博士論文

Relative Navigation for Spacecraft Swarms based on Observability-Driven Adaptive Decentralization of Estimation Algorithm

（観測性評価に基づき推定アルゴリズムの分散度調整を行う宇宙機スワームの相対航法に関する研究）

船曳敦漠

Nobuhiro Funabiki

Department of Aeronautics and Astronautics
School of Engineering
Doctor of Philosophy
at
The University of Tokyo

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Spacecraft Swarm Applications

1.1.1 Constellations, Formation Flying, and Spacecraft Swarms

Multi-spacecraft systems have the potential to achieve a variety of missions that cannot be realized by monolithic spacecraft [11]. Multi-spacecraft systems can be broadly divided into constellation and formation flying systems, as conceptually illustrated in Figure 1.1. [12] In the constellation systems, each spacecraft only takes care of its own state in terms of navigation and control. However, it provides specific services with the other spacecraft in the system. On the other hand, each spacecraft is managed with the interaction with the dynamic states of the other spacecraft in the formation flying systems.

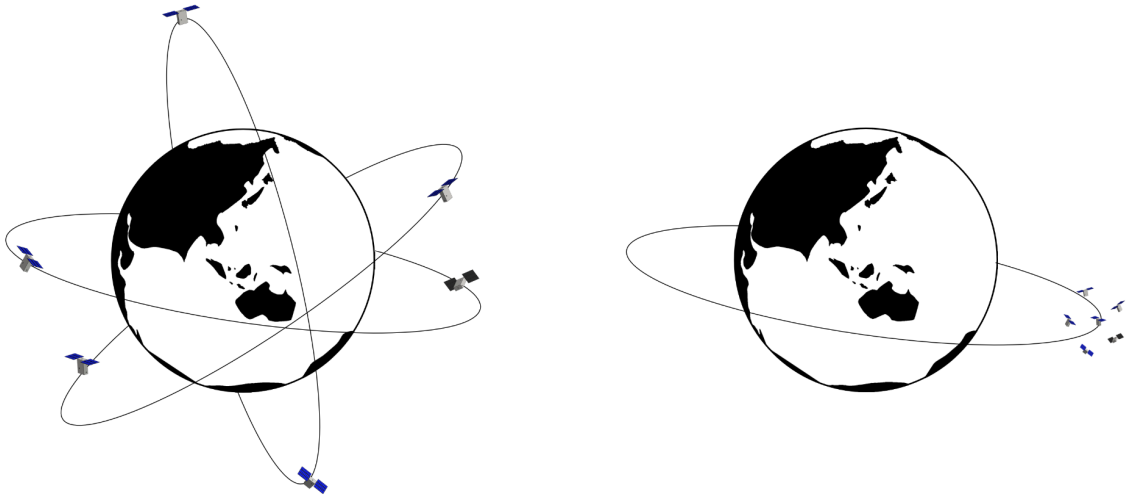


Figure 1.1: Constellation and formation flying systems (left: constellation, right: formation flying).

Among the formation flying, the systems composed of tens or hundreds of spacecraft are called spacecraft swarm systems. Figure 1.2 shows a few examples of spacecraft swarm applications. With a number of spacecraft, swarm systems have the potential to provide larger space occupancy and higher mission redundancy than small-size formation flying systems. However, at the same time, it can be more challenging to manage a number of spacecraft as one coordinated system.

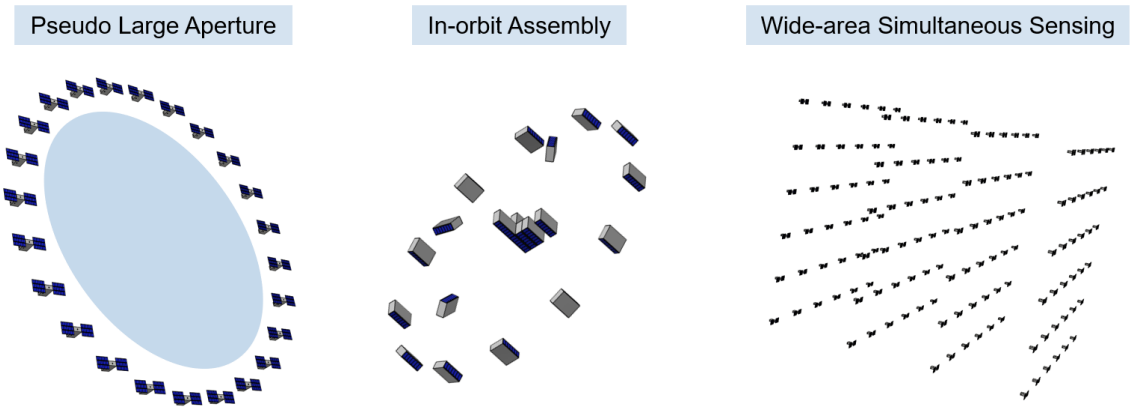


Figure 1.2: Examples of spacecraft swarm systems.

1.1.2 Spacecraft Swarm Missions

In this subsection, we introduce the applications and mission concepts of spacecraft swarms which are proposed in the related work.

Pseudo Large Aperture

One of the applications of spacecraft swarms is a distributed aperture composed of multiple spacecraft. By deploying several spacecraft in a specific area, they can behave like a pseudo large aperture. Figures 1.3 to 1.5 show the examples of distributed aperture missions by spacecraft swarms.

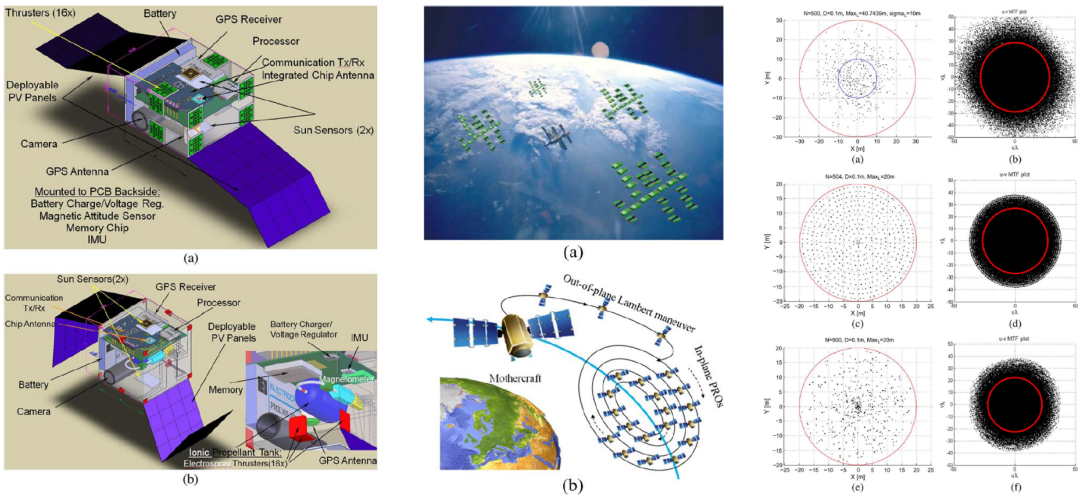


Figure 1.3: Sparse aperture missions by Swarms of Silicon Wafer Integrated Femtosatellites (SWIFT).[1]

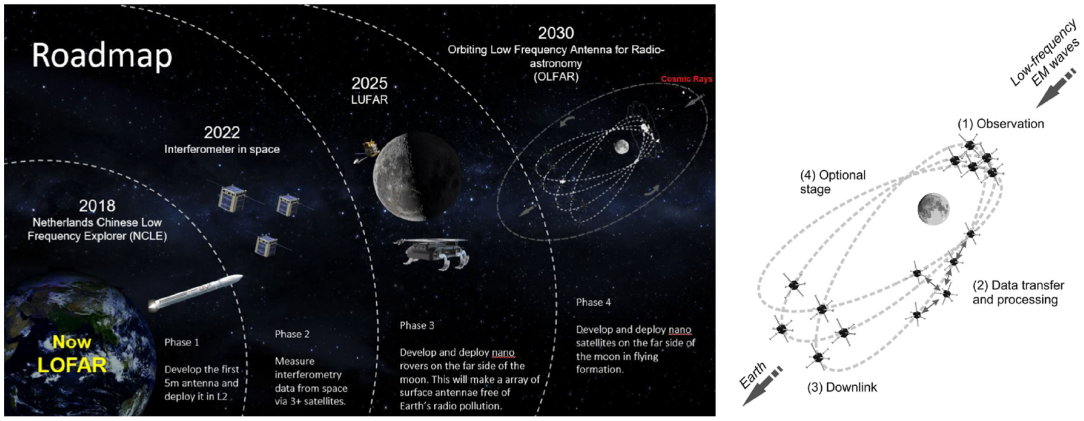


Figure 1.4: Orbiting Low Frequency Antennas for Radio Astronomy (OLFAR).[2, 3]

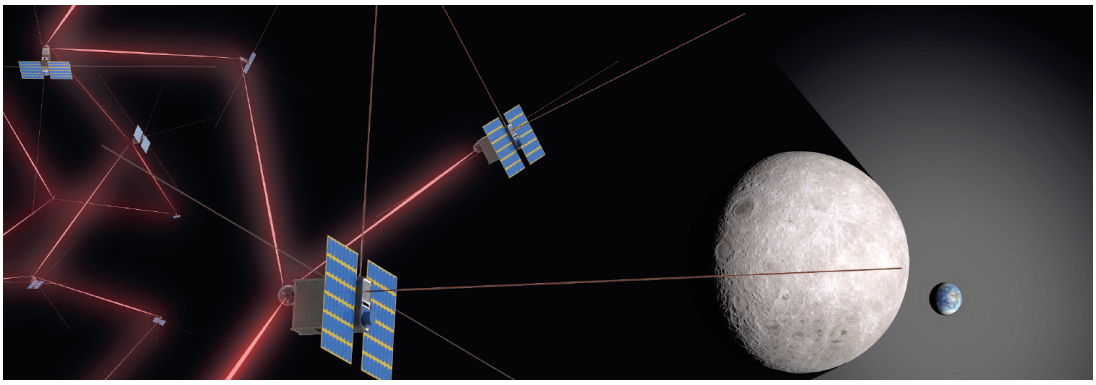


Figure 1.5: CubeSat Array for the Detection of RF Emissions from Exoplanets. [4]

In-Orbit Assembly

The size of space structures is often constrained by the fairing size of launch vehicles, payload mass, or durability against launch loading. In-orbit assembly by multiple spacecraft would solve these launch-related constraints and has the potential to realize large space systems. Figures 1.6 and 1.7 show the examples of in-orbit assembly missions by spacecraft swarms and the related technology.

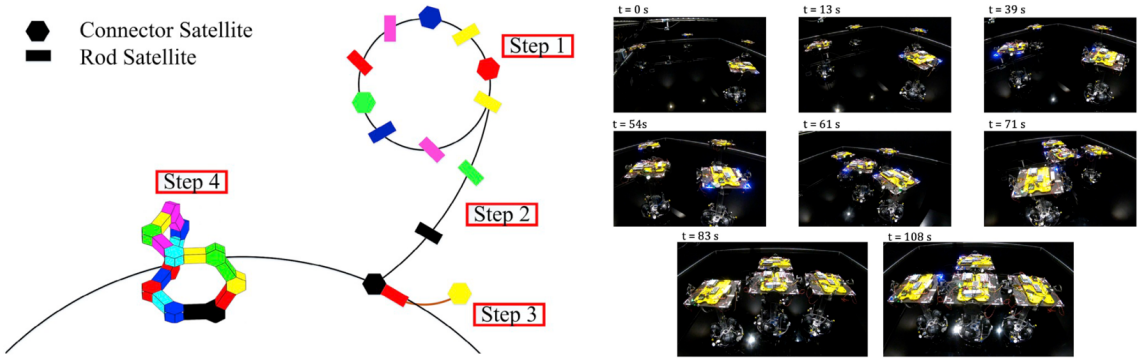


Figure 1.6: Autonomous assembly from heterogeneous modular spacecraft and the ground-based hardware simulation.[5]

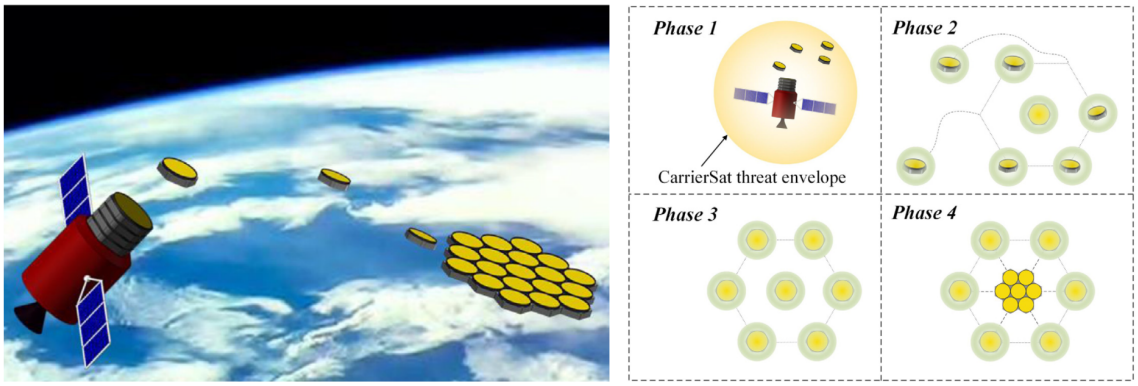


Figure 1.7: Guidance algorithm for autonomous assembly of a large aperture telescope.[6]

Deep Space Exploration

As shown in Figure 1.8, deep space exploration is another field where spacecraft swarms play an essential role. For instance, by deploying several spacecraft around an asteroid, they can monitor observation targets, such as gravitational distribution or asteroid materials, from various locations [13]. Additionally, a swarm system composed of spacecraft with heterogeneous abilities can significantly improve the mission value while keeping high

redundancy and scalability.

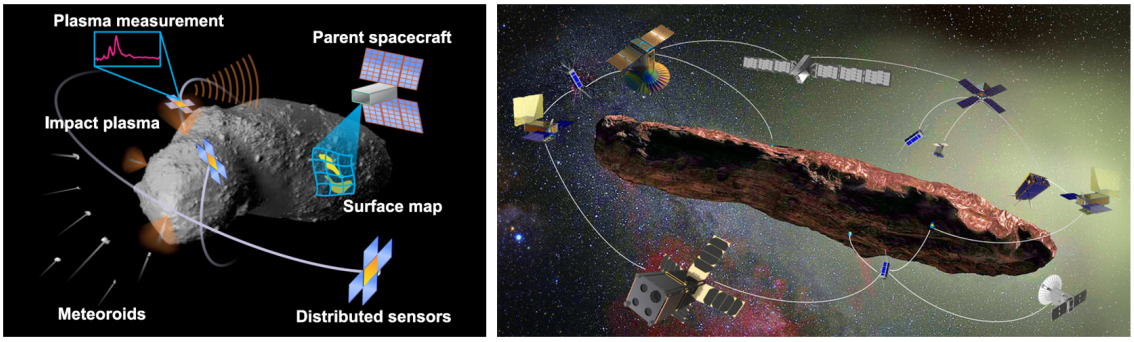


Figure 1.8: Missions around asteroids and comets (left: Meteoroid Impact Detection for Exploration of Asteroids [7], right: Large constellations and formations for exploring interstellar objects and long-period comets [8]).

Deep Space Infrastructure

Spacecraft swarms have gathered attention for supplemental applications for large space infrastructure, such as the Lunar Gateway [14]. By exploiting the advantages of spacecraft swarms, it is possible to realize communication relay and external health monitoring with appropriate resources, as shown in Figure 1.9.

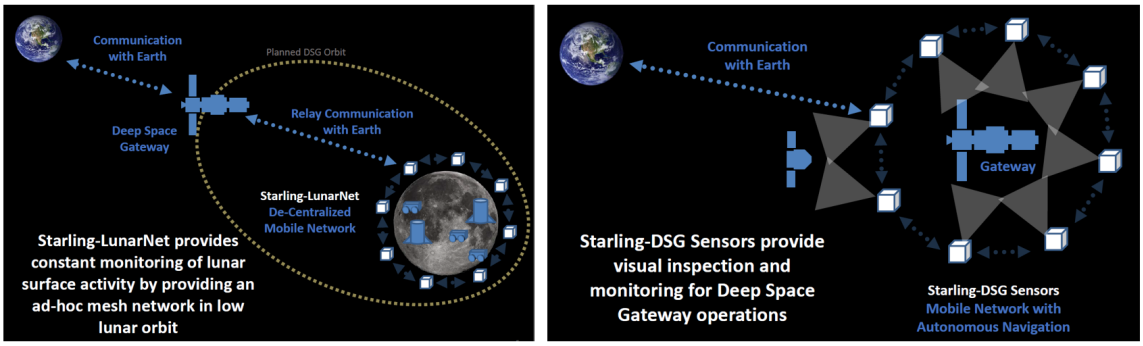


Figure 1.9: Lunar Gateway infrastructures by spacecraft swarms.[9]

1.1.3 Research Targets

To realize spacecraft swarm systems, various technical challenges need to be solved. These challenges include the operation of multiple spacecraft, inter-spacecraft data sharing, collision-free and fuel-efficient trajectory control, and so on [15]. Among those challenges, relative navigation (swarm localization) is one of the essential technical elements. For instance, in typical swarm applications, reconfiguration maneuvers are regularly operated,

as shown in Fig. 1.10, to save thrust fuels for station keeping or optimize the performance as multi-agent systems [16, 17]. Smart motion planning and collision avoidance algorithms, which are vital technologies for reconfiguration operations, cannot work without precise position and velocity information of spacecraft in the swarm systems [18].

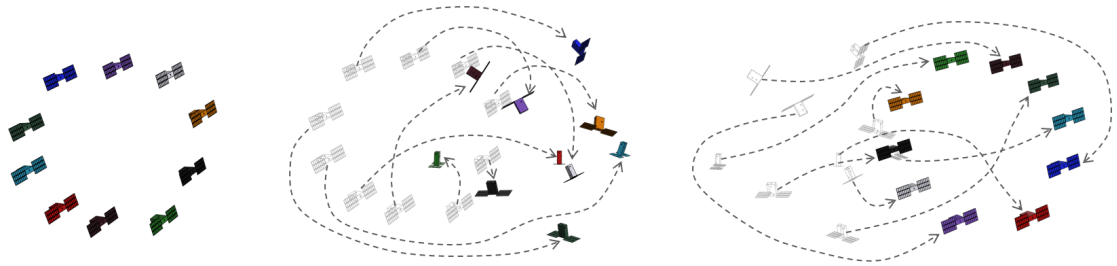


Figure 1.10: Reconfiguration operation of a swarm system.

In this paper, we consider a problem to estimate the relative states of spacecraft in the swarm system as illustrated in Figure 1.11. In the relative navigation for multi-agent systems such as spacecraft swarms, the available information for each agent is limited by inter-agent networks for communication and sensing, as shown in Figure 1.12. Therefore, we investigate an appropriate approach for the swarm navigation under the interaction constraints.

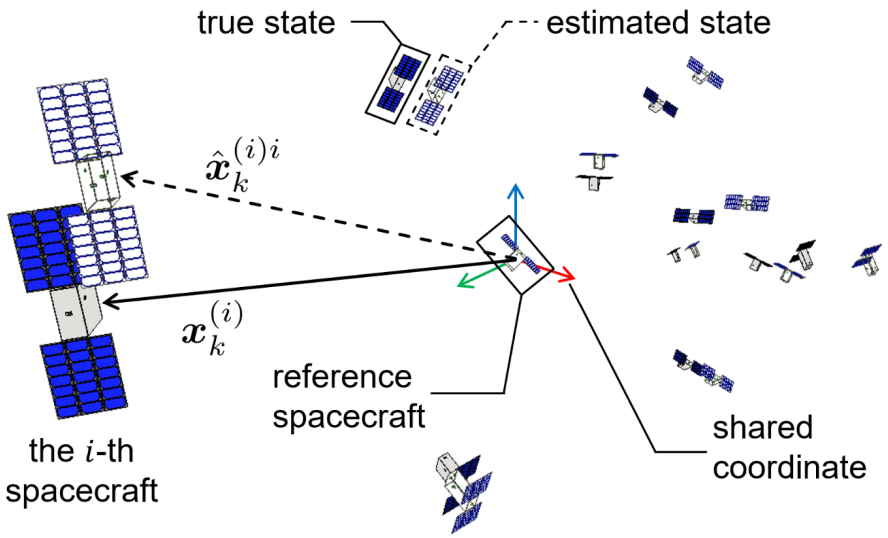


Figure 1.11: Relative navigation problem of a spacecraft swarm.

