying necessary and sufficient conditions of the underlying graph while assuming certain characteristics about the observability at each agent. This work presents a collection of necessary and sufficient conditions for state omniscience of distributed estimators with no underlying assumptions on the graph structure, nor is the observability explicitly considered at each vertex. We draw explicit analogies to the well-known observability matrix, PBH test, and the Observability Gramian, so that the reader may have a clear point-of-reference for the presented material.

Additionally, the present work has developed necessary conditions for a converging distributed estimator under a variety of design cases. By utilizing set theory, a binary quantifier is introduced that serves as the objective to construct both sensing and communication networks to meet the convergence criteria. Unlike similar works, we make use of graph-based descriptions of the sensed information to easily construct a complementary communication network. Future work may be done to expand the theory of state omniscient to systems with linear time-varying or nonlinear dynamics, and time-varying graph topologies.

8.2 Local Catalog Maintenance

This work has presented both the local catalog maintenance problem and cooperative local catalog maintenance problem and posed a decision-making protocol in the form of a centralized supervisor. Multiple measurement modalities were employed such as, full-state, relative-position, and angles-only to consider the performance of the algorithms under a

variety of cases. Additionally, for the cooperative scenarios a centralized supervisor is presented. The centralized supervisor informs each agent of which target to track and constructs a communication graph through which the local state estimates may be passed. Each agent is equipped with a local fusion estimator and local controller to track its appointed target and fuse information coming in from its neighbors in the communication graph. Results show that the local controllers are capable of tracking the appointed targets when communication amongst agents is both disabled and enabled. When communication is enabled, the distributed estimator is able to maintain the desired certainty level on the local state estimate. However, limits to estimator performance are yet to be thoroughly explored. Consequently, performance analysis of the decision-making protocol is a subject of future work. Can a relationship be defined between the certainty threshold, control constraints, number of targets, and communication graph to determine whether a feasible solution exists? There is also an abundance of future work to be done holistically and for the supervisor, controller, and estimator individually. Potential avenues for future work include but are not limited to - including more sophisticated dynamics models, controllers, and estimators, employing a combinatorial approach to communication graph construction, and considering a decentralized supervisor.

8.3 Information Fusion

Chapter 5 presents two novel matrix decomposition schemes for fusing two pieces of information. Where related works use weighting schemes that naturally promote competition between the information pieces, this work attempts to approximate the mutual information through cooperative fusion using matrix square root approaches. Guarantees on upper and lower bounds of the fusion solution are presented. With the upper-bound shown to be the Inverse Covariance Intersection solution. Furthermore, calculation of the mutual mean using the proposed approaches is also presented. Further work is required to evaluate the performance of the fusion solution on a more generalized class of covariance matrices, such as covariance matrices that do not necessarily match the applied form. A

method of optimizing the value of $\mathbf{\Omega}$ within the specified bounds for a specified problem is sought and extensions to multiple information sources will be the subject of future work.

REFERENCES

[1] O'Connor, C. A., and et. al., "Economic Benefits of the Global Positioning System (GPS), Final Report," *RTI Report Number 0215471.*, 2019.

[2] Sgobba, T., and Allahdadi, F. A., "Chapter 8 - Orbital Operations Safety," *Safety Design for Space Operations*, edited by F. A. Allahdadi, I. Rongier, and P. D. Wilde, Butterworth-Heinemann, Oxford, 2013, pp. 411–602.

[3] Hobbs, K. L., Lyons, J. B., Feather, M. S., Bycroft, B. P., Phillips, S., Simon, M., Harter, M., Costello, K., Gawdiak, Y., and Paine, S., "Space Trusted Autonomy Readiness Levels," IEEE Aerospace Conference, 2023. https://doi.org/10.48550/ARXIV.2210.09059, URL https://arxiv.org/abs/2210.09059.

[4] Berger, J., and Lo, N., "An innovative multi-agent search-and-rescue path planning approach," *Computers & Operations Research*, Vol. 53, 2015, pp. 24–31.

[5] Tavakoli, M., Cabrita, G., Faria, R., Marques, L., and de Almeida, A. T., "Cooperative multi-agent mapping of three-dimensional structures for pipeline inspection applications," *The International Journal of Robotics Research*, Vol. 31, No. 12, 2012, pp. 1489–1503.

[6] Chevaleyre, Y., "Theoretical analysis of the multi-agent patrolling problem," *Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004.(IAT 2004).*, IEEE, 2004, pp. 302–308.

[7] Pfingsthorn, M., Slamet, B., and Visser, A., "A scalable hybrid multi-robot SLAM method for highly detailed maps," *Robot Soccer World Cup*, Springer, 2007, pp. 457–464.

[8] Bertsimas, D., and Sim, M., "The price of robustness," *Operations research*, Vol. 52, No. 1, 2004, pp. 35–53.

[9] Prorok, A., Malencia, M., Carlone, L., Sukhatme, G. S., Sadler, B. M., and Kumar, V., "Beyond Robustness: A Taxonomy of Approaches towards Resilient Multi-Robot Systems," *arXiv preprint arXiv:2109.12343*, 2021.

[10] Opromolla, R., Fasano, G., Rufino, G., and Grassi, M., "Pose estimation for spacecraft relative navigation using model-based algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 53, No. 1, 2017, pp. 431–447.