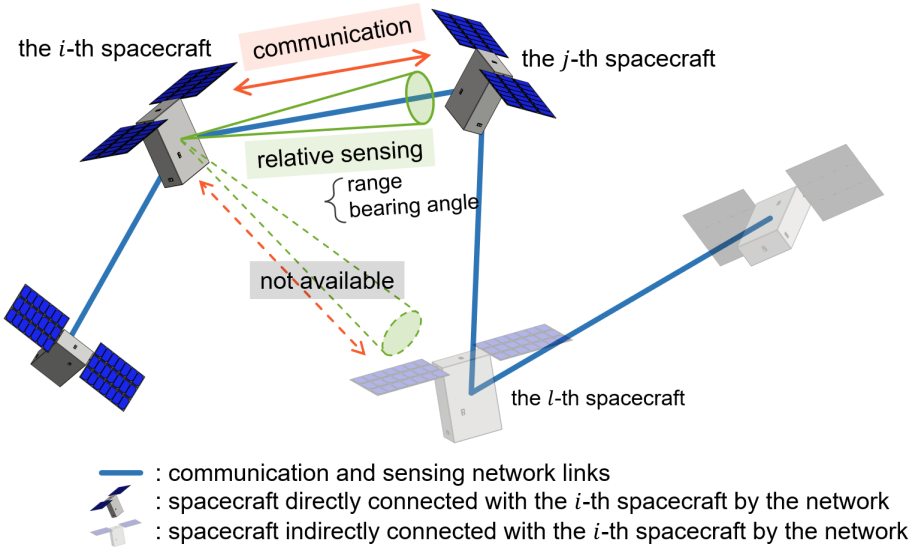


Figure 1.12: Constraints on interaction by communication and sensing networks.

1.2 Related Work

1.2.1 Target Tracking and Cooperative Localization

Multi-agent and multi-sensor localization problems have been widely studied in a variety of application fields as listed below:

- Sensor network tracking (localization) [19, 20, 21, 22, 23]
- Multi-ground-vehicle localization [10, 24, 25]
- UAV (Unmanned Aerial Vehicles) formation tracking (localization) [26, 27, 28, 29, 30, 31, 32, 33]
- Spacecraft formation flying (swarm) localization [34, 35, 36, 37, 38, 39, 40]

These algorithms can be roughly divided into two categories, target tracking [41] and cooperative localization [42]. The main difference between them is the dimensions of estimated states. In the target tracking problem, the dimensions of estimated states are not proportional to the number of agents or sensors. On the other hand, when it comes to the cooperative localization problem, the dimensions of the estimated state are typically proportional to the number of agents or sensors. Therefore, the estimated state can be huge, and an estimator may face challenges to handle a global state in terms of inter-agent communication and computation. The relative navigation problem of a spacecraft swarm is classified as cooperative localization.

1.2.2 Relative Navigation Algorithms for Spacecraft Swarm Systems

Table 1.1 summarizes the algorithmic features of representative relative navigation algorithms for spacecraft swarm systems. Although there are various aspects to classify the algorithms, we focus on sensing information, the size of estimated variables, decentralization levels, constraints on network topology, and algorithmic scalability.

Table 1.1: Characteristics of relative navigation algorithms for spacecraft swarms

Method name	Sensing	Estimated Variables	Decentralization	Network Topology	Scalability
Arambel (2001) [34]	relative range	full state	fully	ring shape	low
McLouglin (2008) [39]	relative position relative angle	full state	fully	ring shape	low
Bezouska (2019) [35]	relative position relative attitude absolute position	full state	fully	no constraint	low
Burchett (2019) [43]	relative range absolute position	single state	fully	no constraint	high
Kaputi (2020) [40]	relative range relative range rate absolute position	single state	fully	no constraint	high
Matsuka (2021) [38]	relative range relative angle absolute position	single state partial state	fully	no constraint	high

Here we briefly review each feature to clarify the requirements or preferable characteristics for spacecraft swarm navigation. In terms of sensing information, whether spacecraft can access absolute position or not significantly affects the requirements for navigation algorithms. For instance, if spacecraft can receive Global Navigation Satellite System (GNSS) signals and locate its absolute position, it is straightforward to compute the relative states of the other spacecraft, and the algorithmic stability is guaranteed even though the estimation accuracy of GNSS is not so high. When a spacecraft swarm system is deployed in low or middle Earth orbits, each spacecraft can receive positioning service from GNSS, as shown on the left side of Figure 1.13. However, when it comes to swarm missions for deep space applications, it is impossible to access to GNSS, and the spacecraft need to rely solely on relative sensing, as illustrated on the right side of Figure 1.13. In the most

recent works, it is assumed that all or some of the spacecraft in the swarm can access to GNSS-based absolute information, and they cannot be applied to swarm missions deployed on GNSS-denied environments such as GEO or deep space, even though they have strong features in network topology constraints, scalability, etc.

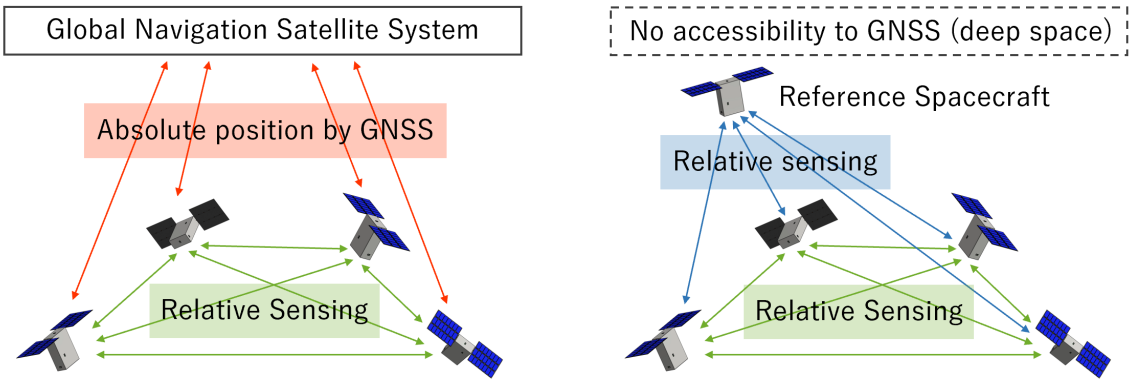


Figure 1.13: Accessibility to the information of absolute position.

The size of estimated variables can also greatly affect the performance of navigation algorithms. Figure 1.14 shows rough algorithmic classifications based on the number of estimation targets. In the case of the single-state estimation, each spacecraft only estimates its own state. This scheme substantially saves communication and computational costs at the expense of the global information of the swarm system.

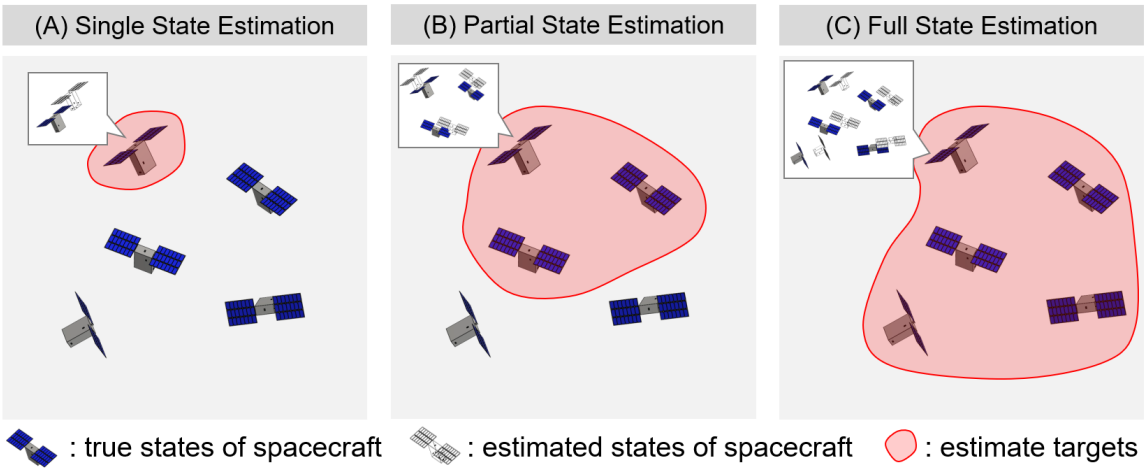


Figure 1.14: Allocation of the estimated targets: (A) single state estimation, (B) partial state estimation, and (C) full state estimation

Navigation algorithms calculate the estimated states based on sensing information and

estimation results shared from the other spacecraft. Therefore, depending on the algorithmic features, they possibly require some communication and sensing networks constraints.

The performance of swarm navigation is greatly affected by spacecraft configurations, initial estimation errors, temporary or permanent spacecraft malfunction, and network structures, as illustrated in Figure 1.15. Therefore, it is essential to prove that a navigation algorithm can adequately handle a variety of conditions for the practical applications of swarm navigation. However, in the previous work, sufficient performance verification has not been carried out against the algorithms proposed in their work as listed in Table 1.2, though it is mentioned that the algorithms have the potential to handle various conditions considered in the swarm systems.

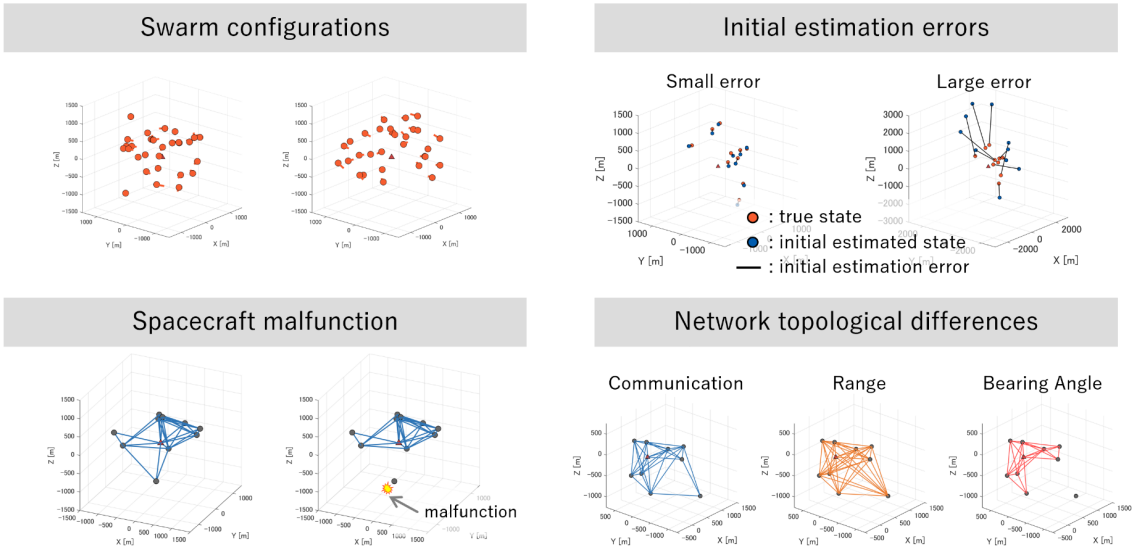


Figure 1.15: Various situations which spacecraft swarms may experience in practical operations.

Table 1.2: Performance verification of relative navigation algorithms in the previous work

Method name	Swarm configurations	Initial estimation errors	Fault tolerance	Network structure	Comparison with other methods
Arambel (2001) [34]	fixed	fixed	none	fixed	none
McLouglin (2008) [39]	fixed	fixed	none	fixed	Centralized KF
Bezouska (2019) [35]	random	fixed	none	4 types	none
Burchett (2019) [43]	several	fixed	none	almost fixed	none
Kaputi (2020) [40]	fixed	fixed	none	fixed	none
Matsuka (2021) [38]	fixed	fixed	none	fixed	Centralized KF

1.2.3 Summary of literature review

The previous work has not developed an algorithm that manages the following conditions while keeping high accuracy and stability:

- All the spacecraft in the swarm system cannot access absolute position (except for the reference spacecraft).
- No constraints on communication and sensing network topology.

as illustrated in Figure 1.16. The relative navigation algorithm with these features will be indispensable, especially for large-size swarm missions operated in GEO or deep space environments in the near future.

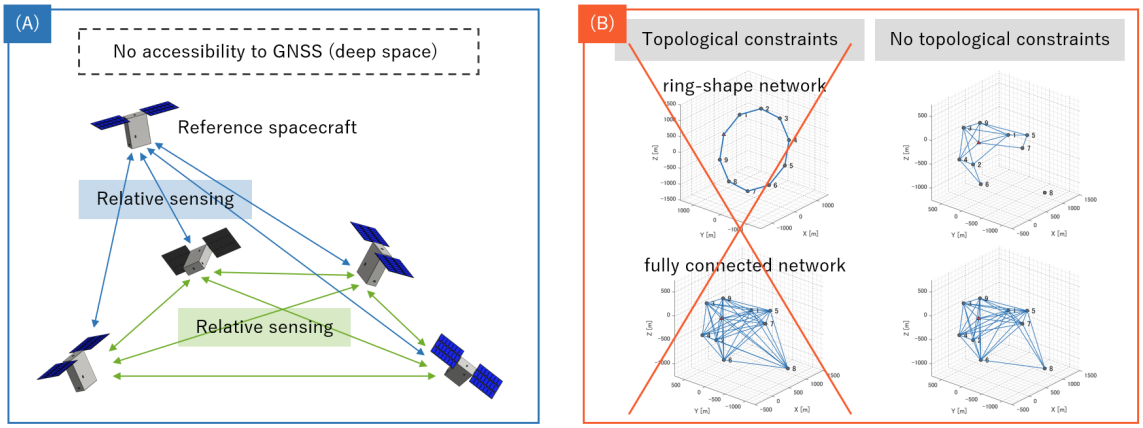


Figure 1.16: Open problems to be solved in this paper: (A) the absolute position measurements are not available. (B) No constraints are applied to the network topology.

1.3 Contributions

The main contributions and highlights of this research are as follows:

- We review the representative existing estimation algorithms for swarm navigation and reformulate them to enable a fair performance comparison with the proposed algorithms.
- We propose new estimation algorithms achieving high accuracy and high stability under the realistic conditions of swarm operations.
  - ADF: Adaptive Decentralized Filter
  - OD-ADF: Observability-Driven Adaptive Decentralized Filter

- We implement the existing and the proposed algorithms in the spacecraft swarm simulator and verify the effectiveness of the ADF and the OD-ADF by the numerical simulations.

1.4 Organization

This dissertation is organized as follows:

- Chapter 1 provides the research background of spacecraft swarm navigation and the objective of this research.
- Chapter 2 provides the basic conditions assumed in this research.
- Chapter 3 provides the fundamental information of the recursive Bayesian filtering used as the main estimation algorithm in this research.
- Chapter 4 provides the overview of the spacecraft swarm simulator, developed for comparative analysis of estimation algorithms for spacecraft swarm navigation.
- Chapter 5 provides the algorithmic details of the existing centralized and decentralized algorithms. (The robust partially decentralized filter, which is one of the algorithms introduced in this chapter, is proposed by the author.)
- Chapter 6 provides the algorithmic detail of the proposed algorithm: the adaptive decentralized filter (ADF).
- Chapter 7 provides the algorithmic detail of the proposed algorithm: the observability-driven adaptive decentralized filter (OD-ADF).
- Chapter 8 provides the details of the numerical simulations: the simulation conditions, the simulation results, and the related discussions.
- Chapter 9 provides the summary of this dissertation.

Chapter 2

Problem Formulation for Swarm Navigation

2.1 Relative Navigation of Spacecraft Swarms

The swarm comprises $N \in \mathbb{N}$ member spacecraft and a single reference spacecraft. All the spacecraft estimate the relative states on the common Local-Vertical Local-Horizontal (LVLH), whose origin is set to the position of the reference spacecraft. The true value of the state of the i-th spacecraft at the k-th time step is defined as:

$$\boldsymbol{x}_k^{(i)} = \begin{bmatrix} p_{x,k}^{(i)} & p_{y,k}^{(i)} & p_{z,k}^{(i)} & v_{x,k}^{(i)} & v_{y,k}^{(i)} & v_{z,k}^{(i)} \end{bmatrix}^T \tag{2.1}$$

$$\boldsymbol{p}_k^{(i)} = \begin{bmatrix} p_{x,k}^{(i)} & p_{y,k}^{(i)} & p_{z,k}^{(i)} \end{bmatrix}^T \tag{2.2}$$

$$\boldsymbol{v}_k^{(i)} = \begin{bmatrix} v_{x,k}^{(i)} & v_{y,k}^{(i)} & v_{z,k}^{(i)} \end{bmatrix}^T \tag{2.3}$$

where $\boldsymbol{p}_k^{(i)} \in \mathbb{R}^3$ is the relative position vector and $\boldsymbol{v}_k^{(i)} \in \mathbb{R}^3$ is the relative velocity vector. Eq. (2.1) is defined with respect to the reference spacecraft. Let the estimated state vector of the i-th spacecraft, which is computed by the j-th spacecraft, be $\hat{\boldsymbol{x}}_k^{(i)j}$, and let the corresponding covariance matrix be $\boldsymbol{P}_k^{(i)j} \in \mathbb{R}^{6\times 6}$. Therefore, for instance, the estimated state vector of the i-th spacecraft, which is computed by the i-th spacecraft itself, is represented as $\hat{\boldsymbol{x}}_k^{(i)i}$.

The navigation of spacecraft swarms can be roughly divided into three steps, relative sensing, inter-spacecraft communication, and estimate calculation, as illustrated in Figure 2.1. One of the influential factors is the availability of measurement data of relative

sensing. One spacecraft needs to obtain the measurement data, which the other spacecraft acquire with its sensor, by the communication relay.

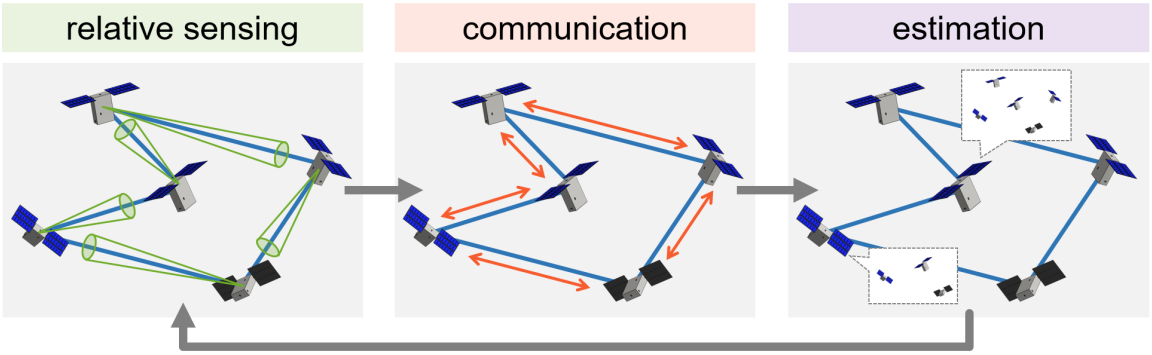


Figure 2.1: Basic navigation flow of spacecraft swarm systems.

2.2 Dynamics Model

The dynamics of the i-th spacecraft in discrete time is given by:

$$\boldsymbol{x}_{k+1}^{(i)} = \boldsymbol{f}_k\left(\boldsymbol{x}_k^{(i)}, \boldsymbol{w}_k^{(i)}, \Delta t\right), \qquad \Delta t = t_{k+1} - t_k \tag{2.4}$$

where $\boldsymbol{f}_k : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}^6$ is the possibly nonlinear time-varying function of the state vector, $\boldsymbol{w}_k^{(i)} \in \mathbb{R}^6$ is a process noise, and $\Delta t \in \mathbb{R}$ is the discretization time step size.

We consider a spacecraft swarm system orbiting around a single planet as illustrated in Figure 2.2. We ignore the other disturbance or perturbation forces for simplicity. The inertial equation of motion of the reference spacecraft is given by [44]:

$$\ddot{\boldsymbol{r}}^{(0)} = -\frac{\mu}{\left(r^{(0)}\right)^3}\boldsymbol{r}^{(0)} \tag{2.5}$$

and the the inertial equation of motion of the i-th spacecraft is also defined in the same form:

$$\ddot{\boldsymbol{r}}^{(i)} = -\frac{\mu}{\left(r^{(i)}\right)^3}\boldsymbol{r}^{(i)} \tag{2.6}$$

where $\mu \in \mathbb{R}$ is the gravitational constant, $\boldsymbol{r}^{(0)} \in \mathbb{R}^3$ and $\boldsymbol{r}^{(i)} \in \mathbb{R}^3$ are the inertial position vectors of the reference and the i-th spacecraft, respectively.

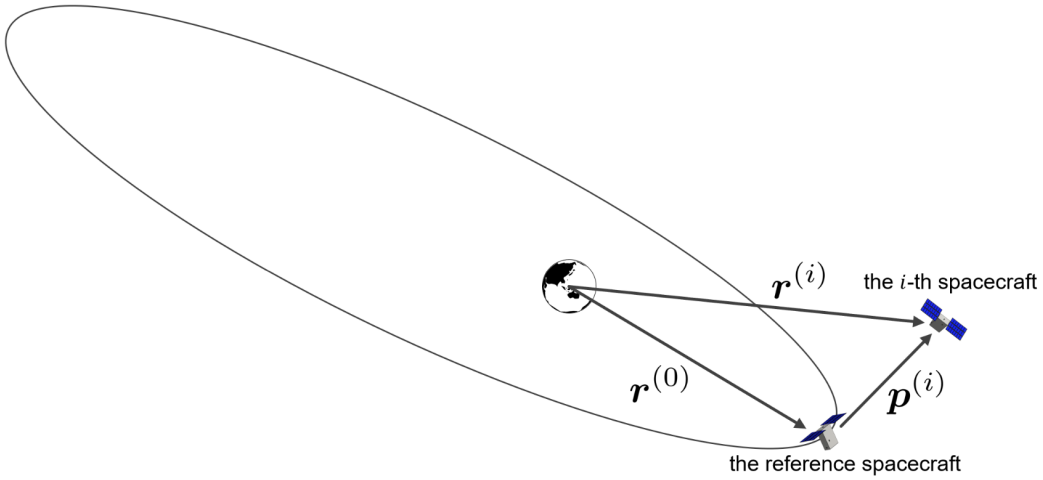


Figure 2.2: Spacecraft positions in the planet-centered inertial coordinate system.

The relative position vector of the i-th spacecraft with respect to the reference spacecraft satisfies the following equation.

$$\boldsymbol{p}^{(i)} = \boldsymbol{r}^{(i)} - \boldsymbol{r}^{(0)} \tag{2.7}$$

Then, substituting Eq. (2.5) and Eq. (2.6) into Eq. (2.7) yields the following equation of relative motion.

$$\ddot{\boldsymbol{p}}^{(i)} = -\frac{\mu\left(\boldsymbol{r}^{(0)} + \boldsymbol{p}^{(i)}\right)}{\left\|\boldsymbol{r}^{(0)} + \boldsymbol{p}^{(i)}\right\|_2^3} + \frac{\mu}{\left(r^{(0)}\right)^3}\boldsymbol{r}^{(0)} \tag{2.8}$$

We assume the following additional conditions on Eq. (2.7) to simplify the problem.

- The reference spacecraft follows a circular orbit.
- The relative distance between the reference and the i-th spacecraft is much shorter than than the inertial position vector of the reference spacecraft $\left(\left\|\boldsymbol{p}^{(i)}\right\| \ll \left\|\boldsymbol{r}^{(0)}\right\|\right)$.

Based on these two assumptions, the relative orbital dynamics of the i-th spacecraft can be reduced to the Hill-Clohessy-Wiltshire (HCW) equations represented by.

$$\boldsymbol{x}_{k+1}^{(i)} = \boldsymbol{F}^{(i)}\boldsymbol{x}_k^{(i)} + \boldsymbol{G}^{(i)}\boldsymbol{u}_k^{(i)} + \boldsymbol{w}_k^{(i)} \tag{2.9}$$

$$\boldsymbol{F}^{(i)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3\omega^2 & 0 & 0 & 0 & 2\omega & 0 \\ 0 & 0 & 0 & -2\omega & 0 & 0 \\ 0 & 0 & -\omega^2 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{G}^{(i)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.10}$$

where $\omega \in \mathbb{R}$ is the mean anomaly of the reference spacecraft orbit and $\boldsymbol{u}_k^{(i)} \in \mathbb{R}^3$ is a control input. [45] We assume that the process noise follows the Gaussian distribution $\boldsymbol{w}_k^{(i)} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{Q}^{(i)}\right)$ and

$$\boldsymbol{Q}^{(i)} = \begin{bmatrix} \sigma_{qp}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{qp}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{qp}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{qv}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{qv}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{qv}^2 \end{bmatrix}, \quad \sigma_{qp}^2 = \alpha_{qp}\Delta t, \quad \sigma_{qv}^2 = \alpha_{qv}\Delta t \tag{2.11}$$

where α_{qp} and α_{qv} are the proportional constants for position and velocity errors, respectively. Although there are higher fidelity dynamics models [46], we consider the above simple HCW model to focus on the performance evaluation of navigation algorithms.

2.3 Measurement Models

As relative sensing, each spacecraft can measure the relative distance and bearing angles of the other spacecraft with its onboard sensors. Figure 2.3 shows the relative sensing of range and bearing angles between the i-th and the j-th spacecraft.

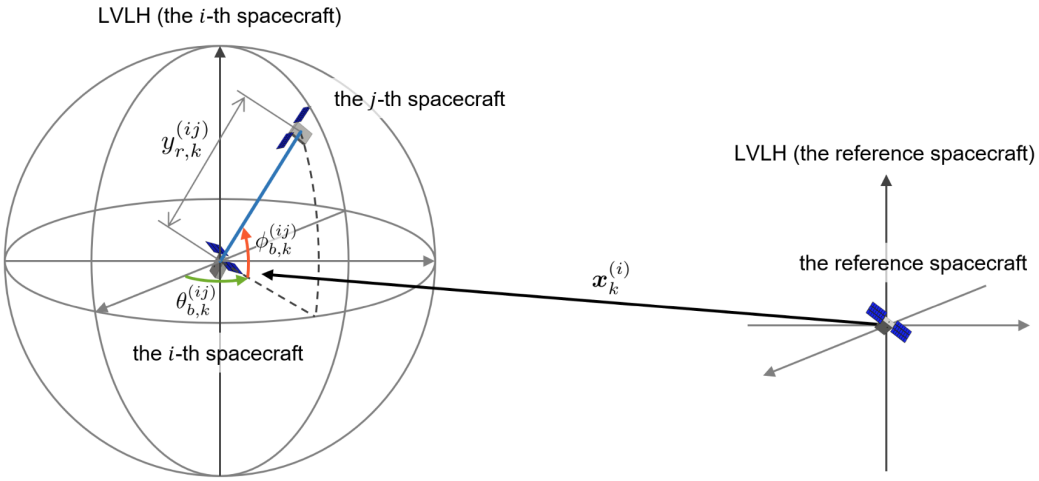


Figure 2.3: Relative sensing between the i-th and the j-th spacecraft.

2.3.1 Range Measurements

The range measurement between two spacecraft is a function of their state vectors, which are represented by:

$$\begin{aligned} y_r^{(ij)} &= h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right) + v_r \\ h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right) &= \left\|\boldsymbol{p}^{(j)} - \boldsymbol{p}^{(i)}\right\|_2 \qquad (2.12) \\ &= \sqrt{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2 + \left(p_z^{(j)} - p_z^{(i)}\right)^2} \qquad (2.13) \end{aligned}$$

where $h_r : \mathbb{R}^6 \to \mathbb{R}$ is a nonlinear function of two state vectors and $v_r \in \mathbb{R}$ is a measurement noise of ranging. We assume that the range noise follows a Gaussian distribution $v_r \sim \mathcal{N}(0, \sigma_r^2)$, where $\sigma_r \in \mathbb{R}$ is a standard deviation.

Here we derive the Jacobian matrix of h_r for the linearization mentioned in the later part. The Jacobian matrix with respect to $\boldsymbol{x}^{(i)}$ is defined as:

$$\begin{aligned} \frac{\partial h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right)}{\partial \boldsymbol{x}^{(i)}} &= \begin{bmatrix} \frac{\partial h_r}{\partial p_x^{(i)}} & \frac{\partial h_r}{\partial p_y^{(i)}} & \frac{\partial h_r}{\partial p_z^{(i)}} & \frac{\partial h_r}{\partial v_x^{(i)}} & \frac{\partial h_r}{\partial v_y^{(i)}} & \frac{\partial h_r}{\partial v_z^{(i)}} \end{bmatrix} \in \mathbb{R}^{1\times 6} \qquad (2.14) \\ &= \begin{bmatrix} \frac{\partial h_r}{\partial p_x^{(i)}} & \frac{\partial h_r}{\partial p_y^{(i)}} & \frac{\partial h_r}{\partial p_z^{(i)}} & 0 & 0 & 0 \end{bmatrix} \qquad (2.15) \end{aligned}$$

where the derivatives with respect to the velocity are zero. The derivative with respect to

the position is formulated as follows.

$$\begin{aligned}\frac{\partial h_r\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)}{\partial p_a^{(i)}} &= \frac{p_a^{(i)}-p_a^{(j)}}{\sqrt{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2+\left(p_z^{(j)}-p_z^{(i)}\right)^2}} \\ &= \frac{p_a^{(i)}-p_a^{(j)}}{h_r\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)}, \qquad (a\in\{x,y,z\}) \end{aligned} \tag{2.16}$$

2.3.2 Bearing Angle Measurements

The bearing angle measurements between two spacecraft is a function of their state vectors, which are represented by:

$$\boldsymbol{y}^{(ij)} = \begin{bmatrix}\theta^{(ij)}\\ \\ \phi^{(ij)}\end{bmatrix} + \boldsymbol{v}_b = \boldsymbol{h}_b\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)+\boldsymbol{v}_b \tag{2.17}$$

$$\boldsymbol{h}_b\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right) = \begin{bmatrix}\tan^{-1}\left(\frac{\Delta p_y^{(ij)}}{\Delta p_x^{(ij)}}\right)\\ \\ \tan^{-1}\left(\frac{\Delta p_z^{(ij)}}{\sqrt{(\Delta p_x^{(ij)})^2+(\Delta p_y^{(ij)})^2}}\right)\end{bmatrix} \tag{2.18}$$

$$\Delta p_{a,k}^{(ij)} = p_{a,k}^{(j)} - p_{a,k}^{(i)} \qquad (a\in\{x,y,z\}) \tag{2.19}$$

where $\boldsymbol{h}_b:\mathbb{R}^6\to\mathbb{R}^2$ is a nonlinear function of state vectors and $\boldsymbol{v}_b\in\mathbb{R}^2$ is the measurement noise of bearing angles. We assume that the measurement noise of bearing angles follows a Gaussian distribution $\boldsymbol{v}_a\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}_b)$, where $\boldsymbol{R}_b=\mathrm{diag}\left(\sigma_b^2,\sigma_b^2\right)\in\mathbb{R}^{2\times2}$ is the covariance matrix. These relative angles are defined with respect to the common LVLH frame.

The Jacobian matrix of the bearing angle measurements is defined as:

$$\frac{\partial \boldsymbol{h}_b\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)}{\partial \boldsymbol{x}^{(i)}} = \begin{bmatrix}\frac{\partial\theta^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial p_z^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial v_x^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial v_y^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial v_z^{(i)}}\\ \frac{\partial\phi^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial p_z^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial v_x^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial v_y^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial v_z^{(i)}}\end{bmatrix} \in\mathbb{R}^{2\times6} \tag{2.20}$$

$$= \begin{bmatrix}\frac{\partial\theta^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial\theta^{(ij)}}{\partial p_z^{(i)}} & 0 & 0 & 0\\ \frac{\partial\phi^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial\phi^{(ij)}}{\partial p_z^{(i)}} & 0 & 0 & 0\end{bmatrix} \tag{2.21}$$

where the derivatives with respect to the velocity are zero as with the range measurement. To express the non-zero terms in simple forms, the following two variables are defined as

follows.

$$\xi \triangleq \frac{p_y^{(j)} - p_y^{(i)}}{p_x^{(j)} - p_x^{(i)}} \tag{2.22}$$

$$\zeta \triangleq \frac{p_z^{(j)} - p_z^{(i)}}{\sqrt{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2}} \tag{2.23}$$

The non-zero derivatives are represented by the following equations with the two parameters ξ, ζ as:

$$\frac{\partial \theta^{(ij)}}{\partial p_x^{(i)}} = \frac{\xi}{1+\xi^2} \cdot \frac{1}{p_x^{(j)} - p_x^{(i)}} \tag{2.24}$$

$$\frac{\partial \theta^{(ij)}}{\partial p_y^{(i)}} = -\frac{1}{1+\xi^2} \cdot \frac{1}{p_x^{(j)} - p_x^{(i)}} \tag{2.25}$$

$$\frac{\partial \theta^{(ij)}}{\partial p_z^{(i)}} = 0 \tag{2.26}$$

$$\frac{\partial \phi^{(ij)}}{\partial p_x^{(i)}} = \frac{\zeta}{1+\zeta^2} \cdot \frac{p_x^{(i)} - p_x^{(j)}}{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2} \tag{2.27}$$

$$\frac{\partial \phi^{(ij)}}{\partial p_y^{(i)}} = \frac{\zeta}{1+\zeta^2} \cdot \frac{p_y^{(i)} - p_y^{(j)}}{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2} \tag{2.28}$$

$$\frac{\partial \phi^{(ij)}}{\partial p_z^{(i)}} = (-1) \cdot \frac{1}{1+\zeta^2} \cdot \frac{1}{\sqrt{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2}} \tag{2.29}$$

$$\frac{\partial \phi^{(ij)}}{\partial p_z^{(j)}} = \frac{1}{1+\zeta^2} \cdot \frac{1}{\sqrt{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2}} \tag{2.30}$$

2.4 Communication and Sensing Networks

We use expressions used in graph theory to represent the inter-spacecraft interaction, such as communication and sensing [47]. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a directed graph, with $\mathcal{V} = \{0, 1, ..., N\}$ the set of spacecraft and $\mathcal{E}$ the set of edges between spacecraft. The node $\{0\}$ represents the reference spacecraft. The edge $(i, j) \in \mathcal{E}$ shows that there is some kind of interaction between the i-th and the j-th spacecraft. The set of spacecraft directly connected with the i-th spacecraft is defined as $\mathcal{V}_i = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}$. The adjacent spacecraft including the i-th spacecraft itself is defined as $\tilde{\mathcal{V}}_i = \mathcal{V}_i \cup \{i\}$. The number of

elements in a spacecraft set or an edge set is computed with an operator $n(\cdot)$, for instance, $n(\mathcal{V}) = 3$ when $\mathcal{V} = \{1, 2, 3\}$.

This research considers two types of networks: communication and sensing networks. In addition, the sensing networks can be classified into ranging and bearing angle networks as shown in Figure 2.4. These networks are distinguished with the superscripts such as $\mathcal{G}^x$ $(x \in \{c, r, b\})$, where $\mathcal{G}^c$ represents the communication network, $\mathcal{G}^r$ represents the sensing network of ranging, and $\mathcal{G}^b$ represents the sensing network of bearing angle measurements. In addition, networks are possibly time-varying; hence a network at the k-th time step is expressed with the subscript as $\mathcal{G}_k$. Therefore, for instance, the network of ranging measurements at the k-th time step is represented as $\mathcal{G}_k^r = (\mathcal{V}, \mathcal{E}_k^r)$, where we assume that the number of spacecraft in the swarm is constant.

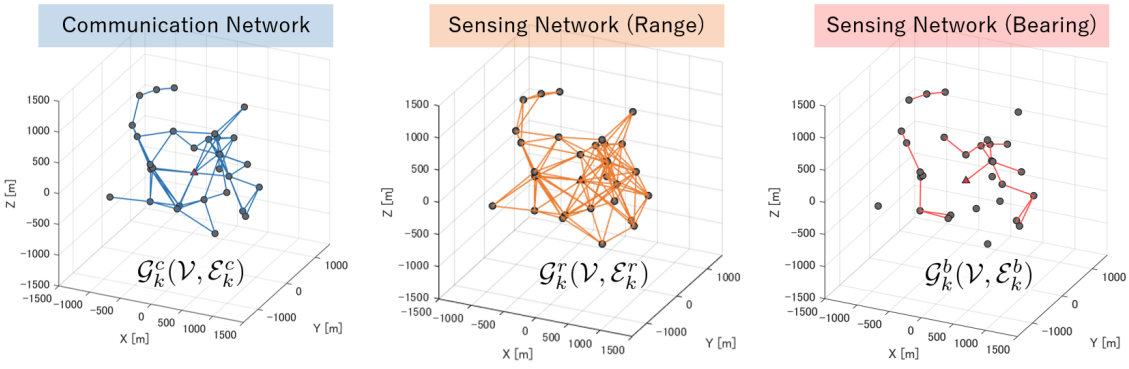


Figure 2.4: Communication and sensing networks.

The network connectivity is governed by inter-spacecraft distances limited by communication signal strength or the field of views of onboard sensors. If the distance between two spacecraft is shorter than ΔR, defined as the network threshold, the two spacecraft is connected. Figure 2.5 shows the network topologies corresponding to different network thresholds. The network connectivity of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ can be mathematically expressed with the adjacency matrix $\mathcal{A}(\mathcal{G}) \in \mathbb{R}^{(N+1)\times(N+1)}$ defined as:

$$[\mathcal{A}(\mathcal{G})]_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \tag{2.31}$$

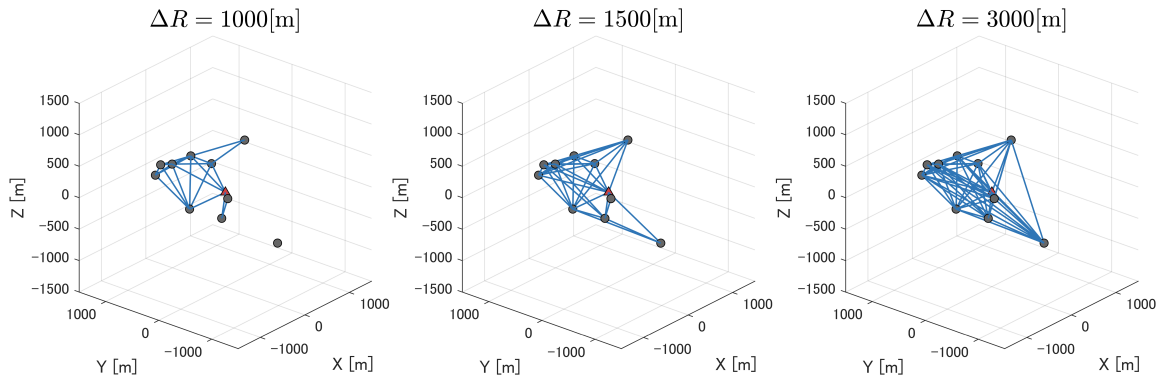


Figure 2.5: Network topologies with different network thresholds.

We can define the ratio of the number of neighbor nodes to the total number of nodes in the graph as the network connection rate. The network connectivity significantly affects the estimation performances and communication loads; hence, it is essential to consider how strongly the communication and sensing networks are connected. The connection rate of the i-th spacecraft $\rho^{(i)}$ is formulated as:

$$\rho^{(i)} = \frac{d\left(\{i\}, \mathcal{G}\right)}{n(\mathcal{V}) - 1} \tag{2.32}$$

$$d\left(\{i\}, \mathcal{G}\right) = \sum_{j \in \mathcal{V}} \left[\mathcal{A}(\mathcal{G})\right]_{ij} \tag{2.33}$$

where $d\left(\{i\}, \mathcal{G}\right)$ is an operator to calculate the degree of the i-th node in the network $\mathcal{G}$. We can also define the average connection rate $\bar{\rho}$ as:

$$\bar{\rho} = \frac{1}{n(\mathcal{V})} \sum_{i \in \mathcal{V}} \rho^{(i)} = \frac{n(\mathcal{E})}{{}_{n(\mathcal{V})}C_2}. \tag{2.34}$$

Additionally, we define the connection rate averaged over swarm configurations as:

$$\breve{\rho} = \frac{1}{M} \sum_{m=1}^{M} {}^{m}\bar{\rho} \tag{2.35}$$

where ${}^{m}\bar{\rho}$ is the connection rate of Eq. (2.34) evaluated for the m-th swarm configuration.

2.5 Variables and Measurements with Network Topology

In an actual navigation process, the collection of spacecraft states is possibly estimated with the collection of sensor measurements. This section summarizes the representation of aggregated states and measurements, which are determined by the communication and

sensing network topologies. We consider a system illustrated in Figure 2.6. The communication and sensing graphs are mathematically expressed as:

$$\mathcal{G}^c = (\mathcal{V}, \mathcal{E}^c) \tag{2.36}$$

$$\mathcal{G}^s = (\mathcal{V}, \mathcal{E}^s) \tag{2.37}$$

$$\mathcal{V} = \{i, j_1, j_2, j_3, j_4, j_5, j_6\} \tag{2.38}$$

$$\mathcal{E}^c = \{(i, j_1), (i, j_2), (j_2, j_3), (j_3, j_4), (j_4, j_5), (j_4, j_6)\} \tag{2.39}$$

$$\mathcal{E}^s = \{(i, j_1), (i, j_2), (j_1, j_3), (j_2, j_3), (j_3, j_4), (j_3, j_6), (j_4, j_5), (j_4, j_6)\} \tag{2.40}$$

where we assume that the ranging network and bearing angle network are the same as $\mathcal{G}^s = \mathcal{G}^r = \mathcal{G}^b$.