[11] Patel, H., Lovell, T. A., Allgeier, S., Russell, R., and Sinclair, A., "Relative navigation for satellites in close proximity using angles-only observations," *AAS/AIAA Space Flight Mechanics Meeting*, American Astronomical Soc. Washington, DC, 2012, pp. 12–202.

[12] Woffinden, D. C., and Geller, D. K., "Observability Criteria for Angles-Only Navigation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 45, No. 3, 2009, pp. 1194–1208. https://doi.org/10.1109/TAES.2009.5259193.

[13] Geller, D. K., and Klein, I., "Angles-only navigation state observability during orbital proximity operations," *Journal of Guidance, Control, and Dynamics*, Vol. 37, No. 6, 2014, pp. 1976–1983.

[14] Hu, Y., Sharf, I., and Chen, L., "Three-spacecraft autonomous orbit determination and observability analysis with inertial angles-only measurements," *Acta Astronautica*, Vol. 170, 2020, pp. 106–121.

[15] Psiaki, M. L., "Autonomous Orbit Determination for Two Spacecraft from Relative Position Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 22, No. 2, 1999, pp. 305–312.

[16] Erwin, R. S., Albuquerque, P., Jayaweera, S. K., and Hussein, I., "Dynamic sensor tasking for Space Situational Awareness," *Proceedings of the 2010 American Control Conference*, 2010, pp. 1153–1158.

[17] Schutz, B., Tapley, B., and Born, G. H., *Statistical orbit determination*, Elsevier, 2004.

[18] Sciotti, M., Besso, P., Flohrer, T., and Krag, H., "Low Earth Orbit objects tracking and orbit determination from ground-based phased array radar systems," *2011 12th International Radar Symposium (IRS)*, IEEE, 2011, pp. 591–596.

[19] Herz, A., and Logic, O., "SSA sensor tasking approach for improved orbit determination accuracies and more efficient use of ground assets," 2017.

[20] Jaunzemis, A. D., Holzinger, M. J., and Jah, M. K., "Evidence-based sensor tasking for space domain awareness," *Advanced Maui Optical and Space Surveillance Technologies Conference*, Maui Economic Development Board Maui, HI, 2016, p. 33.

[21] Endsley, M. R., "Toward a theory of situation awareness in dynamic systems," *Situational awareness*, Routledge, 2017, pp. 9–42.

[22] Holzinger, M., and Jah, M., "Challenges and potential in space domain awareness," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 1, 2018, pp. 15–18.

[23] Jia, B., Pham, K. D., Blasch, E., Shen, D., Wang, Z., and Chen, G., "Cooperative space object tracking using consensus-based filters," *17th International Conference on Information Fusion (FUSION)*, IEEE, 2014, pp. 1–8.

[24] Jia, B., Pham, K. D., Blasch, E., Shen, D., Wang, Z., and Chen, G., "Cooperative space object tracking using space-based optical sensors via consensus-based filters," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 52, No. 4, 2016, pp. 1908–1936.

[25] Hu, Y., Sharf, I., and Chen, L., "Distributed orbit determination and observability analysis for satellite constellations with angles-only measurements," *Automatica*, Vol. 129, 2021, p. 109626.

[26] Hays, C. W., Miller, K., Soderlund, A., Phillips, S., and Henderson, T., "Autonomous Local Catalog Maintenance of Close Proximity Satellite Systems on Closed Natural Motion Trajectories," *45th Rocky Mountain American Astronomical Society Guidance, Navigation & Control Conference*, AAS, 2023.

[27] Hays, C. W., Henderson, T., Phillips, S., and Soderlund, A., "Enabling Proliferated Space Sensor Awareness Constellations Through Topological Observability," *2023 AIAA SciTech Forum*, AIAA, 2023.

[28] Hays, C. W., Miller, K., Soderlund, A. A., Phillips, S., and Henderson, T., "State Omniscience for Cooperative Local Catalog Maintenance of Close Proximity Satellite Systems," *2024 AIAA SCITECH Forum*, 2024, p. 0992. https://doi.org/https://doi.org/10.2514/6.2024-0992.

[29] Dai, P., Yu, W., Wang, H., and Baldi, S., "Distributed actor-critic algorithms for multiagent reinforcement learning over directed graphs," *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 34, No. 10, 2022. https://doi.org/10.1109/TNNLS.2021.3139138.

[30] Lin, Y., Zhang, K., Yang, Z., Wang, Z., Başar, T., Sandhu, R., and Liu, J., "A communication-efficient multi-agent actor-critic algorithm for distributed reinforcement learning," *2019 IEEE 58th Conference on Decision and Control (CDC)*, IEEE, 2019, pp. 5562–5567. https://doi.org/10.1109/CDC40024.2019.9029257.

[31] Choudhury, S., Gupta, J. K., Morales, P., and Kochenderfer, M. J., "Scalable Online Planning for Multi-Agent MDPs," *Journal of Artificial Intelligence Research*, Vol. 73, 2022, pp. 821–846. https://doi.org/10.1613/jair.1.13261.

[32] Fedeler, S. J., Holzinger, M. J., and Whitacre, W. W., "Decentralized Decision Making over Random Graphs," *2023 26th International Conference on Information*

Fusion (FUSION), IEEE, 2023, pp. 1–8. https://doi.org/10.23919/FUSION52260.2023.10224100.

[33] Fedeler, S., Holzinger, M., and Whitacre, W., "Decentralized decision making over random graphs for space domain awareness," *Advances in Space Research*, Vol. In Press, 2024, p. early access. https://doi.org/10.1016/j.asr.2024.01.063.

[34] Kalman, R. E., "On the general theory of control systems," *Proceedings First International Conference on Automatic Control, Moscow, USSR*, 1960, pp. 481–492.

[35] Lin, C.-T., "Structural controllability," *IEEE Transactions on Automatic Control*, Vol. 19, No. 3, 1974, pp. 201–208.

[36] Shields, R., and Pearson, J., "Structural controllability of multiinput linear systems," *IEEE Transactions on Automatic control*, Vol. 21, No. 2, 1976, pp. 203–212.

[37] Godsil, C., "Controllable subsets in graphs," *Annals of Combinatorics*, Vol. 16, No. 4, 2012, pp. 733–744.

[38] Liu, Y.-Y., and Barabási, A.-L., "Control principles of complex systems," *Reviews of Modern Physics*, Vol. 88, No. 3, 2016, p. 035006.

[39] Yuan, Z., Zhao, C., Di, Z., Wang, W.-X., and Lai, Y.-C., "Exact controllability of complex networks," *Nature communications*, Vol. 4, No. 1, 2013, pp. 1–9.

[40] Aguilar, C. O., and Gharesifard, B., "Graph controllability classes for the Laplacian leader-follower dynamics," *IEEE transactions on automatic control*, Vol. 60, No. 6, 2014, pp. 1611–1623.

[41] Rahmani, A., Ji, M., Mesbahi, M., and Egerstedt, M., "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, Vol. 48, No. 1, 2009, pp. 162–186.

[42] Ji, Z., and Yu, H., "A new perspective to graphical characterization of multiagent controllability," *IEEE transactions on cybernetics*, Vol. 47, No. 6, 2016, pp. 1471–1483.

[43] Godsil, C. D., and McKay, B. D., "Spectral conditions for the reconstructibility of a graph," *Journal of Combinatorial Theory, Series B*, Vol. 30, No. 3, 1981, pp. 285–289.

[44] Godsil, C., and Royle, G. F., *Algebraic graph theory*, Vol. 207, Springer Science & Business Media, 2001.

[45] Doostmohammadian, M., and Khan, U. A., "On the genericity properties in networked estimation: Topology design and sensor placement," *arXiv preprint arXiv:1208.3691*, 2012.

[46] Rodríguez, S., De Paz, J. F., Villarrubia, G., Zato, C., Bajo, J., and Corchado, J. M., "Multi-agent Information Fusion System to Manage Data from a WSN in a Residential Home," *Information Fusion*, Vol. 23, 2015, pp. 43–57.

[47] Mirabadi, A., Mort, N., and Schmid, F., "Application of Sensor Fusion to Railway Systems," *1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems (Cat. No. 96TH8242)*, IEEE, 1996, pp. 185–192.

[48] Silva, B., Fisher, R. M., Kumar, A., and Hancke, G. P., "Experimental link quality characterization of wireless sensor networks for underground monitoring," *IEEE Transactions on Industrial Informatics*, Vol. 11, No. 5, 2015, pp. 1099–1110.

[49] Liu, W., Wei, J., Liang, M., Cao, Y., and Hwang, I., "Multi-sensor Fusion and Fault Detection Using Hybrid Estimation for Air Traffic Surveillance," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 49, No. 4, 2013, pp. 2323–2339.

[50] Li, C., and Qu, Z., "Distributed estimation of algebraic connectivity of directed networks," *Systems & Control Letters*, Vol. 62, No. 6, 2013, pp. 517–524.

[51] Vu, T. T., and Rahmani, A. R., "Distributed consensus-based Kalman filter estimation and control of formation flying spacecraft: Simulation and validation," *AIAA guidance, navigation, and control conference*, 2015, p. 1553.

[52] Phillips, S., Erwin, R. S., and Sanfelice, R. G., "Robust Exponential Stability of an Intermittent Transmission State Estimation Protocol," *2018 Annual American Control Conference (ACC)*, 2018, pp. 622–627. https://doi.org/10.23919/ACC.2018.8431144.

[53] He, S., Shin, H.-S., Xu, S., and Tsourdos, A., "Distributed estimation over a low-cost sensor network: A review of state-of-the-art," *Information Fusion*, Vol. 54, 2020, pp. 21–43.

[54] Bar-Shalom, Y., and Campo, L., "The effect of the common process noise on the two-sensor fused-track covariance," *IEEE Transactions on aerospace and electronic systems*, , No. 6, 1986, pp. 803–805.

[55] Kim, K. H., "Development of track to track fusion algorithms," *Proceedings of 1994 American Control Conference-ACC'94*, Vol. 1, IEEE, 1994, pp. 1037–1041.

[56] Mori, S., Chang, K.-C., and Chong, C.-Y., "Comparison of track fusion rules and track association metrics," *2012 15th International Conference on Information Fusion*, IEEE, 2012, pp. 1996–2003.

[57] Olfati-Saber, R., "Distributed Kalman filter with embedded consensus filters," *Proceedings of the 44th IEEE Conference on Decision and Control*, IEEE, 2005, pp. 8179–8184.

[58] Olfati-Saber, R., and Murray, R. M., "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on automatic control*, Vol. 49, No. 9, 2004, pp. 1520–1533.