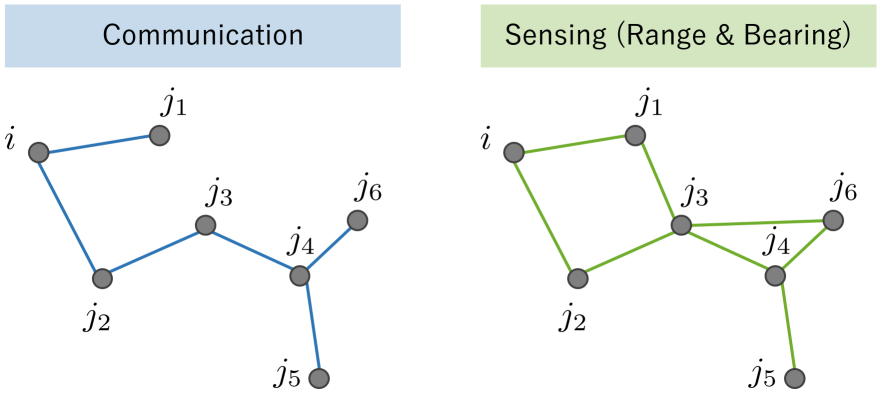


Figure 2.6: Communication and sensing networks for the handling of aggregated states and measurements.

As the first example, we consider locally available information for the i-th spacecraft without inter-spacecraft communication. In this case, the i-th spacecraft can obtain relative sensing information with respect to the j_1-th and the j_2-th spacecraft. The set of sensing edges is $\mathcal{E}_i^s = \{(i, j_1), (i, j_2)\}$. Then, the aggregated function of relative measurements for the i-th spacecraft is defined as:

$$\boldsymbol{h}_{\mathcal{E}_i^s} = \begin{bmatrix} h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_1)}\right) \\ h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_2)}\right) \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_1)}\right) \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_2)}\right) \end{bmatrix} \tag{2.41}$$

where $\boldsymbol{h}_{\mathcal{E}_i^s} : \mathbb{R}^{n(\tilde{\mathcal{V}}_i)n_x} \to \mathbb{R}^{n_y}$, the functions h_r and $\boldsymbol{h}_b$ are previously defined as Eq. (2.12) and Eq. (2.18) respectively. $n(\tilde{\mathcal{V}}_i)n_x$ represents the dimensions of the aggregated state vector defined as:

$$\boldsymbol{x}^{(\tilde{\mathcal{V}}_i)} = \begin{bmatrix} \left(\boldsymbol{x}^{(i)}\right)^T & \left(\boldsymbol{x}^{(j_1)}\right)^T & \left(\boldsymbol{x}^{(j_2)}\right)^T \end{bmatrix}^T \tag{2.42}$$

and n_y means the dimension of the aggregated measurement vector. n_y is computed as:

$$n_y = n\left(\mathcal{E}_i^s\right) \times (n_r + n_b) \tag{2.43}$$

where $n_r = 1$ is the dimension of range measurement and $n_b = 2$ is the dimensions of bearing angle measurements, including azimuth and elevation angles.

As the second example, we consider locally available information for the i-th spacecraft under one-hop communication. Based on the sensing network, the i-th spacecraft can obtain the measurements of (i, j_1) and (i, j_2) in a similar way to the previous example. Through the communication network, the i-th spacecraft can also receive some information that the j_1-th and j_2-th spacecraft possess. The j_1-th spacecraft has the measurement (j_1, j_3) and the j_2-th spacecraft has the measurement (j_2, j_3); hence, the i-th spacecraft can indirectly access to the measurements (j_1, j_3) and (j_2, j_3). Therefore, the set of sensing edges the i-th spacecraft can access by one-hop communication is:

$$\mathcal{E}_i^s = \{(i, j_1), (i, j_2), (j_1, j_3), (j_2, j_3)\} \tag{2.44}$$

In this case, the aggregated measurement function is expressed as:

$$\boldsymbol{h}_{\mathcal{E}_i^s} = \begin{bmatrix} h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_1)}\right) \\ \vdots \\ h_r\left(\boldsymbol{x}^{(j_2)}, \boldsymbol{x}^{(j_3)}\right) \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j_1)}\right) \\ \vdots \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(j_2)}, \boldsymbol{x}^{(j_3)}\right) \end{bmatrix} \tag{2.45}$$

We can also consider a complete graph where an edge connects every pair of distinct nodes. The complete graph and its elements are expressed with the superscript "c" as:

$\mathcal{G}_c^x = (\mathcal{V}, \mathcal{E}_c^x)$ where $x \in \{c, s\}$. For instance, in the complete graph of Figure 2.7, the aggregated measurement function is formulated as:

$$\boldsymbol{h}_{\mathcal{E}_k^{s[c]}} = \begin{bmatrix} \boldsymbol{h}_{\mathcal{E}_k^{r[c]}} \\ \boldsymbol{h}_{\mathcal{E}_k^{b[c]}} \end{bmatrix} = \begin{bmatrix} h_r\left(\boldsymbol{x}^{(j_1)}, \boldsymbol{x}^{(j_2)}\right) \\ \vdots \\ h_r\left(\boldsymbol{x}^{(j_3)}, \boldsymbol{x}^{(j_4)}\right) \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(j_1)}, \boldsymbol{x}^{(j_2)}\right) \\ \vdots \\ \boldsymbol{h}_b\left(\boldsymbol{x}^{(j_3)}, \boldsymbol{x}^{(j_4)}\right) \end{bmatrix} \tag{2.46}$$

where $\boldsymbol{h}_{\mathcal{E}_k^{s[c]}} : \mathbb{R}^{n(\mathcal{V})n_x} \to \mathbb{R}^{n_y}$ and $n(y)$ is commputed as:

$$n_y = n\left(\mathcal{E}_k^{s[c]}\right) \times (n_r + n_b) = {}_{n(\mathcal{V})}C_2 \times (n_r + n_b) \tag{2.47}$$

hence, $n_y = {}_4C_2 \times (1 + 2) = 18$.

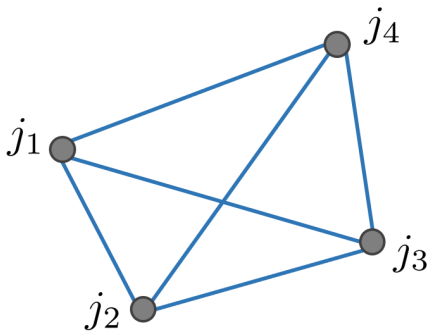


Figure 2.7: Example of a complete graph.

Chapter 3

Recursive Bayesian Filtering

3.1 Bayesian Filter

This chapter reviews the fundamental information of the recursive Bayesian filtering used as the main estimation algorithm in this research. Let an estimated state vector be $\boldsymbol{x} \in \mathbb{R}^{n_x}$ and let a measurement vector be $\boldsymbol{y} \in \mathbb{R}^{n_y}$. These two vectors assumed to be random variables and the state vector $\boldsymbol{x}$ has a known prior probability function $p(\boldsymbol{x})$. Additionally, the condition density (likelihood function) $p(\boldsymbol{y}|\boldsymbol{x})$ is also pre-known. The posterior conditional probability is calculated by Bayes' law as [48]:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})} \tag{3.1}$$

where the denominator is a scalar normalization constant computed as:

$$p(\boldsymbol{y}) = \int_{\mathbb{R}^{n_x}} p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x} \tag{3.2}$$

Bayesian recursive filters are divided into the prediction and the measurement update.

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}) = \int_{\mathbb{R}^{n_x}} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1})d\boldsymbol{x}_{k-1} \tag{3.3}$$

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})}{\int_{\mathbb{R}^{n_x}} p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})d\boldsymbol{x}_k} \tag{3.4}$$

3.2 Gaussian Filter

3.2.1 Gaussian Probability Density Function

The Gaussian probability density function of the random variable $\boldsymbol{x}$ is defined with the two moment parameters, the mean vector $\hat{\boldsymbol{x}}$ and the covariance matrix $\boldsymbol{P}$, as,

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\,\hat{\boldsymbol{x}}, \boldsymbol{P}) = \frac{1}{\sqrt{(2\pi)^{n_x}\,|\boldsymbol{P}|}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^T \boldsymbol{P}^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right\} \tag{3.5}$$

where the operator $|\cdot|$ provides the determinant of a matrix.

3.2.2 Gaussian Assumption

In this section, we summarize the algorithmic background of the Gaussian filters. [49] We consider the nonlinear filtering problem of the following system expressed as:

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k \tag{3.6}$$

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{3.7}$$

where $\boldsymbol{w} \in \mathbb{R}^{n_x}$ is the process noise and $\boldsymbol{v} \in \mathbb{R}^{n_y}$ is the measurement noise. Here we assume that the noises are independent from the state vector and they follow the Gaussian distributions as $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{w}|\boldsymbol{0}, \boldsymbol{Q})$ and $\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{v}|\boldsymbol{0}, \boldsymbol{R})$, where $\boldsymbol{Q} \in \mathbb{R}^{n_x \times n_x}$ and $\boldsymbol{Q} \in \mathbb{R}^{n_y \times n_y}$.

In the framework of the Gaussian filters, the following conditions are assumed:

- The posterior probability $p\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1}\right)$ follows the Gaussian distribution.
- The prior probability $p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}\right)$ follows the Gaussian distribution.
- The posterior probability $p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}\right)$ follows the Gaussian distribution.

Under these Gaussian assumptions, the above probability density functions are represented as:

$$p\left(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1}\right) = \mathcal{N}\left(\boldsymbol{x}_{k-1}|\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{P}_{k-1|k-1}\right) \tag{3.8}$$

$$p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}\right) = \mathcal{N}\left(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}_{k|k-1}\right) \tag{3.9}$$

$$p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}\right) = \mathcal{N}\left(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{P}_{k|k}\right) \tag{3.10}$$

where the right sides are defined as Eq. (3.5). This research deals with probability as the Gaussian distributions to enable the onboard computation for the estimation algorithms.

3.2.3 Formulation of the Gaussian Filter

We reformulate the computation process of the Bayesian recursive filter, defined as Eqs. (3.3) and (3.4), under the Gaussian assumptions of Eqs. (3.8) to (3.10). The detail of the derivation process is summarized in Appendix A.1.1. The prediction step of the Gaussian filter is formulated as:

$$\hat{\boldsymbol{x}}_{k|k-1} = \mathrm{E}\left[\boldsymbol{f}(\boldsymbol{x}_{k-1})\right] = \int_{\mathbb{R}^{n_x}} \boldsymbol{f}(\boldsymbol{x}_{k-1})\mathcal{N}(\boldsymbol{x}_{k-1}|\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{P}_{k-1})d\boldsymbol{x}_{k-1} \tag{3.11}$$

$$\begin{aligned}\boldsymbol{P}_{k|k-1} &= \mathrm{Cov}\left[\boldsymbol{f}(\boldsymbol{x}_{k-1}), \boldsymbol{f}(\boldsymbol{x}_{k-1})\right] = \mathrm{E}\left[(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))(\boldsymbol{x}_k - \boldsymbol{f}(\boldsymbol{x}_{k-1}))^T\right] + \boldsymbol{Q} \\ &= \int_{\mathbb{R}^{n_x}} \left(\boldsymbol{f}(\boldsymbol{x}_{k-1}) - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\boldsymbol{f}(\boldsymbol{x}_{k-1}) - \hat{\boldsymbol{x}}_{k|k-1}\right)^T \\ &\qquad \times \mathcal{N}(\boldsymbol{x}_{k-1}|\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{P}_{k-1})d\boldsymbol{x}_{k-1} + \boldsymbol{Q}\end{aligned} \tag{3.12}$$

and the measurement update step is formulated as:

$$\hat{\boldsymbol{y}}_{k|k-1} = \mathrm{E}[\boldsymbol{h}(\boldsymbol{x}_k)] = \int_{\mathbb{R}^{n_x}} \boldsymbol{h}(\boldsymbol{x}_k)\mathcal{N}(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}_{k|k-1})d\boldsymbol{x}_k \tag{3.13}$$

$$\begin{aligned}\boldsymbol{P}^{yy}_{k|k-1} &= \mathrm{E}\left[(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T\right] + \boldsymbol{R} \\ &= \int_{\mathbb{R}^{n_x}} \left(\boldsymbol{h}(\boldsymbol{x}_k) - \hat{\boldsymbol{y}}_{k|k-1}\right)\left(\boldsymbol{h}(\boldsymbol{x}_k) - \hat{\boldsymbol{y}}_{k|k-1}\right)^T \\ &\qquad \times\mathcal{N}(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}_{k|k-1})d\boldsymbol{x}_k + \boldsymbol{R}\end{aligned} \tag{3.14}$$

$$\begin{aligned}\boldsymbol{P}^{xy}_{k|k-1} &= \mathrm{E}\left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T\right] \\ &= \int_{\mathbb{R}^{n_x}} \left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}\right)\left(\boldsymbol{h}(\boldsymbol{x}_k) - \hat{\boldsymbol{y}}_{k|k-1}\right)^T \mathcal{N}(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{P}_{k|k-1})d\boldsymbol{x}_k\end{aligned} \tag{3.15}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xy}_{k|k-1}\left(\boldsymbol{P}^{yy}_{k|k-1}\right)^{-1} \tag{3.16}$$

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}\right) \tag{3.17}$$

$$\boldsymbol{P}_k = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{yy}_{k|k-1}\boldsymbol{K}_k^T. \tag{3.18}$$

where $\boldsymbol{K}$ represents the gain of the measurement update.

3.3 Extended Information Filter

When the system functions $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{h}(\boldsymbol{x})$ are nonlinear, it is difficult to compute the following integrals appeared in the Gaussian filter:

$$\mathrm{E}\left[\boldsymbol{g}(\boldsymbol{x})\right] = \int_{\mathbb{R}^{n_x}} \boldsymbol{g}(\boldsymbol{x})\mathcal{N}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}, \boldsymbol{P}\right)d\boldsymbol{x} \tag{3.19}$$

$$\mathrm{Cov}\left[\boldsymbol{g}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})\right] = \int_{\mathbb{R}^{n_x}} \left(\boldsymbol{g}(\boldsymbol{x}) - \mathrm{E}\left[\boldsymbol{g}(\boldsymbol{x})\right]\right)\left(\boldsymbol{g}(\boldsymbol{x}) - \mathrm{E}\left[\boldsymbol{g}(\boldsymbol{x})\right]\right)^T \mathcal{N}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}, \boldsymbol{P}\right)d\boldsymbol{x} \tag{3.20}$$

where $\boldsymbol{g}(\boldsymbol{x}) \in \{\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{h}(\boldsymbol{x})\}$. In this research, we introduce the linear approximation for the function $\boldsymbol{g}(\boldsymbol{x})$ to satisfy the Gaussian assumption mentioned in Section 3.2.2. The detail of the linear approximation of Eqs. (3.19) and (3.20) are discussed in Appendix A.2.

Additionally, we consider the different parameterizations to formulate the Gaussian filters. A Gaussian distribution is typically represented by the two moments, mean and covariance. On the other hand, a Gaussian distribution is also expressed with an information matrix and an information vector in the canonical parameterization. The different parameterization of Gaussian distributions leads to the different formulations of the Gaussian filters in the prediction and the measurement update. [50] The information matrix $\boldsymbol{Z} \in \mathbb{R}^{n_x \times n_x}$ and the information vector $\hat{\boldsymbol{z}} \in \mathbb{R}^{n_x}$ of the Gaussian distribution $\mathcal{N}(\boldsymbol{x}|\hat{\boldsymbol{x}}, \boldsymbol{P})$

are defined as:

$$\boldsymbol{Z} = \boldsymbol{P}^{-1} \tag{3.21}$$

$$\hat{\boldsymbol{z}} = \boldsymbol{P}^{-1}\hat{\boldsymbol{x}} \tag{3.22}$$

where the distribution is also expressed by the canonical parameters as $\mathcal{N}(\boldsymbol{x}|\boldsymbol{Z}^{-1}\hat{\boldsymbol{z}}, \boldsymbol{Z}^{-1})$. The canonical parameterization provides a simpler formulation of the measurement update than the moment parameterization. Therefore, it would enable the efficient implementation of decentralized estimation algorithms introduced in the following chapter.

The Gaussian filter is reformulated as the extended information filter by combining the above-mentioned linear approximation of probability integrals and the canonical parameterization. The algorithmic detail of the extended information filter is shown in Algorithm 1.

Algorithm 1: Extended Information Filter

At the k-th time step;

input : $\hat{\boldsymbol{z}}_{k-1}, \boldsymbol{Z}_{k-1}$

output: $\hat{\boldsymbol{z}}_k, \boldsymbol{Z}_k$

Prediction step;

$\boldsymbol{F}_k = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{Z}_{k-1}^{-1}\hat{\boldsymbol{z}}_{k-1}}$;

$\boldsymbol{Z}_{k|k-1} = \left(\boldsymbol{F}_k \boldsymbol{Z}_{k-1}^{-1} \boldsymbol{F}_k^T + \boldsymbol{Q}\right)^{-1}$;

$\hat{\boldsymbol{z}}_{k|k-1} = \boldsymbol{Z}_{k|k-1}\left(\boldsymbol{F}_k \boldsymbol{Z}_{k-1}^{-1}\hat{\boldsymbol{z}}_{k-1} + \boldsymbol{G}_k \boldsymbol{u}_k\right)$;

$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{Z}_{k|k-1}^{-1}\hat{\boldsymbol{z}}_{k|k-1}$;

Measurement update step ;

$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{k|k-1}}$;

$\boldsymbol{Z}_k = \boldsymbol{Z}_{k|k-1} + \boldsymbol{H}_k^T \boldsymbol{R}^{-1} \boldsymbol{H}_k$;

$\hat{\boldsymbol{z}}_k = \hat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{H}_k^T \boldsymbol{R}^{-1}\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\hat{\boldsymbol{x}}_{k|k-1}\right) + \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}\right]$;

Chapter 4

Spacecraft Swarm Simulator

4.1 Overview of the Simulator

Figure 4.1 illustrates the significant functionalities and the data flow in the simulator. We briefly summarize the essential points in the simulator. The network manager and the measurement generator provide the corresponding network and measurement information based on the true state vectors. The measurement generator adds proper random noises for the measurement information to simulate actual sensor noises. Then, the measurements are trimmed based on the network connectivity and delivered to the estimation algorithms. Each algorithm executes its estimation process and provides the estimated state vectors. In parallel, the communication manager computes the communication load by monitoring the network manager and the estimation algorithms. Finally, all the information is collected to the single visualization platform enabling various performance analyses.

4.2 General Settings for the Simulations

4.2.1 Common Settings

We consider a swarm system composed of 10 spacecraft ($N = 9$ member spacecraft and a reference spacecraft). The reference spacecraft moves along a circular orbit around the Earth at an altitude of 500km. Table 4.1 summarizes the parameters used in the simulations.

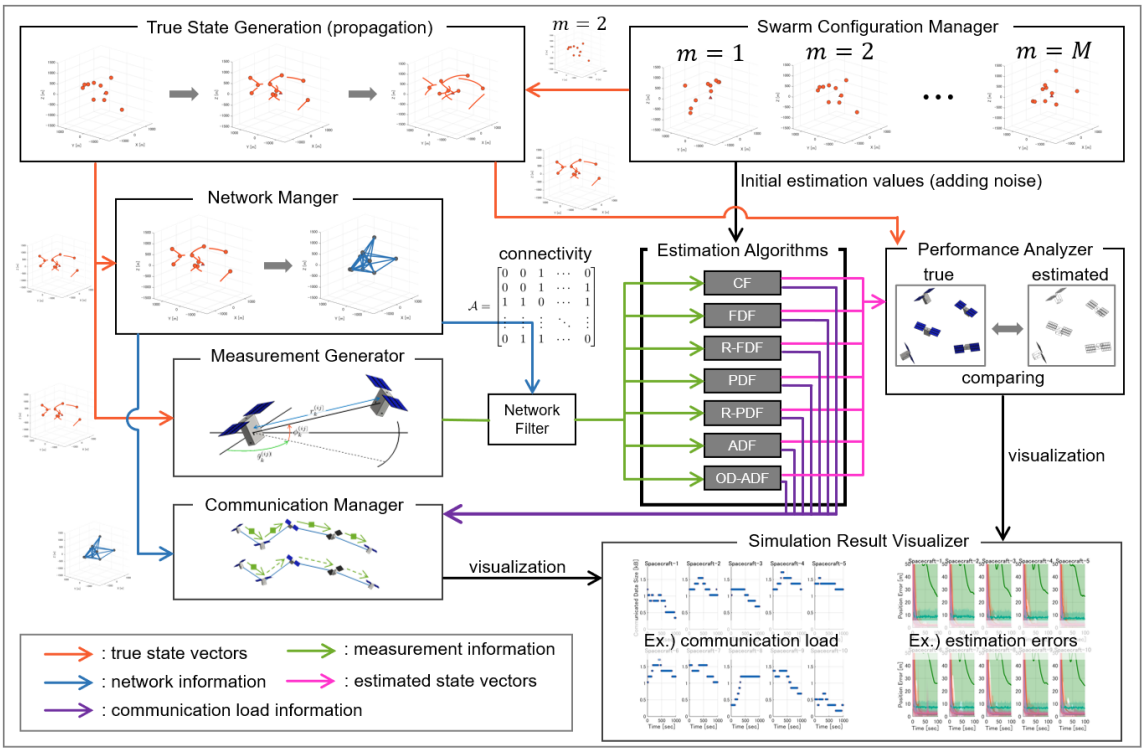


Figure 4.1: Overview of the spacecraft swarm simulator.