[59] Olfati-Saber, R., "Distributed Kalman filtering for sensor networks," *2007 46th IEEE Conference on Decision and Control*, IEEE, 2007, pp. 5492–5498.

[60] Olfati-Saber, R., "Kalman-consensus filter: Optimality, stability, and performance," *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, Ieee, 2009, pp. 7036–7042.

[61] Ma, K., Wu, S., Wei, Y., and Zhang, W., "Gossip-based distributed tracking in networks of heterogeneous agents," *IEEE Communications Letters*, Vol. 21, No. 4, 2016, pp. 801–804.

[62] Boyd, S., Ghosh, A., Prabhakar, B., and Shah, D., "Randomized gossip algorithms," *IEEE transactions on information theory*, Vol. 52, No. 6, 2006, pp. 2508–2530.

[63] Cattivelli, F. S., and Sayed, A. H., "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on automatic control*, Vol. 55, No. 9, 2010, pp. 2069–2084.

[64] Cattivelli, F., and Sayed, A. H., "Diffusion distributed Kalman filtering with adaptive weights," *2009 conference record of the forty-third asilomar conference on signals, systems and computers*, IEEE, 2009, pp. 908–912.

[65] Hu, J., Xie, L., and Zhang, C., "Diffusion Kalman filtering based on covariance intersection," *IEEE Transactions on Signal Processing*, Vol. 60, No. 2, 2011, pp. 891–902.

[66] Phillips, S., Petersen, C., and Fierro, R., *Robust, Resilient, and Energy-Efficient Satellite Formation Control*, Springer International Publishing, 2022.

[67] Chen, G., Zhang, Y., Gu, S., and Hu, W., "Resilient state estimation and control of cyber-physical systems against false data injection attacks on both actuator and sensors," *IEEE Transactions on Control of Network Systems*, Vol. 9, No. 1, 2021, pp. 500–510.

[68] Byrne, R. H., Nguyen, T. A., Copp, D. A., Chalamala, B. R., and Gyuk, I., "Energy management and optimization methods for grid energy storage systems," *IEEE Access*, Vol. 6, 2017, pp. 13231–13260.

[69] Nugroho, S. A., Taha, A. F., Gatsis, N., and Zhao, J., "Observers for differential algebraic equation models of power networks: Jointly estimating dynamic and algebraic states," *IEEE transactions on control of network systems*, Vol. 9, No. 3, 2022, pp. 1531–1543.

[70] Mutambara, A. G., *Decentralized estimation and control for multisensor systems*, CRC press, 1998.

[71] Smith, R. S., and Hadaegh, F. Y., "Closed-loop dynamics of cooperative vehicle formations with parallel estimators and communication," *IEEE Transactions on Automatic Control*, Vol. 52, No. 8, 2007, pp. 1404–1414.

[72] Battistelli, G., Chisci, L., Morrocchi, S., and Papi, F., "An information-theoretic approach to distributed state estimation," *IFAC Proceedings Volumes*, Vol. 44, No. 1, 2011, pp. 12477–12482.

[73] Khan, U. A., and Moura, J. M., "Distributing the Kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, Vol. 56, No. 10, 2008, pp. 4919–4935.

[74] Mitra, A., and Sundaram, S., "An approach for distributed state estimation of LTI systems," *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2016, pp. 1088–1093.

[75] Kalman, R. E., and Bucy, R. S., "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, Vol. 83, No. 1, 1961, pp. 95–108.

[76] Doostmohammadian, M., and Khan, U. A., "On the genericity properties in distributed estimation: Topology design and sensor placement," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 7, No. 2, 2013, pp. 195–204. https://doi.org/10.1109/JSTSP.2013.2246135.

[77] Doostmohammadian, M., and Khan, U. A., "Communication strategies to ensure generic networked observability in multi-agent systems," *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, IEEE, 2011, pp. 1865–1868.

[78] Zelazo, D., and Mesbahi, M., "On the observability properties of homogeneous and heterogeneous networked dynamic systems," *2008 47th IEEE conference on decision and control*, IEEE, 2008, pp. 2997–3002.

[79] Li, Y., Phillips, S., and Sanfelice, R. G., "Robust Distributed Estimation for Linear Systems Under Intermittent Information," *IEEE Transactions on Automatic Control*, Vol. 63, No. 4, 2018, p. 973–988.

[80] Khan, U. A., and Jadbabaie, A., "Coordinated networked estimation strategies using structured systems theory," *2011 50th IEEE Conference on Decision and Control and European Control Conference*, IEEE, 2011, pp. 2112–2117.

[81] Rego, F. F., Pascoal, A. M., Aguiar, A. P., and Jones, C. N., "Distributed state estimation for discrete-time linear time invariant systems: A survey," *Annual Reviews in Control*, Vol. 48, 2019, pp. 36–56.

[82] Park, S., and Martins, N. C., "Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers," *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, IEEE, 2012, pp. 7431–7436.

[83] Park, S., and Martins, N. C., "Design of distributed LTI observers for state omniscience," *IEEE Transactions on Automatic Control*, Vol. 62, No. 2, 2016, pp. 561–576.

[84] Li, Y., Phillips, S., and Sanfelice, R. G., "Robust distributed estimation for linear systems under intermittent information," *IEEE Transactions on Automatic Control*, Vol. 63, No. 4, 2017, pp. 973–988.