4.2.2 Generation of Initial Estimation Errors

The mean value of the initial estimation is randomly generated from the Gaussian distribution, that is defined as:

$$p\left(\hat{\boldsymbol{x}}_0^{(i)}\right) = \mathcal{N}\left(\hat{\boldsymbol{x}}_0^{(i)} \middle| \boldsymbol{x}_0^{(i)}, \boldsymbol{P}_{init}\right) \tag{4.1}$$

where $\boldsymbol{x}_0^{(i)}$ represents the true initial state of the i-th spacecraft and $\boldsymbol{P}_{init} \in \mathbb{R}^{6\times6}$ is the covariance matrix defined as:

$$\boldsymbol{P}_{init} = \mathrm{diag}\left(\sigma_{p_{init}}^2, \sigma_{p_{init}}^2, \sigma_{p_{init}}^2, \sigma_{v_{init}}^2, \sigma_{v_{init}}^2, \sigma_{v_{init}}^2\right). \tag{4.2}$$

4.2.3 Network Connectivity

The network connectivity is one of the most influential factors affecting estimation performance. In the practical operation of spacecraft swarms, the network connectivity is governed by various constraints, such as the applicable range of onboard sensors or the communication capacity of transceivers. Based on these constraints, we consider distance-based connectivity.

Table 4.1: Common simulation parameters

Parameter	Value	Units
Number of the member spacecraft (N)	9	[-]
Total simulation time	100.0	[s]
Estimation interval	1.0	[s]
STD of the range measurement noise (σ_r)	1.0	[m]
STD of the bearing angle measurement noise (σ_b)	0.01	[rad]
Data size of a single measurement	256	[bit]
Data size of a single state vector	264	[bit]
Data size of a single state covariance	1224	[bit]

STD: standard deviation

When the distance-based connectivity determines the inter-spacecraft links, a pair of spacecraft is connected if their physical distance is shorter than the threshold ΔR. Therefore, the degree of each spacecraft is not necessarily identical.

4.3 Performance Evaluation Metrics

4.3.1 Convergence Rate

The estimation stability is assessed with the convergence rate. If the 3σ error range, which is calculated from the covariance matrix, bounds an estimation error, the result is regarded as a converged estimate. We introduce the flag variable of convergence ${}^{m}\mu_{x_f}^{(i)} \in \{0, 1\}$, that represents the convergence of the filter $x_f \in X_f$ for the estimate of the i-th spacecraft under the m-th swarm configuration. X_f is a set of estimation filters to be evaluated. The value of the convergence flag is defined as:

$$ {}^{m}\mu_{x_f}^{(i)} = \begin{cases} 1 & \text{(The estimate is converged.)} \\ 0 & \text{(The estimate is diverged.)} \end{cases} \tag{4.3} $$

In addition, we introduce a gross convergence flag that evaluates whether the estimates are converged for all the spacecraft or not. The gross convergence flag is defined as:

$$ {}^{m}\mu_{x_f} = \begin{cases} 1 & \text{if} \quad {}^{m}\mu_{x_f}^{(i)} = 1 \quad (\forall i \in \mathcal{V}) \\ 0 & \text{otherwise} \end{cases} \tag{4.4} $$

The judgement of the convergence in Eqs. 4.3 and 4.4 are executed by Algorithm 2, where k_e is the index of the last time step. $\mathcal{X}^2_\nu$ represents the inverse chi-squared cumulative distribution function defined as:

$$ \mathcal{X}^2_\nu(x) = \int_0^x \frac{t^{\frac{\nu-2}{2}} e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} dt \tag{4.5} $$

where ν is the dimension of the variable x, and Γ represents the gamma function.

Algorithm 2: Convergence judgment

input : $\boldsymbol{x}_{k_e}^{(i)},\ \mathcal{N}\left(\boldsymbol{x}_{k_e}^{(i)} \mid \hat{\boldsymbol{x}}_{k_e}^{(i)}, \boldsymbol{P}_{k_e}^{(i)}\right)\ (i \in \mathcal{V}), \quad x_f \in X_f$

output: ${}^{m}\mu_{x_f}, {}^{m}\mu_{x_f}^{(i)}\ (i \in \mathcal{V})$

```
m_mu_xf = 1;
for i ∈ V do
    Δe_{k_e}^{(i)} = ‖ m_p̂_{k_e}^{(i)} − m_p_{k_e}^{(i)} ‖_2;
    Δe_{99%}^{(i)} = sqrt( X_3^2(99%) · Υ_p^{(i)}( P_{k_e}^{(i)} ) );
    if Δe_{k_e}^{(i)} > Δe_{99%}^{(i)} then
        m_mu_xf^{(i)} = 0;
        m_mu_xf = 0;
    else
        m_mu_xf^{(i)} = 1;
    end
end
```

The convergence rate is defined as the ratio of the simulation cases with converged estimation results to the total number of the simulation cases. With the gross convergence flag, the convergence rate of a filter x_f is expressed as:

$$ \mathrm{CR}_{x_f} = \frac{1}{M} \sum_{m=1}^{M} {}^{m}\mu_{x_f} \tag{4.6} $$

4.3.2 Estimation Accuracy

Secondly, the estimation accuracy is evaluated by the two parameters, root mean square error (RMSE) and root of the trace of error covariance (RTEC), which are defined as:

$$\text{RMSE} = \frac{1}{n(\mathcal{M}_c)} \sum_{m \in \mathcal{M}_c} \left(\frac{1}{N} \sum_{i=1}^{N} \left| {}^{m}\hat{\boldsymbol{p}}_k^{(i)} - {}^{m}\boldsymbol{p}_k^{(i)} \right| \right), \tag{4.7}$$

$$\text{RTEC} = \frac{1}{n(\mathcal{M}_c)} \sum_{m \in \mathcal{M}_c} \left(\frac{1}{N} \sum_{i=1}^{N} \sqrt{\text{trace}\left({}^{m}\boldsymbol{P}_{p,k}^{(i)} \right)} \right). \tag{4.8}$$

where m is the index of swarm configurations. $\mathcal{M}_c$ represents the set of swarm configurations considered in the performance evaluation of estimation accuracy, which is generated by Algorithm 3. $\mathcal{M}$ is the set of all the swarm configuration to be evaluated, ${}^{m}\mu_{x_f}$ is computed by Algorithm 2, X_f is the set of filtering algorithms, and X_f^{ex} is the set of filtering algorithms which are not considered in the evaluation of estimation accuracy. For instance, $\mathcal{M}_c$ has few swarm configurations, and the meaningful evaluation becomes difficult when X_f has a filtering algorithm with a remarkably low convergence rate. In such a case, we set X_f^{ex} to exclude the algorithm with a low convergence rate from the evaluated algorithm set X_f.

Algorithm 3: Set of swarm configurations for the performance evaluation

input : $\mathcal{M}$, ${}^{m}\mu_{x_f}$, X_f, X_f^{ex}

output: $\mathcal{M}_c$

```
M_c = ∅;
for m ∈ M do
    if ^m μ_{x_f} = 1 (∀x_f ∈ X_f \ X_f^{ex}) then
        M_c = M_c ∪ {m};
    end
end
```

4.3.3 Communication Load

For the evaluation of inter-spacecraft communication load (CL), we assess the total amount of communicated data. Specifically, we consider three values, the maximum (maxCL), the minimum (minCL), and the average communication load (aveCL), which are

defined as:

$$\text{maxCL} = \frac{1}{M}\sum_{m=1}^{M}\left(\max\left({}^{m}\text{CL}^{(i)}\right)\right), \tag{4.9}$$

$$\text{minCL} = \frac{1}{M}\sum_{m=1}^{M}\left(\min\left({}^{m}\text{CL}^{(i)}\right)\right), \tag{4.10}$$

$$\text{aveCL} = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{N}\sum_{i=1}^{N}{}^{m}\text{CL}^{(i)}\right). \tag{4.11}$$

Chapter 5

Centralized and Decentralized Estimation

5.1 Centralized Filter

In the framework of centralized estimation, one leader spacecraft collects measurement information from all the other spacecraft and runs estimate calculation. Figures 5.1 and 5.2 show the data flow and the algorithmic flow of the centralized filter, respectively.

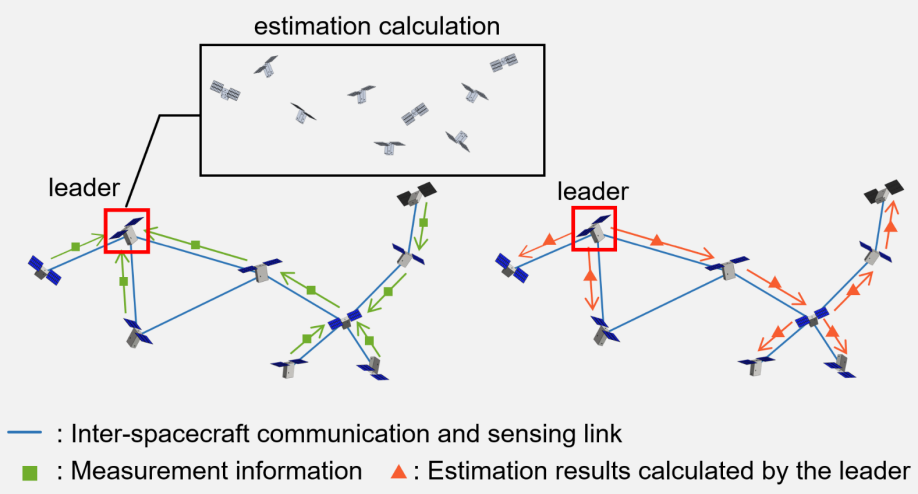


Figure 5.1: Inter-spacecraft communication in the centralized filter.

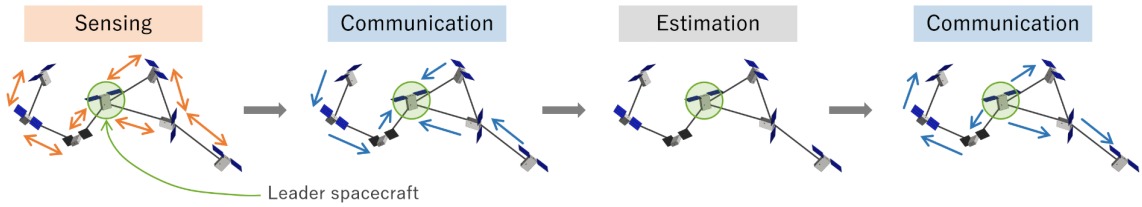


Figure 5.2: Algorithmic flow of the centralized filter.

The centralized filter estimates the aggregated state of all the spacecraft in the swarm. The aggregated state vector is defined as:

$$\boldsymbol{x}_k^{(c)} = \left[\left(\boldsymbol{x}_k^{(1)}\right)^T \quad \cdots \quad \left(\boldsymbol{x}_k^{(i)}\right)^T \quad \cdots \quad \left(\boldsymbol{x}_k^{(N)}\right)^T \right]^T \tag{5.1}$$

and the corresponding covariance matrix is

$$\boldsymbol{P}_k^{(c)} = \mathrm{E}\left[\left(\boldsymbol{x}_k^{(c)} - \hat{\boldsymbol{x}}_k^{(c)}\right) \left(\boldsymbol{x}_k^{(c)} - \hat{\boldsymbol{x}}_k^{(c)}\right)^T \right]. \tag{5.2}$$

Then, the information vector and the information matrix are computed as $\boldsymbol{z}_k^{(c)} = \left(\boldsymbol{P}_k^{(c)}\right)^{-1} \boldsymbol{x}_k^{(c)}$ and $\boldsymbol{Z}_k^{(c)} = \left(\boldsymbol{P}_k^{(c)}\right)^{-1}$ respectively.

For simplicity, we formulate the centralized estimation as the extended information filter, reviewed in Section 3.3. The prediction step is formulated as follows.

$$\boldsymbol{Z}_{k|k-1}^{(c)c} = \left(\boldsymbol{F}^{(c)} \left(\boldsymbol{Z}_{k-1}^{(c)c}\right)^{-1} \left(\boldsymbol{F}^{(c)}\right)^T + \boldsymbol{Q}^{(c)} \right)^{-1} \tag{5.3}$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(c)c} = \boldsymbol{Z}_{k|k-1}^{(c)c} \left(\boldsymbol{F}^{(c)} (\boldsymbol{Z}_{k-1}^{(c)c})^{-1} \hat{\boldsymbol{z}}_{k-1}^{(c)c} + \boldsymbol{G}^{(c)} \boldsymbol{u}_k^{(c)} \right) \tag{5.4}$$

where the aggregated matrices $\boldsymbol{F}^{(c)}$, $\boldsymbol{G}^{(c)}$, $\boldsymbol{Q}^{(c)}$ are defined as:

$$\boldsymbol{X}^{(c)} = \begin{bmatrix} \boldsymbol{X}^{(1)} & & 0 \\ & \ddots & \\ & \boldsymbol{X}^{(i)} & \\ & & \ddots \\ 0 & & \boldsymbol{X}^{(N)} \end{bmatrix}, \qquad \boldsymbol{X} \in \{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{Q}\}. \tag{5.5}$$

The non-diagonal block elements in Eq. (5.5) are zero since each spacecraft's motion and control input are independent.

To formulate the measurement update step, we firstly consider the measurements under the complete sensing graph. The measurement vector of the complete sensing graph is represented as:

$$\begin{aligned} \boldsymbol{y}_k^{(c)} &= \left[\left(\boldsymbol{y}_k^{(01)}\right)^T \quad \cdots \quad \left(\boldsymbol{y}_k^{(ij)}\right)^T \quad \cdots \quad \left(\boldsymbol{y}_k^{((N-1)N)}\right)^T \right]^T && (5.6) \\ &= \boldsymbol{h}_{\mathcal{E}_k^{s[c]}}\left(\boldsymbol{x}_k^{(c)}\right) + \boldsymbol{v}_k^{(c)} && (5.7) \end{aligned}$$

when we consider the range and bearing angle measurements separately, each measurement vector is represented as:

$$\begin{aligned} \boldsymbol{y}_{r,k}^{(c)} &= \boldsymbol{h}_{\mathcal{E}_k^{r[c]}}\left(\boldsymbol{x}_k^{(c)}\right) + \boldsymbol{v}_{r,k}^{(c)} && (5.8) \\ \boldsymbol{y}_{b,k}^{(c)} &= \boldsymbol{h}_{\mathcal{E}_k^{b[c]}}\left(\boldsymbol{x}_k^{(c)}\right) + \boldsymbol{v}_{b,k}^{(c)}. && (5.9) \end{aligned}$$

The Jacobian matrices of the aggregated measurement functions on the complete sensing graphs $\boldsymbol{h}_{\mathcal{E}_k^{r[c]}}$ and $\boldsymbol{h}_{\mathcal{E}_k^{b[c]}}$ are defined as:

$$\begin{aligned} \boldsymbol{H}_{r,k}^{(c)} &= \left.\frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{r[c]}}}{\partial \boldsymbol{x}_k^{(c)}}\right|_{\boldsymbol{x}_k^{(c)}=\hat{\boldsymbol{x}}_{k|k-1}^{(c)}} && (5.10) \\ \boldsymbol{H}_{b,k}^{(c)} &= \left.\frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{b[c]}}}{\partial \boldsymbol{x}_k^{(c)}}\right|_{\boldsymbol{x}_k^{(c)}=\hat{\boldsymbol{x}}_{k|k-1}^{(c)}} && (5.11) \end{aligned}$$

In actual situations, the sensing networks are possibly incomplete graphs. Hence, some algebraic processes ignore the sensing information corresponding to the unconnected edges. It is possible to adjust the contribution of each measurement by modifying the error covariance matrix of measurements. The adjusted covariance matrix is defined as:

$$\tilde{\boldsymbol{R}}_{s,k}^{(c)} = \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_k^s\right), \qquad s \in \{r, b\} \tag{5.12}$$

where Γ is a function to change the element in $\boldsymbol{R}_{s,k}^{(c)}$ to the infinity when the corresponding sensing edge is not included in $\mathcal{E}_k^s$. The aggregated measurement error covariance matrices $\boldsymbol{R}_{s,k}^{(c)},\ s \in \{r, b\}$ are defined as:

$$\boldsymbol{R}_{r,k}^{(c)} = \begin{bmatrix} \sigma_r^2 & & 0 \\ & \ddots & \\ & & \sigma_r^2 \\ & & & \ddots \\ 0 & & & & \sigma_r^2 \end{bmatrix}, \qquad \boldsymbol{R}_{b,k}^{(c)} = \begin{bmatrix} \boldsymbol{R}_b & & 0 \\ & \ddots & \\ & & \boldsymbol{R}_b \\ & & & \ddots \\ 0 & & & & \boldsymbol{R}_b \end{bmatrix} \tag{5.13}$$

where $\boldsymbol{R}_{r,k}^{(c)} \in \mathbb{R}^{n_r n_y \times n_r n_y}$, $\boldsymbol{R}_{b,k}^{(c)} \in \mathbb{R}^{n_b n_y \times n_b n_y}$, and $n_y = {}_{(N+1)}C_2$. An infinite element in the measurement error covariance suggests that the corresponding measurement has zero information gain and it does not affect the updated estimate result. Finally, the measurement update step of the centralized extended information filter is formulated as:

$$\boldsymbol{Z}_{k|k}^{(c)c} = \boldsymbol{Z}_{k|k-1}^{(c)c} + \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(c)}\right)^{-1} \boldsymbol{H}_{s,k}^{(c)} \tag{5.14}$$

$$\hat{\boldsymbol{z}}_{k|k}^{(c)c} = \hat{\boldsymbol{z}}_{k|k-1}^{(c)c} + \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(c)}\right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c}\right). \tag{5.15}$$

Algorithm 4 summarizes the computational flow of the CF.

Algorithm 4: CF: Centralized Filter

At the k-th time step on the leader spacecraft (fusion center) ;

input : $\hat{\boldsymbol{z}}_{k-1}^{(c)c}, \boldsymbol{Z}_{k-1}^{(c)c}$

output: $\hat{\boldsymbol{z}}_{k}^{(c)c}, \boldsymbol{Z}_{k}^{(c)c}$

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(c)c} = \left(\boldsymbol{F}^{(c)} \left(\boldsymbol{Z}_{k-1}^{(c)c}\right)^{-1} \left(\boldsymbol{F}^{(c)}\right)^T + \boldsymbol{Q}^{(c)}\right)^{-1};$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(c)c} = \boldsymbol{Z}_{k|k-1}^{(c)c} \left(\boldsymbol{F}^{(c)} (\boldsymbol{Z}_{k-1}^{(c)c})^{-1} \hat{\boldsymbol{z}}_{k-1}^{(c)c} + \boldsymbol{G}^{(c)} \boldsymbol{u}_k^{(c)}\right);$$

Measurement update step ;

Collect the measurements $\boldsymbol{y}_k^{(c)}$ from all the spacecraft ;

Compute the Jacobian matrices $\boldsymbol{H}_{r,k}^{(c)}$ and $\boldsymbol{H}_{b,k}^{(c)}$;

$$\boldsymbol{H}_{r,k}^{(c)} = \left.\frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{r[c]}}}{\partial \boldsymbol{x}_k^{(c)}}\right|_{\boldsymbol{x}_k^{(c)}=\hat{\boldsymbol{x}}_{k|k-1}^{(c)}}, \quad \boldsymbol{H}_{b,k}^{(c)} = \left.\frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{b[c]}}}{\partial \boldsymbol{x}_k^{(c)}}\right|_{\boldsymbol{x}_k^{(c)}=\hat{\boldsymbol{x}}_{k|k-1}^{(c)}};$$

Compute the adjusted error covariance based on the sensing networks ;

$$\tilde{\boldsymbol{R}}_{s,k}^{(c)} = \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_k^s\right), \quad s \in \{r, b\} ;$$

$$\boldsymbol{Z}_{k|k}^{(c)c} = \boldsymbol{Z}_{k|k-1}^{(c)c} + \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(c)}\right)^{-1} \boldsymbol{H}_{s,k}^{(c)} ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(c)c} = \hat{\boldsymbol{z}}_{k|k-1}^{(c)c} + \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(c)}\right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c}\right) ;$$

5.2 Fully Decentralized Filter

In the CF introduced in the previous section, all the measurement information should be collected to a single node in the swarm system. The necessity of information collection possibly results in huge communication costs and low durability against spacecraft malfunction, especially for large swarm systems. In such a situation, estimation algorithms based on decentralized architectures are highly effective.

In this section, we consider a fully-decentralized filter, in which each spacecraft only estimates its own state with locally available information. Each spacecraft executes the estimation with measurements obtained by its own sensors and the state information shared from the adjacent spacecraft. As illustrated in Figure 5.3, each spacecraft only needs to interact with the adjacent spacecraft; hence the communication cost is significantly reduced, and high scalability is achieved. Figure 5.4 shows the algorithmic flow of the fully decentralized filter.