[85] Wang, L., and Morse, A. S., "A distributed observer for a time-invariant linear system," *IEEE Transactions on Automatic Control*, Vol. 63, No. 7, 2017, pp. 2123–2130.

[86] Han, W., Trentelman, H. L., Wang, Z., and Shen, Y., "A simple approach to distributed observer design for linear systems," *IEEE Transactions on Automatic Control*, Vol. 64, No. 1, 2018, pp. 329–336.

[87] Kim, T., Lee, C., and Shim, H., "Completely decentralized design of distributed observer for linear systems," *IEEE Transactions on Automatic Control*, Vol. 65, No. 11, 2019, pp. 4664–4678.

[88] Phillips, S., and Sanfelice, R. G., "Observer-based synchronization of multi-agent systems using intermittent measurements," *Proceedings of the IEEE Conference on Decision and Control*, 2019.

[89] Park, S., and Martins, N. C., "An augmented observer for the distributed estimation problem for LTI systems," *2012 American Control Conference (ACC)*, IEEE, 2012, pp. 6775–6780.

[90] Khan, U. A., and Jadbabaie, A., "Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication," *Automatica*, Vol. 50, No. 7, 2014, pp. 1909–1914.

[91] Mitra, A., and Sundaram, S., "Distributed observers for LTI systems," *IEEE Transactions on Automatic Control*, Vol. 63, No. 11, 2018, pp. 3689–3704.

[92] Li, Y., Phillips, S., and Sanfelice, R. G., "On distributed observers for linear time-invariant systems under intermittent information constraints," *IFAC-PapersOnLine*, Vol. 49, No. 18, 2016, pp. 654–659.

[93] Phillips, S., Li, Y., and Sanfelice, R. G., "A hybrid consensus protocol for pointwise exponential stability with intermittent information," *IFAC-PapersOnLine*, Vol. 49, No. 18, 2016, pp. 146–151.

[94] Zegers, F. M., Phillips, S., and Dixon, W. E., "Consensus over clustered networks with asynchronous inter-cluster communication," *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 4249–4254.

[95] Parlangeli, G., and Notarstefano, G., "Graph reduction based observability conditions for network systems running average consensus algorithms," *18th Mediterranean Conference on Control and Automation, MED'10*, IEEE, 2010, pp. 689–694.

[96] Tuna, S. E., "Observability through a matrix-weighted graph," *IEEE Transactions on Automatic Control*, Vol. 63, No. 7, 2017, pp. 2061–2074.

[97] Montanari, A. N., and Aguirre, L. A., "Observability of network systems: A critical review of recent results," *Journal of Control, Automation and Electrical Systems*, Vol. 31, No. 6, 2020, pp. 1348–1374.

[98] Tanner, H. G., "On the controllability of nearest neighbor interconnections," *2004 43rd IEEE conference on decision and control (CDC)(IEEE Cat. No. 04CH37601)*, Vol. 3, IEEE, 2004, pp. 2467–2472.

[99] Boukhobza, T., Hamelin, F., Martinez-Martinez, S., and Sauter, D., "Structural analysis of the partial state and input observability for structured linear systems:

Application to distributed systems," *European Journal of Control*, Vol. 15, No. 5, 2009, pp. 503–516.

[100] Aguirre, L. A., Portes, L. L., and Letellier, C., "Structural, dynamical and symbolic observability: From dynamical systems to networks," *PLoS One*, Vol. 13, No. 10, 2018, p. e0206180.

[101] Gutiérrez, R., Ramperez, V., Paggi, H., Lara, J. A., and Soriano, J., "On the Use of Information Fusion Techniques to Improve Information Quality: Taxonomy, Opportunities and Challenges," *Information Fusion*, Vol. 78, 2022, pp. 102–137.

[102] Grime, S., and Durrant-Whyte, H. F., "Data Fusion in Decentralized Sensor Networks," *Control Engineering Practice*, Vol. 2, No. 5, 1994, pp. 849–863.

[103] Boström, H., Andler, S. F., Brohede, M., Johansson, R., Karlsson, A., Van Laere, J., Niklasson, L., Nilsson, M., Persson, A., and Ziemke, T., "On the Definition of Information Fusion as a Field of Research," *Institutionen för Kommunikation och Information*, 2007.

[104] Shen, Y., Wang, Z., Dong, H., and Liu, H., "Multi-sensor Multi-rate Fusion Estimation for Networked Systems: Advances and Perspectives," *Information Fusion*, 2022, pp. 19–27.

[105] Sun, S., Lin, H., Ma, J., and Li, X., "Multi-sensor Distributed Fusion Estimation with Applications in Networked Systems: A Review Paper," *Information Fusion*, Vol. 38, 2017, pp. 122–134.

[106] Julier, S. J., and Uhlmann, J. K., "A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations," *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, Vol. 4, IEEE, 1997, pp. 2369–2373.

[107] Daum, F., "Handbook of Multisensor Data Fusion [Book Review]," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 16, No. 10, 2001, pp. 15–16.

[108] Sijs, J., Lazar, M., and Bosch, P., "State Fusion with Unknown Correlation: Ellipsoidal Intersection," *Proceedings of the 2010 American Control Conference*, IEEE, 2010, pp. 3992–3997.

[109] Noack, B., Sijs, J., Reinhardt, M., and Hanebeck, U. D., "Decentralized Data Fusion With Inverse Covariance Intersection," *Automatica*, Vol. 79, 2017, pp. 35–41.