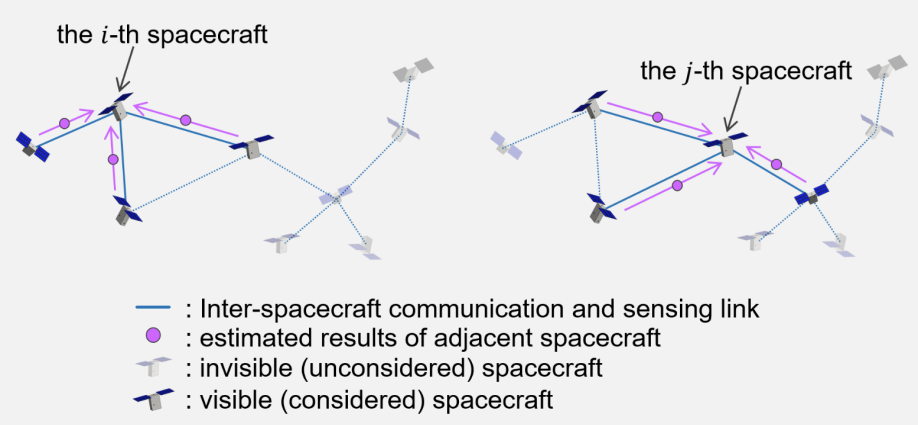


Figure 5.3: Inter-spacecraft communication in the fully decentralized filter

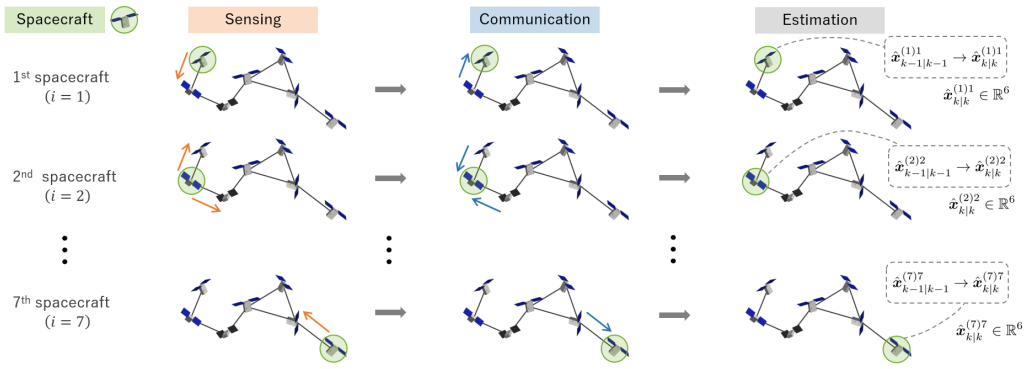


Figure 5.4: Algorithmic flow of the fully decentralized filter.

Firstly, we consider the prediction step of the fully decentralized filter on the i-th spacecraft, that is one of the spacecraft ($i \in \mathcal{V}$) in the swarm. As previously described, the i-th spacecraft only takes care of its own state, therefore, the prediction step is formulated as:

$$\boldsymbol{Z}^{(i)i}_{k|k-1} = \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}^{(i)i}_{k-1} \right)^{-1} \left(\boldsymbol{F}^{(i)} \right)^{T} + \boldsymbol{Q}^{(i)} \right)^{-1} \tag{5.16}$$

$$\hat{\boldsymbol{z}}^{(i)i}_{k|k-1} = \boldsymbol{Z}^{(i)i}_{k|k-1} \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}^{(i)i}_{k-1} \right)^{-1} \hat{\boldsymbol{z}}^{(i)}_{k-1} + \boldsymbol{G}^{(i)} \boldsymbol{u}^{(i)}_{k} \right). \tag{5.17}$$

Next, we consider the measurement update step. For the enhanced stability of update calculation, the range and bearing angle measurements are converted to the single position measurement as follows.

$$\hat{\boldsymbol{p}}^{(i)j}_{k|k-1} = \boldsymbol{g}\left(\hat{\boldsymbol{p}}^{(j)j}_{k|k-1}, \boldsymbol{y}^{(ij)j}_{k} \right) \tag{5.18}$$

$$\boldsymbol{g}\left(\boldsymbol{p}^{(j)}_{k|k-1}, \boldsymbol{y}^{(ij)}_{k} \right) = \boldsymbol{p}^{(j)}_{k|k-1} + \begin{bmatrix} y^{(ij)}_{r,k} \cos\phi^{(ij)}_{k} \cos\theta^{(ij)}_{k} \\ y^{(ij)}_{r,k} \cos\phi^{(ij)}_{k} \sin\theta^{(ij)}_{k} \\ y^{(ij)}_{r,k} \sin\phi^{(ij)}_{k} \end{bmatrix} \tag{5.19}$$

where $\boldsymbol{y}^{(ij)j}_{k}$ is the relative measurements between the i-th and the j-th spacecraft which the j-th spacecraft obtains. The physical meaning of Eqs. (5.18) and (5.19) is visualized in Figure 5.5.

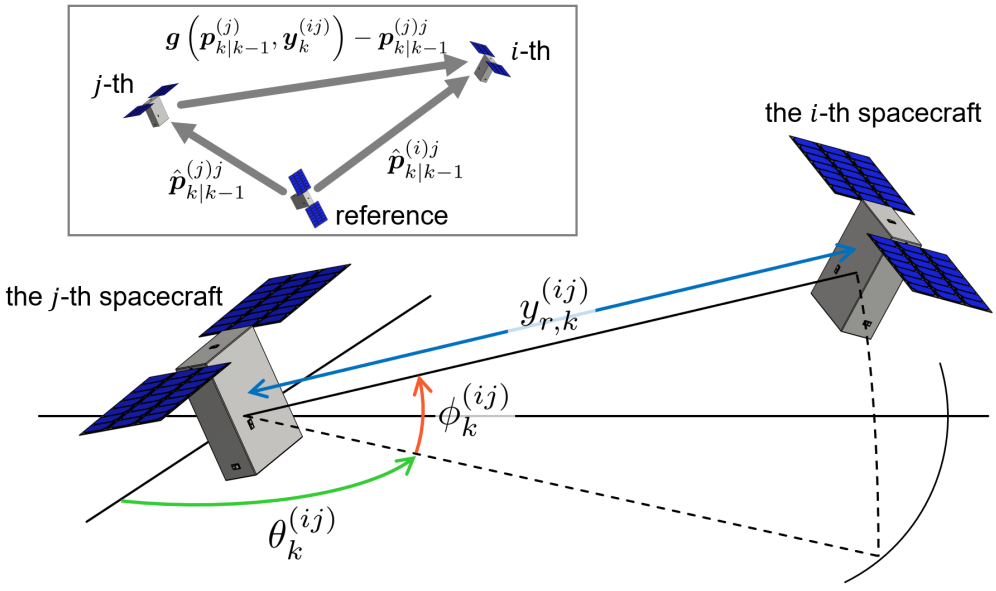


Figure 5.5: Position measurement conversion between two spacecraft in the fully decentralized filter.

The corresponding error covariance matrix of the estimated i-th position $\tilde{\boldsymbol{P}}^{(i)j}_{p,k|k-1} \in \mathbb{R}^{3\times3}$ is computed as:

$$\tilde{\boldsymbol{P}}^{(i)j}_{p,k|k-1} = \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{p}^{(j)}_k}\right) \boldsymbol{P}^{(j)j}_{p,k|k-1} \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{p}^{(j)}_k}\right)^T + \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{y}^{(ij)}_k}\right) \boldsymbol{R}^{(ij)}_k \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{y}^{(ij)}_k}\right)^T \tag{5.20}$$

$$\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{p}^{(j)}_k} \triangleq \left.\frac{\partial \boldsymbol{g}\left(\boldsymbol{p}^{(j)}_k, \boldsymbol{y}^{(ij)}_k\right)}{\partial \boldsymbol{p}^{(j)}_k}\right|_{\boldsymbol{p}^{(j)}_k = \hat{\boldsymbol{p}}^{(j)j}_k} \tag{5.21}$$

$$\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{y}^{(ij)}_k} \triangleq \left.\frac{\partial \boldsymbol{g}\left(\boldsymbol{p}^{(j)}_k, \boldsymbol{y}^{(ij)}_k\right)}{\partial \boldsymbol{y}^{(ij)}_k}\right|_{\boldsymbol{p}^{(j)}_k = \hat{\boldsymbol{p}}^{(j)j}_k} \tag{5.22}$$

where $\boldsymbol{P}^{(j)j}_{p,k|k-1}$ is the prior error covariance matrix of the j-th spacecraft state estimated on the j-th spacecraft. $\boldsymbol{R}^{(ij)}_k \in \mathbb{R}^{3\times3}$ is the error covariance matrix of the range and bearing angle measurements between the i-th and the j-th spacecraft, which is defined as:

$$\boldsymbol{R}^{(ij)}_k = \begin{bmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_b^2 \end{bmatrix}. \tag{5.23}$$

As the superscripts of Eqs. (5.18) to (5.22) suggest, they are calculated on the j-th spacecraft and transferred to the i-th spacecraft through the inter-spacecraft communication link. Eq. (5.21) is calculated as:

$$\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{p}^{(j)}_k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.24}$$

and the first row elements of Eq. (5.22) is calculated as:

$$\frac{\partial g^{(ij)j}[1]}{\partial y^{(ij)}_{r,k}} = \frac{\partial}{\partial y^{(ij)}_{r,k}}\left[p^{(j)}_{x,k} + y^{(ij)}_{r,k}\cos\phi^{(ij)}_k \cos\theta^{(ij)}_k\right] = \cos\phi^{(ij)}_k \cos\theta^{(ij)}_k \tag{5.25}$$

$$\frac{\partial g^{(ij)j}[1]}{\partial \theta^{(ij)}_k} = -y^{(ij)}_{r,k}\cos\phi^{(ij)}_k \sin\theta^{(ij)}_k \tag{5.26}$$

$$\frac{\partial g^{(ij)j}[1]}{\partial \phi^{(ij)}_k} = -y^{(ij)}_{r,k}\sin\phi^{(ij)}_k \cos\theta^{(ij)}_k \tag{5.27}$$

and the second row elements of Eq. (5.22) is calculated as:

$$\frac{\partial g^{(ij)j}[2]}{\partial y_{r,k}^{(ij)}} = \frac{\partial}{\partial y_{r,k}^{(ij)}}\left[p_{y,k}^{(j)} + y_{r,k}^{(ij)}\cos\phi_k^{(ij)}\sin\theta_k^{(ij)}\right] = \cos\phi_k^{(ij)}\sin\theta_k^{(ij)} \tag{5.28}$$

$$\frac{\partial g^{(ij)j}[2]}{\partial \theta_k^{(ij)}} = y_{r,k}^{(ij)}\cos\phi_k^{(ij)}\cos\theta_k^{(ij)} \tag{5.29}$$

$$\frac{\partial g^{(ij)j}[2]}{\partial \phi_k^{(ij)}} = -y_{r,k}^{(ij)}\sin\phi_k^{(ij)}\sin\theta_k^{(ij)} \tag{5.30}$$

and the third row elements of Eq. (5.22) is calculated as:

$$\frac{\partial g^{(ij)j}[3]}{\partial y_{r,k}^{(ij)}} = \frac{\partial}{\partial y_{r,k}^{(ij)}}\left[p_{z,k}^{(j)} + y_{r,k}^{(ij)}\sin\phi_k^{(ij)}\right] = \sin\phi_k^{(ij)} \tag{5.31}$$

$$\frac{\partial g^{(ij)j}[3]}{\partial \theta_k^{(ij)}} = 0 \tag{5.32}$$

$$\frac{\partial g^{(ij)j}[3]}{\partial \phi_k^{(ij)}} = y_{r,k}^{(ij)}\cos\phi_k^{(ij)}. \tag{5.33}$$

Finally, the information contribution matrix $\boldsymbol{U}_{k|k-1}^{(i)j}$ and the information contribution vector $\hat{\boldsymbol{u}}_{k|k-1}^{(i)j}$ are computed as:

$$\boldsymbol{U}_{k|k-1}^{(i)j} = \begin{bmatrix} \left(\tilde{\boldsymbol{P}}_{p,k|k-1}^{(i)j}\right)^{-1} & \boldsymbol{O}_{3\times3} \\ \boldsymbol{O}_{3\times3} & \boldsymbol{O}_{3\times3} \end{bmatrix} \tag{5.34}$$

$$\hat{\boldsymbol{u}}_{k|k-1}^{(i)j} = \begin{bmatrix} \left(\tilde{\boldsymbol{P}}_{p,k|k-1}^{(i)j}\right)^{-1}\hat{\boldsymbol{p}}_{k|k-1}^{(i)j} \\ \boldsymbol{O}_{3\times1} \end{bmatrix} \tag{5.35}$$

where the velocity-related elements are set as zero since range and bearing angle measurements obtained at a single time step have no information gain for velocity estimates.

We consider the measurement update step based on the converted position measurements expressed as Eqs. (5.34) and (5.35). The i-th spacecraft obtains the measurement information from the adjacent spacecraft. To acquire the position measurement, both range and bearing angle measurements must be available. Let the set of the adjacent spacecraft that can get range and bearing angle measurements with respect to the i-th spacecraft at the k-th time step be $\mathcal{V}_{i,k}^{rb}$, that is defined as:

$$\mathcal{V}_{i,k}^{rb} = \left\{ j \in \mathcal{V} \mid (i,j) \in \mathcal{E}_{i,k}^{r} \cap \mathcal{E}_{i,k}^{b} \right\} \tag{5.36}$$

where $\mathcal{E}^r_{i,k} \cap \mathcal{E}^b_{i,k}$ represents the set of inter-spacecraft links where both range and bearing angle measurements are available.

Then, we consider fusing the prior estimate with the information contribution. It is essential to note that the estimates $\hat{\boldsymbol{z}}^{(i)i}_{k|k-1}, \boldsymbol{Z}^{(i)i}_{k|k-1}$ and $\hat{\boldsymbol{u}}^{(i)j}_{k|k-1}, \boldsymbol{U}^{(i)j}_{k|k-1}$, for the random variable $\boldsymbol{z}^{(i)}_k$, are not independent. If the i-th and the j-th spacecraft have communicated measurement information before the k-th time step and execute the measurement update, their estimates are correlated. In the fully decentralized architecture, the cross-correlation terms for any two spacecraft are not tracked. Therefore, the measurement information is possibly overcounted if the information contribution probability is purely merged into the prior estimates. It possibly leads to estimation inconsistency and induces estimate divergence.

The Covariance Intersection (CI) is highly effective in managing this issue. The CI is a fusion method taking a convex combination of the means and covariances of the source estimates. [51] The main advantage of the CI is that it certainly provides consistent fusion results even if there exists an unknown correlation between source estimates. Then, the measurement update step based on the CI is formulated as:

$$\boldsymbol{Z}^{(i)i}_{k|k} = \omega_i \boldsymbol{Z}^{(i)i}_{k|k-1} + \sum_{j \in \mathcal{V}^{rb}_{i,k}} \omega_j \boldsymbol{U}^{(i)j}_{k|k-1} \tag{5.37}$$

$$\hat{\boldsymbol{z}}^{(i)i}_{k|k} = \omega_i \hat{\boldsymbol{z}}^{(i)i}_{k|k-1} + \sum_{j \in \mathcal{V}^{rb}_{i,k}} \omega_j \hat{\boldsymbol{u}}^{(i)j}_{k|k-1} \tag{5.38}$$

where ω_i, ω_j are the CI parameters satisfying the condition: $\omega_i + \sum_{j \in \mathcal{V}^{rb}_{i,k}} \omega_j = 1$. These coefficient parameters are calculated to minimize the trace of the posterior covariance matrix $(\boldsymbol{Z}^{(i)i}_{k|k})^{-1}$.

Algorithm 5 shows the computational detail of the fully decentralized filter (FDF), and Algorithm 6 represents one of the robust fully decentralized filters (R-FDF). The main difference between the FDF and the R-FDF is whether the covariance information of each spacecraft is transferred or not. In the FDF, each spacecraft sends only measurement information; hence, the communication cost is kept low. On the other hand, the R-FDF also sends the covariance information in the form of the relative position measurements for more consistent estimation though it consumes communication resources.

Algorithm 5: FDF: Fully Decentralized Filter

At the k-th time step on the i-th spacecraft;

(The following process equally runs on all the spacecraft.);

input : $\hat{\boldsymbol{z}}_{k-1}^{(i)i}, \boldsymbol{Z}_{k-1}^{(i)i}$

output: $\hat{\boldsymbol{z}}_{k}^{(i)i}, \boldsymbol{Z}_{k}^{(i)i}$

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(i)i} = \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \left(\boldsymbol{F}^{(i)} \right)^{T} + \boldsymbol{Q}^{(i)} \right)^{-1} ;$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} = \boldsymbol{Z}_{k|k-1}^{(i)i} \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \hat{\boldsymbol{z}}_{k-1}^{(i)} + \boldsymbol{G}^{(i)} \boldsymbol{u}_{k}^{(i)} \right);$$

Measurement update step ;

Receive the measurements $\boldsymbol{y}_{k}^{(ij)j}$ from the adjacent spacecraft ;

Compute the relative position measurements ;

$$\hat{\boldsymbol{p}}_{k|k-1}^{(i)j} = \boldsymbol{g} \left(\boldsymbol{p}_{k|k-1}^{(i)j}, \boldsymbol{y}_{k}^{(ij)j} \right) ;$$

$$\tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)i} = \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{y}_{k}^{(ij)}} \right) \boldsymbol{R}_{k}^{(ij)} \left(\frac{\partial \boldsymbol{g}^{(ij)j}}{\partial \boldsymbol{y}_{k}^{(ij)}} \right)^{T} ;$$

Convert $\hat{\boldsymbol{p}}_{k|k-1}^{(i)j}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(i)j}$ to $\hat{\boldsymbol{u}}_{k|k-1}^{(i)j}, \boldsymbol{U}_{k|k-1}^{(i)j}$;

Execute the covariance intersection ;

$$\boldsymbol{Z}_{k|k}^{(i)i} = \omega_i \boldsymbol{Z}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \boldsymbol{U}_{k|k-1}^{(i)j} ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \omega_i \hat{\boldsymbol{z}}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \hat{\boldsymbol{u}}_{k|k-1}^{(i)j} ;$$

Algorithm 6: R-FDF: Robust Fully Decentralized Filter

At the k-th time step on the i-th spacecraft;

(The following process equally runs on all the spacecraft.);

input : $\hat{\boldsymbol{z}}_{k-1}^{(i)i}, \boldsymbol{Z}_{k-1}^{(i)i}$

output: $\hat{\boldsymbol{z}}_{k}^{(i)i}, \boldsymbol{Z}_{k}^{(i)i}$

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(i)i} = \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \left(\boldsymbol{F}^{(i)} \right)^{T} + \boldsymbol{Q}^{(i)} \right)^{-1};$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} = \boldsymbol{Z}_{k|k-1}^{(i)i} \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \hat{\boldsymbol{z}}_{k-1}^{(i)} + \boldsymbol{G}^{(i)} \boldsymbol{u}_{k}^{(i)} \right);$$

Measurement update step ;

Compute the relative position measurements for the adjacent spacecraft ;

$$\hat{\boldsymbol{p}}_{k|k-1}^{(j)i} = \boldsymbol{g}\left(\boldsymbol{p}_{k|k-1}^{(i)i}, \boldsymbol{y}_{k}^{(ji)i} \right) ;$$

$$\tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)i} = \left(\frac{\partial \boldsymbol{g}^{(ji)i}}{\partial \boldsymbol{p}_{k}^{(i)}} \right) \boldsymbol{P}_{p,k|k-1}^{(i)i} \left(\frac{\partial \boldsymbol{g}^{(ji)i}}{\partial \boldsymbol{p}_{k}^{(i)}} \right)^{T} + \left(\frac{\partial \boldsymbol{g}^{(ji)i}}{\partial \boldsymbol{y}_{k}^{(ji)}} \right) \boldsymbol{R}_{k}^{(ji)} \left(\frac{\partial \boldsymbol{g}^{(ji)i}}{\partial \boldsymbol{y}_{k}^{(ji)}} \right)^{T} ;$$

Transfer $\hat{\boldsymbol{p}}_{k|k-1}^{(j)i}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)i}$ to the j-th spacecraft ($\forall j \in \mathcal{V}_{i,k}^{rb}$) ;

Receive $\hat{\boldsymbol{p}}_{k|k-1}^{(i)j}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(i)j}$ from the j-th spacecraft ($\forall j \in \mathcal{V}_{i,k}^{rb}$) ;

Convert $\hat{\boldsymbol{p}}_{k|k-1}^{(i)j}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(i)j}$ to $\hat{\boldsymbol{u}}_{k|k-1}^{(i)j}, \boldsymbol{U}_{k|k-1}^{(i)j}$;

Execute the covariance intersection ;

$$\boldsymbol{Z}_{k|k}^{(i)i} = \omega_i \boldsymbol{Z}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \boldsymbol{U}_{k|k-1}^{(i)j} ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \omega_i \hat{\boldsymbol{z}}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \hat{\boldsymbol{u}}_{k|k-1}^{(i)j} ;$$

5.3 Partially Decentralized Filter

The partially decentralized filter (PDF) is the intermediate approach between the CF and the FDF. In the PDF, each spacecraft estimates the subset of states, including its own state and the states of the adjacent spacecraft, as shown in Figure 5.6. The states of spacecraft covered by the red region are estimated similarly to the CF. On the other hand, support spacecraft outside the red area but connected by one-hop communication links are utilized to leverage measurements between the estimated spacecraft and the support spacecraft. By estimating the states of several spacecraft simultaneously, the PDF can explicitly consider the correlations among spacecraft. Figure 5.7 shows the algorithmic flow of the PDF.

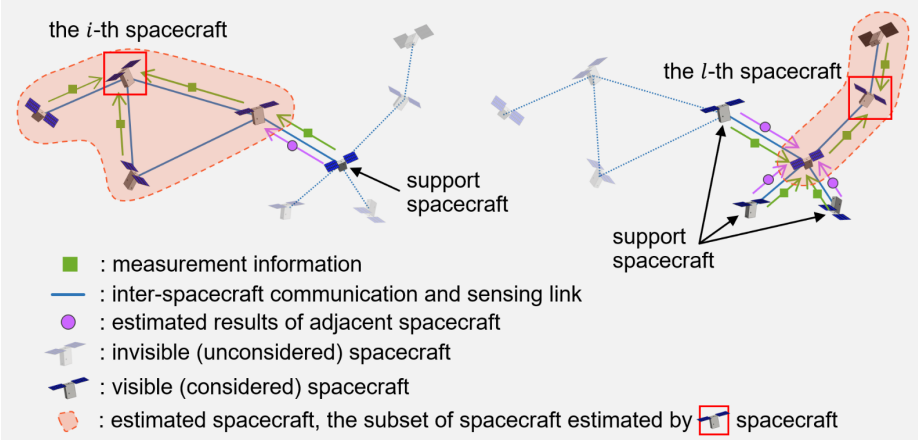


Figure 5.6: Inter-spacecraft communication in the partially decentralized filter

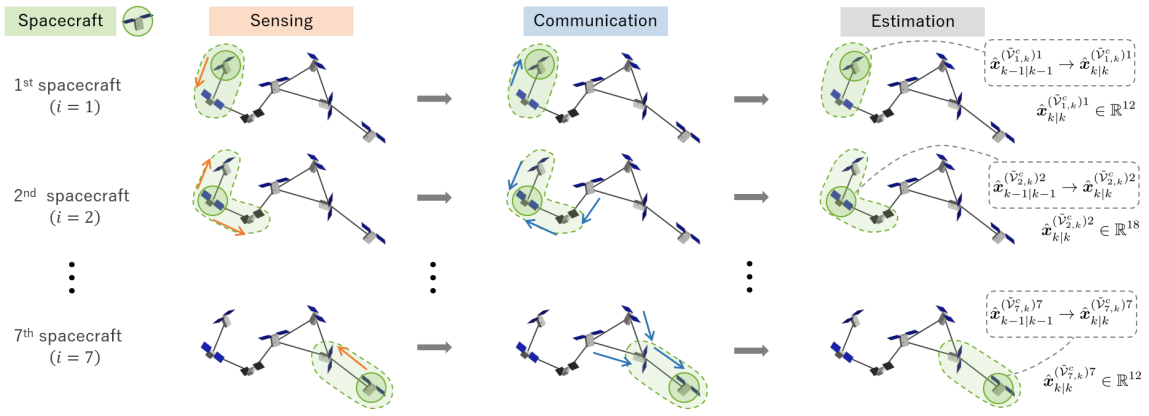


Figure 5.7: Algorithmic flow of the partially decentralized filter.

The PDF manages a collection of spacecraft states defined as $\boldsymbol{x}^{(\tilde{\mathcal{V}}^c_{i,k})}$, where the space-

craft set $\tilde{\mathcal{V}}_{i,k}^c$ is defined in Section 2.5. On the estimator of the i-th spacecraft, the propagation step is formulated as:

$$\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \right)^{-1} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^T + \boldsymbol{Q}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^{-1} \tag{5.39}$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \right)^{-1} \boldsymbol{z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} + \boldsymbol{G}^{(\tilde{\mathcal{V}}_{i,k}^c)} \boldsymbol{u}_k^{(\tilde{\mathcal{V}}_{i,k}^c)} \right). \tag{5.40}$$

where the aggregated matrices $\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)}$, $\boldsymbol{G}^{(\tilde{\mathcal{V}}_{i,k}^c)}$, $\boldsymbol{Q}^{(\tilde{\mathcal{V}}_{i,k}^c)}$ are defined as:

$$\boldsymbol{X}^{(\tilde{\mathcal{V}}_{i,k}^c)} = \begin{bmatrix} \ddots & & 0 \\ & \boldsymbol{X}^{(j)} & \\ 0 & & \ddots \end{bmatrix}, \qquad j \in \tilde{\mathcal{V}}_{i,k}^c, \qquad \boldsymbol{X} \in \{\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{Q}\}. \tag{5.41}$$

Next, we consider the measurement update step. For the preparation, a few node and edge sets are defined as:

$$\mathcal{E}_{i,k}^c \triangleq \left\{ (j,l) \in \mathcal{E}_k^c \mid j,l \in \tilde{\mathcal{V}}_{i,k}^c \right\} \tag{5.42}$$

$$\tilde{\mathcal{E}}_{i,k}^c \triangleq \left\{ (j,l) \in \mathcal{E}_k^c \mid j \in \mathcal{V}_{i,k}^c, l \in \tilde{\mathcal{M}}_{i,k}^c \right\} \tag{5.43}$$

$$\tilde{\mathcal{M}}_{i,k}^c \triangleq \left\{ j \in \mathcal{V} \mid (j,l) \in \mathcal{E}_k^c, j \notin \tilde{\mathcal{V}}_{i,k}^c, l \in \tilde{\mathcal{V}}_{i,k}^c \right\}. \tag{5.44}$$

In the case of Figure 5.8, the node and edge sets are represented as:

$$\begin{aligned} \mathcal{E}_{i,k}^c &= \{(i,j_1),(i,j_2),(i,j_3),(j_2,j_3)\} \\ \tilde{\mathcal{E}}_k^c &= \{(j_3,j_4)\} \\ \mathcal{V}_{i,k}^c &= \{j_1,j_2,j_3\} \\ \tilde{\mathcal{V}}_{i,k}^c &= \{i,j_1,j_2,j_3\} \\ \tilde{\mathcal{M}}_{i,k}^c &= \{j_4\}. \end{aligned}$$

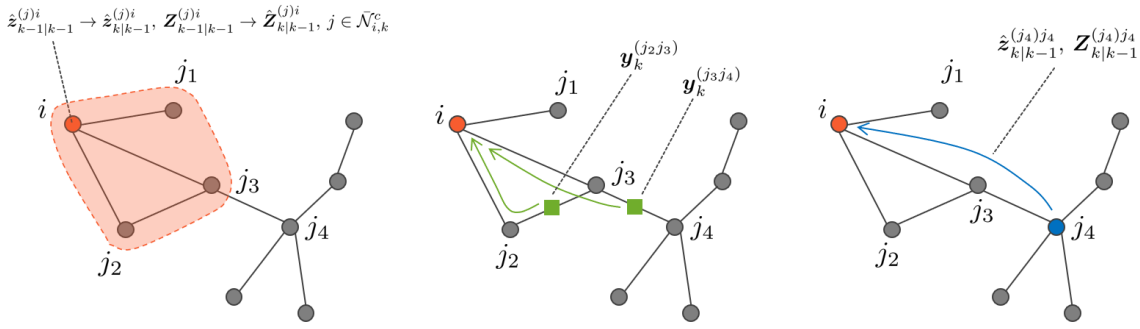


Figure 5.8: Estimation sequence of the partial-state decentralized filter

As with the CF, we firstly consider that the complete sensing graph is adjusted by modifying the error covariance matrix of measurements. The adjusted covariance matrix is computed in the same manner with Eq. (5.12) as:

$$\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} = \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^{s}\right), \qquad s \in \{r, b\} \tag{5.45}$$

where the available sensing network of each spacecraft differs as shown in Figure 5.9.

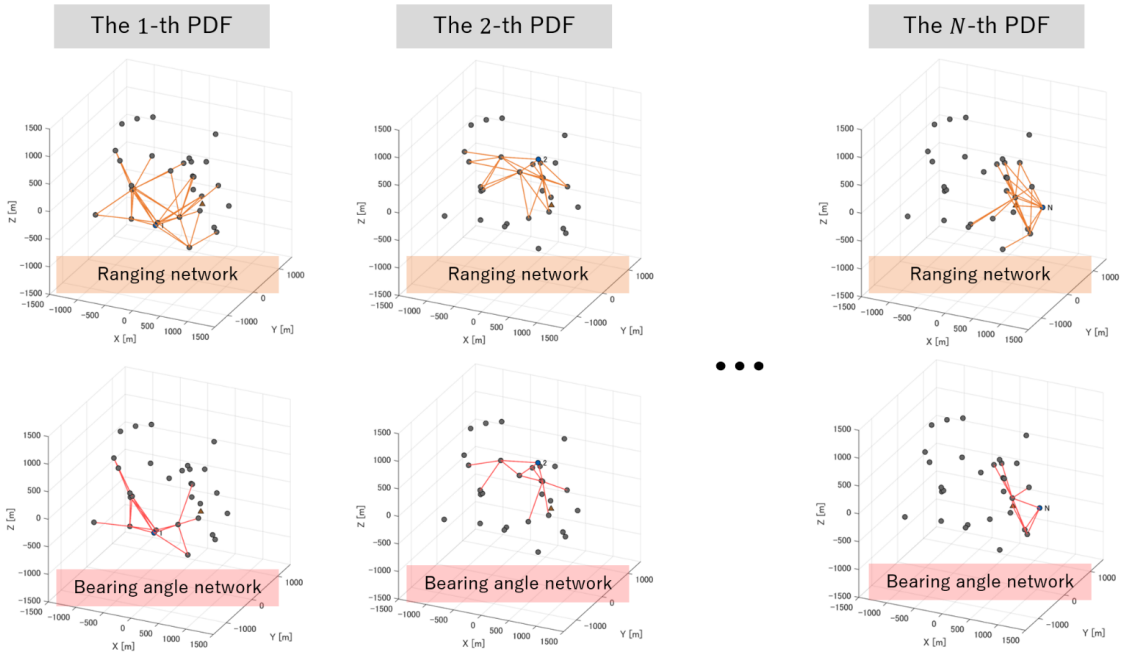


Figure 5.9: Sensing networks in the partially decentralized filter

Therefore, the information contribution matrix $\boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}$ and vector $\boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}$ are formulated as:

$$\boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \boldsymbol{H}_{s,k}^{(c)} \tag{5.46}$$

$$\boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c}\right) \tag{5.47}$$

In an actual scenario, the elements of $\boldsymbol{H}_{s,k}^{(c)}$, $\hat{\boldsymbol{y}}_{s,k|k-1}^{(c)}$, $\hat{\boldsymbol{x}}_{k|k-1}^{(c)c}$ corresponding to spacecraft that are not included in $\tilde{\mathcal{V}}_{i,k}^{c}$ cannot be calculated. However, they are ignored when multiplying the inverse of $\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}$, so they can be replaced with some random values, such as zero.

In addition to the inter-spacecraft measurement information expressed as Eqs. (5.46) and (5.47), the measurements related to the support spacecraft ($l \in \tilde{\mathcal{M}}_{i,k}^{c}$) are also available.

They are integrated as additional likelihood probability in the CI manner since the cross-correlation is unknown. Finally, the measurement update step of the partially decentralized filter is summarized as:

$$\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_i \left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{k}^{c}} \omega_p \boldsymbol{U}_{k|k-1}^{(j)l} \tag{5.48}$$

$$\boldsymbol{z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_i \left(\boldsymbol{z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{k}^{c}} \omega_p \boldsymbol{u}_{k|k-1}^{(j)l} \tag{5.49}$$

where ω_i, ω_p are the CI parameters satisfying the condition: $\omega_i + \sum_{p=1}^{n(\tilde{\mathcal{E}}_{k}^{c})} \omega_p = 1$. These coefficient parameters are calculated to minimize the trace of the posterior covariance matrix $(\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i})^{-1}$.

In the partially decentralized filter, the set of estimated spacecraft may change over time along with the update of the communication network. We consider the network update as shown in Figure 5.10. At the k-th time step, the adjacent set of the i-th spacecraft changes from $\{1,2,3,4\}$ to $\{1,2,4,5\}$. In this case, the estimator on the i-th spacecraft stops estimating the state of the 3-rd spacecraft and starts to track the state of the 5-th spacecraft alternatively. The estimated information vector is updated as:

$$\boldsymbol{z}_{k}^{1(-)} = \begin{bmatrix} \boldsymbol{z}_{k}^{(1)1} & \boldsymbol{z}_{k}^{(2)1} & \boldsymbol{z}_{k}^{(3)1} & \boldsymbol{z}_{k}^{(4)1} \end{bmatrix}^{T} \tag{5.50}$$

$$\boldsymbol{z}_{k}^{1(+)} = \begin{bmatrix} \boldsymbol{z}_{k}^{(1)1} & \boldsymbol{z}_{k}^{(2)1} & \boldsymbol{z}_{k}^{(4)1} & \boldsymbol{z}_{k}^{(5)1} \end{bmatrix}^{T}, \quad \boldsymbol{z}_{k}^{(5)1} = \boldsymbol{z}_{k|k-1}^{(5)5} \tag{5.51}$$

and the information matrix is updated as:

$$\boldsymbol{Z}_{k}^{1(-)} = \begin{bmatrix} \boldsymbol{Z}_{k}^{(11)1} & \boldsymbol{Z}_{k}^{(12)1} & \boldsymbol{Z}_{k}^{(13)1} & \boldsymbol{Z}_{k}^{(14)1} \\ \boldsymbol{Z}_{k}^{(21)1} & \boldsymbol{Z}_{k}^{(22)1} & \boldsymbol{Z}_{k}^{(23)1} & \boldsymbol{Z}_{k}^{(24)1} \\ \boldsymbol{Z}_{k}^{(31)1} & \boldsymbol{Z}_{k}^{(32)1} & \boldsymbol{Z}_{k}^{(33)1} & \boldsymbol{Z}_{k}^{(34)1} \\ \boldsymbol{Z}_{k}^{(41)1} & \boldsymbol{Z}_{k}^{(42)1} & \boldsymbol{Z}_{k}^{(43)1} & \boldsymbol{Z}_{k}^{(44)1} \end{bmatrix} \tag{5.52}$$

$$\boldsymbol{Z}_{k}^{1(+)} = \begin{bmatrix} \boldsymbol{Z}_{k}^{(11)1} & \boldsymbol{Z}_{k}^{(12)1} & \boldsymbol{Z}_{k}^{(14)1} & \boldsymbol{O} \\ \boldsymbol{Z}_{k}^{(21)1} & \boldsymbol{Z}_{k}^{(22)1} & \boldsymbol{Z}_{k}^{(24)1} & \boldsymbol{O} \\ \boldsymbol{Z}_{k}^{(41)1} & \boldsymbol{Z}_{k}^{(42)1} & \boldsymbol{Z}_{k}^{(44)1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{Z}_{k}^{(55)1} \end{bmatrix}, \quad \boldsymbol{Z}_{k}^{(55)1} = \boldsymbol{Z}_{k|k-1}^{(55)5} \tag{5.53}$$

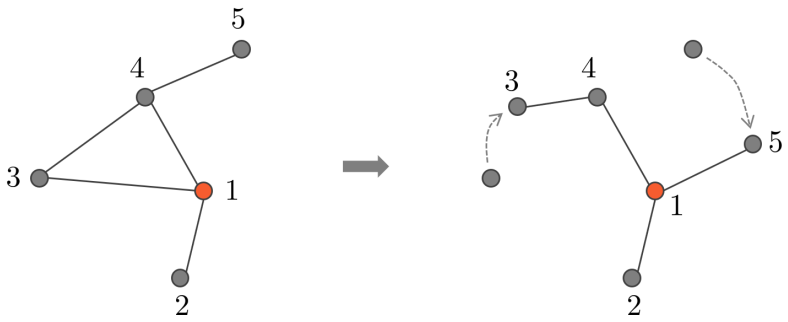


Figure 5.10: Network update in the partially decentralized filter

We consider the two types of algorithms: the partially decentralized filter (PDF) defined as Algorithm 7 and the robust partially decentralized filter (R-PDF) defined as Algorithm 8.

Algorithm 7: PDF: Partially Decentralized Filter

At the k-th time step on the i-th spacecraft ;

(The following process equally runs on all the spacecraft.) ;

input : $\hat{\boldsymbol{z}}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}, \boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}$

output: $\hat{\boldsymbol{z}}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}, \boldsymbol{Z}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}$

Grouping step;

Check the update of the communication network. ;

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} \right)^{-1} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right)^{T} + \boldsymbol{Q}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right)^{-1} ;$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} \right)^{-1} \hat{\boldsymbol{z}}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{G}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right) ;$$

Measurement update step ;

Compute the adjusted error covariance based on the local sensing networks ;

$$\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} = \Gamma \left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^{s} \cup \tilde{\mathcal{E}}_{i,k}^{s} \right), \quad s \in \{r, b\} ;$$

Compute the information contribution in the estimated subset ;

$$\boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)} \right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right)^{-1} \boldsymbol{H}_{s,k}^{(c)} ;$$

$$\boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)} \right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} \right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c} \right) ;$$

Execute the measurement update ;

$$\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} ;$$

Algorithm 8: R-PDF: Robust Partially Decentralized Filter

At the k-th time step on the i-th spacecraft ;

(The following process equally runs on all the spacecraft.) ;

input : $\hat{\boldsymbol{z}}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i}, \boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i}$

output: $\hat{\boldsymbol{z}}_{k}^{(\tilde{\mathcal{V}}_{i,k}^c)i}, \boldsymbol{Z}_{k}^{(\tilde{\mathcal{V}}_{i,k}^c)i}$

Grouping step;

Check the update of the communication network. ;

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \right)^{-1} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^T + \boldsymbol{Q}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^{-1} ;$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \left(\boldsymbol{F}^{(\tilde{\mathcal{V}}_{i,k}^c)} \left(\boldsymbol{Z}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} \right)^{-1} \hat{\boldsymbol{z}}_{k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} + \boldsymbol{G}^{(\tilde{\mathcal{V}}_{i,k}^c)} \boldsymbol{u}_k^{(\tilde{\mathcal{V}}_{i,k}^c)} \right) ;$$

Measurement update step ;

Compute the adjusted error covariance based on the local sensing networks ;

$$\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^c)} = \Gamma \left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^s \right), \quad s \in \{r, b\} ;$$

Compute the information contribution in the estimated subset ;

$$\boldsymbol{U}_k^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)} \right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^{-1} \boldsymbol{H}_{s,k}^{(c)} ;$$

$$\boldsymbol{u}_k^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \textstyle\sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)} \right)^T \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^c)} \right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c} \right) ;$$

Compute the information contribution of the support spacecraft $(l \in \tilde{\mathcal{M}}_{i,k}^c, j \in \tilde{\mathcal{E}}_k^c)$;

$$\hat{\boldsymbol{p}}_{k|k-1}^{(j)l} = \boldsymbol{g} \left(\boldsymbol{p}_{k|k-1}^{(l)l}, \boldsymbol{y}_k^{(jl)l} \right) ;$$

$$\tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l} = \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_k^{(l)}} \right) \boldsymbol{P}_{p,k|k-1}^{(l)l} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_k^{(l)}} \right)^T + \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_k^{(jl)}} \right) \boldsymbol{R}_k^{(jl)} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_k^{(jl)}} \right)^T ;$$

Convert $\hat{\boldsymbol{p}}_{k|k-1}^{(j)l}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l}$ to $\hat{\boldsymbol{u}}_{k|k-1}^{(j)l}, \boldsymbol{U}_{k|k-1}^{(j)l}$;

Execute the covariance intersection ;

$$\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \omega_i \left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} + \boldsymbol{U}_k^{(\tilde{\mathcal{V}}_{i,k}^c)} \right) + \textstyle\sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^c} \omega_p \boldsymbol{U}_{k|k-1}^{(j)l} ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^c)i} = \omega_i \left(\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^c)i} + \boldsymbol{u}_k^{(\tilde{\mathcal{V}}_{i,k}^c)} \right) + \textstyle\sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^c} \omega_p \boldsymbol{u}_{k|k-1}^{(j)l} ;$$

5.4 Performance Comparison of Existing Algorithms

In this section, we evaluate the performances of the estimation algorithms with the different levels of decentralization: the centralized filter, the fully decentralized filter, and the partially decentralized filter, as summarized in Table 5.1. The detailed information on simulation conditions is summarized in Chapter 4. In the performance verification, we evaluate the convergence rate, the estimation accuracy, and the communication load.