[110] Noack, B., Sijs, J., and Hanebeck, U. D., "Inverse Covariance Intersection: New Insights and Properties," *2017 20th International Conference on Information Fusion (Fusion)*, IEEE, 2017, pp. 1–8.

[111] Ajgl, J., and Straka, O., "On Weak Points of the Ellipsoidal Intersection Fusion," *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, IEEE, 2017, pp. 28–33.

[112] Hanebeck, U. D., and Briechle, K., "New Results for Stochastic Prediction and Filtering with Unknown Correlations," *Conference Documentation International Conference on Multisensor Fusion and Integration for Intelligent Systems. MFI 2001 (Cat. No. 01TH8590)*, IEEE, 2001, pp. 147–152.

[113] Amin, A. A., and Hasan, K. M., "A review of fault tolerant control systems: advancements and applications," *Measurement*, Vol. 143, 2019, pp. 58–68.

[114] Hays, C. W., and Henderson, T., "Matrix decomposition approaches for mutual information approximation with applications to covariance intersection techniques," *Information Fusion*, Vol. 95, 2023, pp. 446–453.

[115] Biggs, N., Biggs, N. L., and Norman, B., *Algebraic graph theory*, 67, Cambridge university press, 1993.

[116] West, D. B., et al., *Introduction to graph theory*, Vol. 2, Prentice hall Upper Saddle River, 2001.

[117] Xiang, L., Wang, P., Chen, F., and Chen, G., "Controllability of directed networked MIMO systems with heterogeneous dynamics," *IEEE Transactions on Control of Network Systems*, Vol. 7, No. 2, 2019, pp. 807–817.

[118] Pecora, L. M., and Carroll, T. L., "Master stability functions for synchronized coupled systems," *Physical review letters*, Vol. 80, No. 10, 1998, p. 2109.

[119] Griffith, E. W., and Kumar, K., "On the observability of nonlinear systems: I," *Journal of Mathematical Analysis and Applications*, Vol. 35, No. 1, 1971, pp. 135–147.

[120] Kou, S. R., Elliott, D. L., and Tarn, T. J., "Observability of nonlinear systems," *Information and Control*, Vol. 22, No. 1, 1973, pp. 89–99.

[121] Hermann, R., and Krener, A., "Nonlinear controllability and observability," *IEEE Transactions on automatic control*, Vol. 22, No. 5, 1977, pp. 728–740.

[122] Liu, Y., and Zhang, W., "Necessary and sufficient condition for global controllability of nonlinear systems," *Asian Journal of Control*, 2022.

[123] Lukes, D., "Global controllability of nonlinear systems," *SIAM journal on Control*, Vol. 10, No. 1, 1972, pp. 112–126.

[124] Friedland, B., "Controllability index based on conditioning number," *Journal of Dynamic Systems, Measurement and Control*, Vol. 97, No. 4, 1975, pp. 444–445.

[125] Aguirre, L. A., "Controllability and observability of linear systems: some noninvariant aspects," *IEEE Transactions on Education*, Vol. 38, No. 1, 1995, pp. 33–39.

[126] Liu, F., and Morse, A. S., "A graphical characterization of structurally controllable linear systems with dependent parameters," *IEEE Transactions on Automatic Control*, Vol. 64, No. 11, 2019, pp. 4484–4495.

[127] Jantzen, J., "Tutorial on the digraph approach to multivariable control," *IFAC Proceedings Volumes*, Vol. 26, No. 2, 1993, pp. 711–714.

[128] Clohessy, W., and Wiltshire, R., "Terminal guidance system for satellite rendezvous," *Journal of the Aerospace Sciences*, Vol. 27, No. 9, 1960, pp. 653–658.

[129] Hill, G. W., "Researches in the lunar theory," *American journal of Mathematics*, Vol. 1, No. 1, 1878, pp. 5–26.

[130] Lei, H. H., Shubert, M., Damron, N., Lang, K., and Phillips, S., "Deep reinforcement Learning for Multi-agent Autonomous Satellite Inspection," *In Proceedings of AAS Guidance Navigation and Control Conference*, 2022.

[131] Petersen, C. D., Phillips, S., Hobbs, K. L., and Lang, K., "Challenge Problem: Assured Satellite Proximity Operations," *In Proceedings of AAS/AIAA Space Flight Mechanics Meeting*, 2021.

[132] Shibata, M., and Ichikawa, A., "Orbital rendezvous and flyaround based on null controllability with vanishing energy," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 4, 2007, pp. 934–945.

[133] Mote, M. L., Hays, C. W., Collins, A., Feron, E., and Hobbs, K. L., "Natural motion-based trajectories for automatic spacecraft collision avoidance during proximity operations," *2021 IEEE Aerospace Conference (50100)*, IEEE, 2021, pp. 1–12.

[134] Frey, G. R., Petersen, C. D., Leve, F. A., Kolmanovsky, I. V., and Girard, A. R., "Incorporating periodic and non-periodic natural motion trajectories into constrained invariance-based spacecraft relative motion planning," *2017 IEEE Conference on Control Technology and Applications (CCTA)*, IEEE, 2017, pp. 1811–1816.

[135] Hobbs, K. L., and Feron, E. M., "A Taxonomy for Aerospace Collision Avoidance with Implications for Automation in Space Traffic Management," *AIAA Scitech 2020 Forum*, 2020, p. 0877.

[136] Soderlund, A. A., and Phillips, S., "Autonomous Rendezvous and Proximity Operations of an Underactuated Spacecraft via Switching Controls," *AIAA SCITECH 2022 Forum*, 2022, p. 0956.

[137] Kim, S. C., Shepperd, S. W., Norris, H. L., Goldberg, H. R., and Wallace, M. S., "Mission design and trajectory analysis for inspection of a host spacecraft by a microsatellite," *2007 IEEE Aerospace Conference*, IEEE, 2007, pp. 1–23.

[138] Frey, G. R., Petersen, C. D., Leve, F. A., Kolmanovsky, I. V., and Girard, A. R., "Constrained spacecraft relative motion planning exploiting periodic natural motion trajectories and invariance," *Journal of Guidance, Control, and Dynamics*, Vol. 40, No. 12, 2017, pp. 3100–3115.

[139] Uhlmann, J. K., "Covariance Consistency Methods for Fault-Tolerant Distributed Data Fusion," *Information Fusion*, Vol. 4, No. 3, 2003, pp. 201–215.

[140] Ajgl, J., and Straka, O., "Lower Bounds In Estimation Fusion With Partial Knowledge of Correlations," *2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, IEEE, 2021, pp. 1–6.

[141] Reinhardt, M., Kulkarni, S., and Hanebeck, U. D., "Generalized Covariance Intersection Based on Noise Decomposition," *2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems (MFI)*, IEEE, 2014, pp. 1–8.

[142] Ajgl, J., and Straka, O., "Comparison of Fusions Under Unknown and Partially Known Correlations," *IFAC-PapersOnLine*, Vol. 51, No. 23, 2018, pp. 295–300.

[143] Ajgl, J., and Straka, O., "On Fusion of Partial Estimates Under Implicit Partial Knowledge of Correlation," *2019 22th International Conference on Information Fusion (FUSION)*, IEEE, 2019, pp. 1–8.

[144] Forsling, R., Hansson, A., Gustafsson, F., Sjanic, Z., Löfberg, J., and Hendeby, G., "Conservative Linear Unbiased Estimation Under Partially Known Covariances," *IEEE Transactions on Signal Processing*, Vol. 70, 2022, pp. 3123–3135.

[145] Wu, Z., Cai, Q., and Fu, M., "Covariance Intersection for Partially Correlated Random Vectors," *IEEE Transactions on Automatic Control*, Vol. 63, No. 3, 2017, pp. 619–629.

[146] Julier, S. J., "Estimating and Exploiting the Degree of Independent Information in Distributed Data Fusion," *2009 12th International Conference on Information Fusion*, IEEE, 2009, pp. 772–779.

[147] Uhlmann, J., "Dynamic Map Building and Localization: New Theoretical Foundations," Ph.D. thesis, University of Oxford, 1995.

[148] Julier, S. J., "Fusion Without Independence," *IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications*, 2008.

[149] Crassidis, J. L., and Junkins, J. L., *Optimal estimation of dynamic systems*, Vol. 2, Chapman & Hall/CRC Boca Raton, FL, 2004.

[150] Ly, A., Marsman, M., Verhagen, J., Grasman, R. P., and Wagenmakers, E.-J., "A tutorial on Fisher information," *Journal of Mathematical Psychology*, Vol. 80, 2017, pp. 40–55.

[151] Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., and Diehl, M., "CasADi – A software framework for nonlinear optimization and optimal control," *Mathematical*

Programming Computation, Vol. 11, No. 1, 2019, pp. 1–36. https://doi.org/10.1007/s12532-018-0139-4.

[152] Khan, U. A., and Jadbabaie, A., "On the stability and optimality of distributed Kalman filters with finite-time data fusion," *Proceedings of the 2011 American control conference*, IEEE, 2011, pp. 3405–3410.

[153] Belevitch, V., "Classical network theory," *Holden-day*, Vol. 7, 1968.

[154] Popov, V.-M., and Georgescu, R., *Hyperstability of control systems*, Springer, 1973.

[155] Parlangeli, G., and Notarstefano, G., "On the observability of path and cycle graphs," *49th IEEE Conference on Decision and Control (CDC)*, IEEE, 2010, pp. 1492–1497.

[156] Cvetković, D. M., Rowlinson, P., and Simić, S., *An introduction to the theory of graph spectra*, Vol. 75, Cambridge University Press Cambridge, 2010.

[157] Maybee, J. S., Olesky, D., and Van Den Driessche, P., "Partly zero eigenvectors," *Linear and Multilinear Algebra*, Vol. 28, No. 1-2, 1990, pp. 83–92.

[158] Ugrinovskii, V., "Conditions for detectability in distributed consensus-based observer networks," *IEEE Transactions on Automatic Control*, Vol. 58, No. 10, 2013, pp. 2659–2664.

[159] Hou, M., Zítek, P., and Patton, R. J., "An observer design for linear time-delay systems," *IEEE Transactions on Automatic Control*, Vol. 47, No. 1, 2002, pp. 121–125.

[160] Hohn, F. E., *Elementary matrix algebra*, Courier Corporation, 2013.

[161] Heath, M., Laub, A., Paige, C., and Ward, R., "Computing the singular value decomposition of a product of two matrices," *SIAM Journal on Scientific and Statistical Computing*, Vol. 7, No. 4, 1986, pp. 1147–1159.

[162] Kahan, W., "Circumscribing an Ellipsoid About the Intersection of Two Ellipsoids," *Canadian Mathematical Bulletin*, Vol. 11, No. 3, 1968, pp. 437–441.

[163] Björck, Å., and Hammarling, S., "A Schur Method for the Square Root of a Matrix," *Linear algebra and its applications*, Vol. 52, 1983, pp. 127–140.

[164] Deadman, E., Higham, N. J., and Ralha, R., "Blocked Schur Algorithms for Computing the Matrix Square Root," *International Workshop on Applied Parallel Computing*, Springer, 2012, pp. 171–182.

[165] Rhudy, M., Gu, Y., Gross, J., and Napolitano, M. R., "Evaluation of Matrix Square Root Operations for UKF Within a UAV GPS/INS Sensor Fusion Application," *International Journal of Navigation and Observation*, Vol. 2011, 2011.

[166] Chakravarti, I.-M., "On a characterization of irreducibility of a non-negative matrix," *Linear Algebra and Its Applications*, Vol. 10, No. 2, 1975, pp. 103–109.

PUBLICATIONS

Dissertation

Hays, C.W., Phillips, S., Henderson, T., "Necessary and Sufficient Conditions for State Omniscience of Linear Time-Invariant Distributed Estimators," IEEE Transactions on Automatic Control. [In Preparation].

Hays, C.W., Phillips, S., Henderson, T., "A Necessary and Sufficient Condition for State Omniscience of Distributed Estimators," Proceedings of the 2024 American Control Conference. July 8-12, 2024 [To Appear].

Hays, C.W., Henderson, T., Miller, K., Phillips, S., Soderlund, A., "State Omniscience for Cooperative Local Catalog Maintenance of Close Proximity Satellite Systems," AIAA Journal of Guidance, Control, and Dynamics Virtual Collection on Distributed Space Systems. [Under Review].

Hays, C.W., Phillips, S., Henderson, T., "Cooperative Estimator Convergence: Ubiquitous Observability, Ubiquitous Detectability, and State Omniscience," Transactions on Control of Network Systems. [Under Review].

Hays, C.W., Henderson, T., Miller, K., Phillips, S., Soderlund, A., "State Omniscience for Cooperative Local Catalog Maintenance of Close Proximity Satellite Systems," Proceedings of the 2024 AIAA SciTech Forum. January 8-12, 2024..

Hays, C.W., Soderlund, A., Phillips, S., Henderson, T., "Ubiquitous Controllability of Single Input Linear Time-Invariant Systems," Proceedings of the 2023 American Control Conference. May 31 - June 2, 2023.

Hays, C.W., Henderson, T., "Matrix Decomposition Approaches for Mutual Information Approximation with Applications to Covariance Intersection Techniques," Information Fusion. Vol. 95 pgs. 446-453.

Hays, C.W., Miller, K., Soderlund, A., Phillips, S., Henderson, T., "Local Catalog Maintenance of Close Proximity Satellite Systems on Closed Natural Motion Trajectories," Proceedings of the 45^{th} Rocky Mountain AAS GN&C Conference. February 2-8, 2023.

Hays, C.W., Henderson, T., Phillips, S., Soderlund, A., "Enabling Proliferated Space Sensor Awareness Constellations Through Topological Observability," Proceedings of the 2023 AIAA SciTech Forum. January 23-27, 2023.

EagleCam

Hays, C.W., Posada, D., Malik, A., Korczyk, D., Dafoe, B., Henderson, T., "Structural Design and Impact Analysis of a 1.5U Cubesat on the Lunar Surface," . .

Sahr, J., Posada, D., Miguélez-Gómez, N., Korczyk, D., Pepin, K., Parkhurst, J., **Hays, C.W.**, Henderson, T., Rojas-Nastrucci, E., "Wireless Payload Thermal-Vacuum Testing for Lunar Harsh Environment," 2021 IEEE Space Hardware and Radio Conference (SHaRC). .

Posada, D., **Hays, C.W.**, Jordan, J., Lopez, D., Yow, T., Malik, A., Henderson, T., "EagleCam: A 1.5u Low-Cost CubeSat Mission for a Novel Third-Person View of a Lunar Landing," 2023 IEEE Aerospace Conference. .

Other

Yow, T., **Hays, C.W.**, Malik, A., Henderson, T., "Orbital Acceleration Using Product of Exponentials," . .

Zuehlke, D., Yow, T., Posada, D., Nicolich, J., **Hays, C.W.**, Malik, A., Henderson, T., "Initial Orbit Determination for the CR3BP Using Particle Swarm Optimization," . .

Hays, C.W., Henderson, T., "Evaluations on Uncertainty Characterization of an SE(3) Variational Filter," 2022 American Control Conference. .

Hays, C.W., Tiwari, M., Henderson, T., Prazenica, R., "A Geometric Mechanics Approach to Direct Adaptive Model Predictive Control," 2022 AIAA SciTech Forum. .

Malik, A., Canales, D., Yow, T., Posada, D., **Hays, C.W.**, Zuehlke, D., Henderson, T., "Computation of Relative Orbital Motion Using Product of Exponentials Mapping," . .

Mote, M.L., **Hays, C.W.**, Collins, A., Feron, E., Hobbs, K., "Natural Motion-based Trajectories for Automatic Spacecraft Collision Avoidance During Proximity Operations," 2021 IEEE Aerospace Conference. .