Table 5.1: Estimation algorithms with the different levels of decentralization

No.	Acronym	Full name	Detail
(i)	CF	Centralized Filter	Algorithm 4
(ii)	R-FDF	Robust Fully Decentralized Filter	Algorithm 6
(iii)	R-PDF	Robust Partially Decentralized Filter	Algorithm 8

Figure 5.11 shows the relationship between the connection rate and the network threshold. We evaluate various metrics with respect to the average connection rate instead of the network threshold.

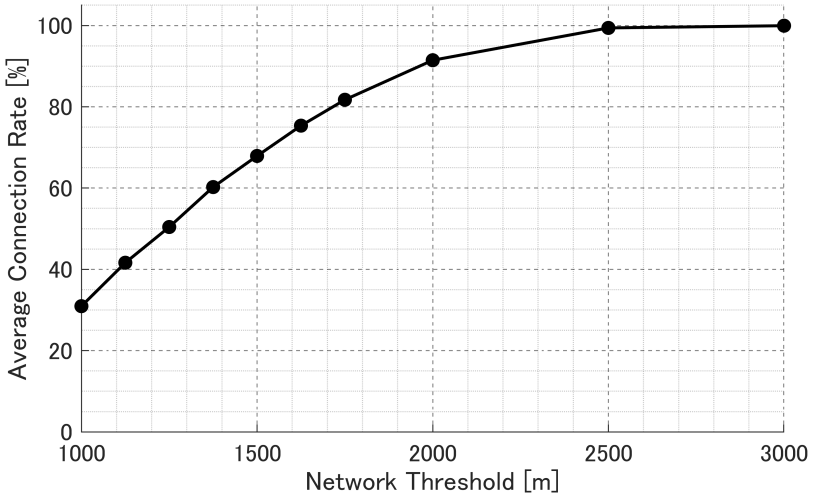


Figure 5.11: Network connection rate (the performance comparison of the existing algorithms).

Figure 5.12 shows the convergence rate with respect to the average connection rate defined as Eq. (2.35). In this evaluation, if the distance between two spacecraft is shorter than the network threshold, there exists the communication or sensing link between them

regardless of any other limitation, such as communication window number. As shown in Figure 5.12, all the filters achieve a 100% convergence rate when the average connection rate is higher than 90%. However, when the connection rate is lower than 90%, the performances differ among the three filters. The fully decentralized filter achieves a 100% convergence rate regardless of the average connection rate. On the other hand, the centralized and partially decentralized filters provide a lower convergence rate as the network connection rate decreases.

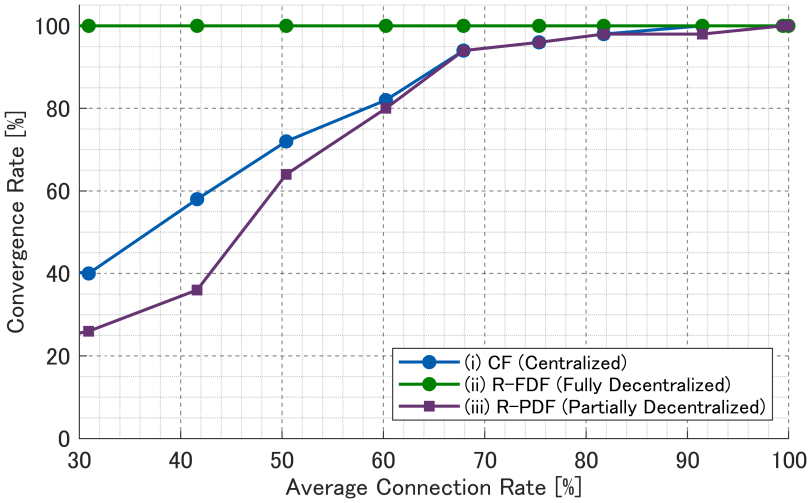


Figure 5.12: Convergence rate (the performance comparison of the existing algorithms).

We consider the estimation results with the same conditions for the fair performance comparison. Figure 5.13 shows the gross convergence flags ${}^{m}\mu_{x_f}$ of Eq. (4.4).

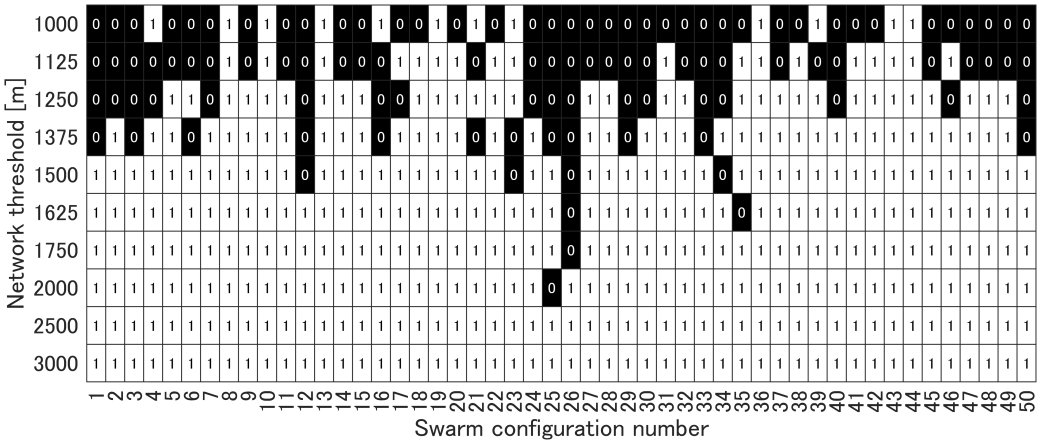
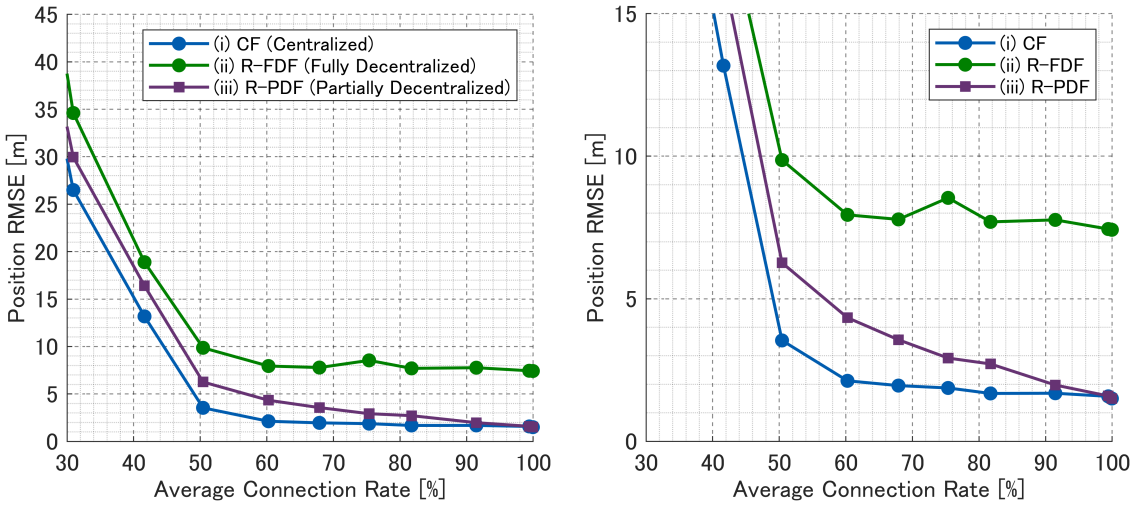


Figure 5.13: Gross convergence flags (the performance comparison of the existing algorithms).

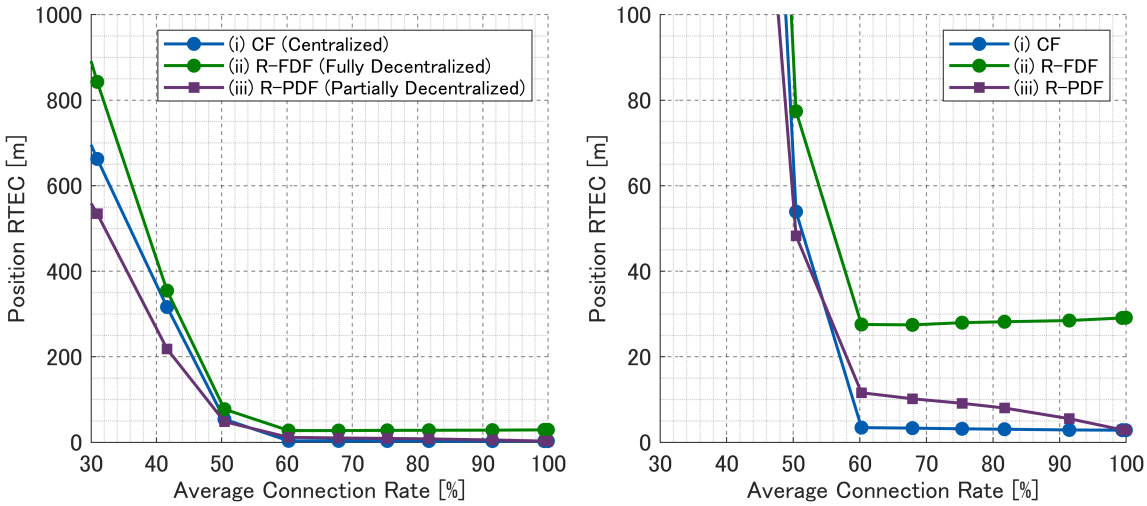
Regarding estimation accuracy, as shown in Figures 5.14 and 5.15, the centralized filter

provides the best results. On the other hand, the estimation accuracy of the fully decentralized filter is lower than the other two filters though it achieves the highest stability. Another remarkable point is that the partially decentralized filter provides the same result as the centralized filter when the network is fully connected. In that condition, the algorithm of the partially decentralized filter gets identical to that of the centralized filter.



(a) Overall view **(b)** Extended view

Figure 5.14: Root mean square errors of the estimated positions (the performance comparison of the existing algorithms).



(a) Overall view **(b)** Extended view

Figure 5.15: Root of the trace of the position error covariance matrix (the performance comparison of the existing algorithms).

As shown in Figure 5.16, the centralized filter takes much higher costs for inter-spacecraft communication than the other decentralized filters. In the framework of the centralized filter, all the information related to the estimation process should be collected to the leader spacecraft by inter-spacecraft communication; therefore, it leads to the high communication cost. This is one of the main reasons to make the estimation process decentralized in large multi-agent systems such as spacecraft swarms. When comparing the fully and the partially decentralized filters, the fully decentralized filter interacts with fewer adjacent spacecraft than the partially decentralized filter. Therefore, the fully decentralized filter needs the smallest costs for inter-spacecraft communication.

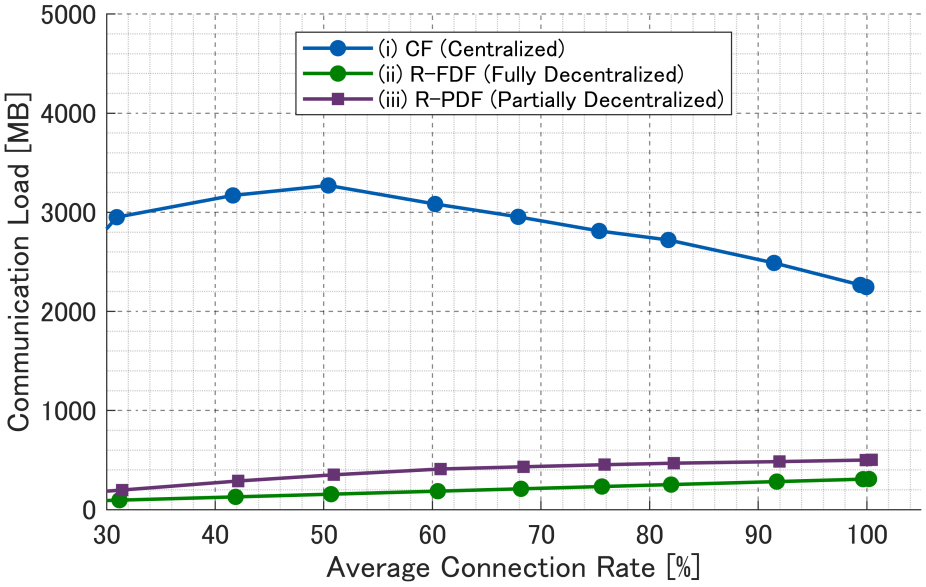


Figure 5.16: Communication load (the performance comparison of the existing algorithms).

Finally, the advantages and the disadvantages of the three estimation filters are summarized in the following tables with respect to the accuracy, stability, and communication load. Table 5.2 shows the features of estimation performances when the communication and sensing networks are dense, that is $\bar{\rho} \geq 70\%$. When the spacecraft in the swarm are strongly connected by the networks, high stability is achieved regardless of the decentralization levels. However, the stability of the centralized and the partially decentralized filters are notably degraded when the networks are sparse, that is $\bar{\rho} < 70\%$, as shown in Figure 5.12 and Table 5.3. If the fully decentralized filter is applied for the swarm navigation, high stability is achieved even with the sparse networks. However, the estimation accuracy is much worse than the centralized and the partially decentralized filters, as shown in Figures 5.14 and 5.15. In summary, it is demanded to develop a new algorithm to realize the

feature listed in the last row of Table 5.3.

Table 5.2: Characteristics of the three estimation algorithms for dense networks

Decentralization level	Accuracy	Stability	Communication load
Centralized	high	high	high
Fully decentralized	low	high	low
Partially decentralized	high ∼ medium	high	low ∼ medium

Color meanings: advantage ← medium → disadvantage

Table 5.3: Characteristics of the three estimation algorithms for sparse networks

Decentralization level	Accuracy	Stability	Communication load
Centralized	high	medium ∼ low	high
Fully decentralized	low	high	low
Partially decentralized	high ∼ medium	medium ∼ low	low
Desired algorithm	high ∼ medium	high	low

Color meanings: advantage ← medium → disadvantage

Chapter 6

Adaptive Decentralized Filter

6.1 Overview of Adaptive Decentralized Filter

As we analyze in Section 5.4, the fully decentralized filters have the advantage of high estimation stability; however, the estimation accuracy is relatively low. On the other hand, the centralized and partially decentralized filters have a challenge in estimation stability while achieving high estimation accuracy. Although both accuracy and stability are demanded in practical swarm applications, the existing algorithms do not satisfy both requirements with sufficient levels. To solve this issue, we propose a new estimation algorithm, the adaptive decentralized filter (ADF), that combines the advantages of the centralized and the decentralized estimation filters and realizes high accuracy and high stability at the same time.

The basic algorithmic flow of the ADF is illustrated in Figure 6.1. The ADF independently executes two filtering algorithms, the internal fully decentralized filter (iFDF) and the internal partially decentralized filter (iPDF). The ADF properly merges two estimates from the iFDF and the iPDF to make use of the advantages of two filters, the high stability of the iFDF and the high accuracy of the iPDF.

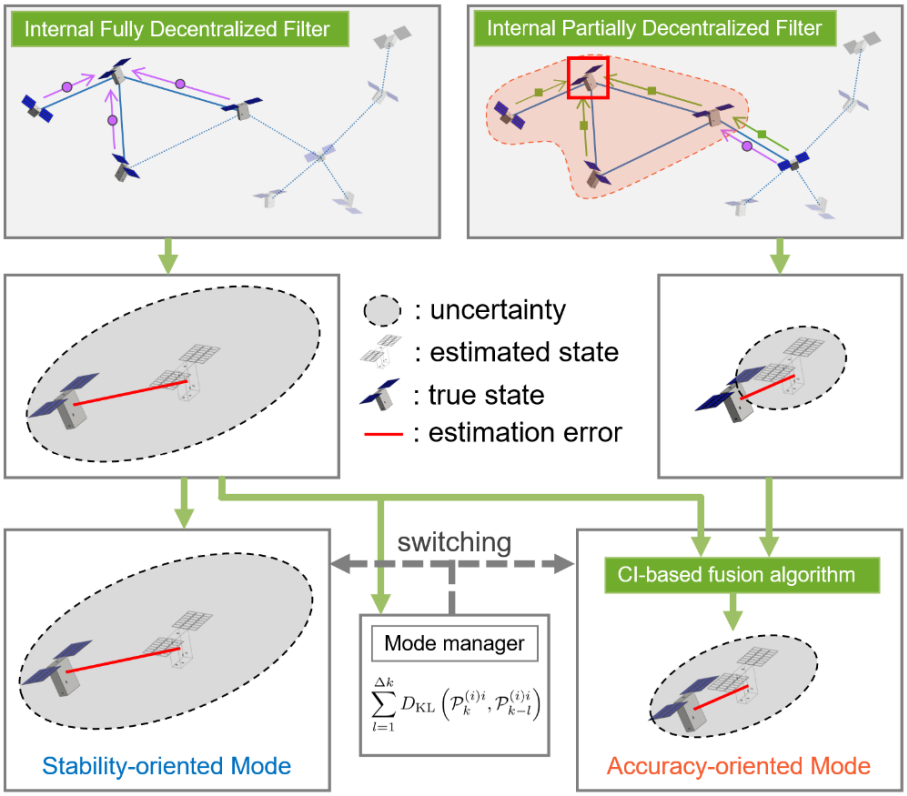


Figure 6.1: Algorithmic flow of the adaptive decentralized filter.

6.2 Algorithmic Details of Adaptive Decentralized Filter

This section summarizes the algorithmic detail of the ADF. The algorithm of the ADF can be divided into two phases, the pre-filtering phase and the fusing phase.

6.2.1 Pre-filtering Phase

Prediction

Before staring the measurement update, the prediction step is executed as a common process for both the iFDF and the iPDF. On the i-th spacecraft, its ADF only takes care of its own state in the prediction step, that is formulated as:

$$\boldsymbol{Z}_{k|k-1}^{(i)i} = \left(\boldsymbol{F}^{(i)}\left(\boldsymbol{Z}_{k-1}^{(i)i}\right)^{-1}\left(\boldsymbol{F}^{(i)}\right)^{T} + \boldsymbol{Q}^{(i)}\right)^{-1}, \tag{6.1}$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} = \boldsymbol{Z}_{k|k-1}^{(i)i}\left(\boldsymbol{F}^{(i)}\left(\boldsymbol{Z}_{k-1}^{(i)i}\right)^{-1}\hat{\boldsymbol{z}}_{k-1}^{(i)i} + \boldsymbol{G}^{(i)}\boldsymbol{u}_k^{(i)}\right). \tag{6.2}$$

Internal Fully Decentralized Filtering (iFDF)

In the iFDF, the prior estimates of Eqs. (6.1) and (6.2) are updated with the measurements obtained on the local sensing network. Figure 6.2 shows the local sensing networks that are visible from the iFDF on the i-th spacecraft. The i-th spacecraft communicates with the set of adjacent spacecraft in $\mathcal{V}_{i,k}^{rb} = \left\{ j \in \mathcal{V} \mid (i,j) \in \mathcal{E}_{i,k}^{r} \cap \mathcal{E}_{i,k}^{b} \right\}$. The measurement update step of the iFDF is formulated as:

$$
\begin{aligned}
{}^{f}\boldsymbol{Z}_{k|k}^{(i)i} &= \omega_i \boldsymbol{Z}_{k|k-1}^{(i)i} + \sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \boldsymbol{Z}_{k|k-1}^{(i)j}, && (6.3)\\
{}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i} &= \omega_i \hat{\boldsymbol{z}}_{k|k-1}^{(i)i} + \sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j \hat{\boldsymbol{z}}_{k|k-1}^{(i)j}, && (6.4)\\
\boldsymbol{\omega} &= \begin{bmatrix} \omega_i & \omega_{j_1} & \cdots & \omega_{j_m} \end{bmatrix} = \arg\min_{\boldsymbol{\omega}} \left[\mathrm{trace} \left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i} \right)^{-1} \right], \\
& \quad \{j_1, \cdots, j_m\} \in \mathcal{V}_{i,k}^{rb}. && (6.5)
\end{aligned}
$$

where ω_i, ω_j are the CI parameters satisfying the condition: $\omega_i + \sum_{j \in \mathcal{V}_{i,k}^{rb}} \omega_j = 1$. The left-superscript f represents the value computed by the iFDF. The likelihood function $\mathcal{N}\left((\boldsymbol{Z}_{k|k-1}^{(i)j})^{-1} \hat{\boldsymbol{z}}_{k|k-1}^{(i)j}, (\boldsymbol{Z}_{k|k-1}^{(i)j})^{-1} \right)$ is provided from the adjacent spacecraft in $\mathcal{V}_{i,k}^{rb}$ by the inter-spacecraft communication.

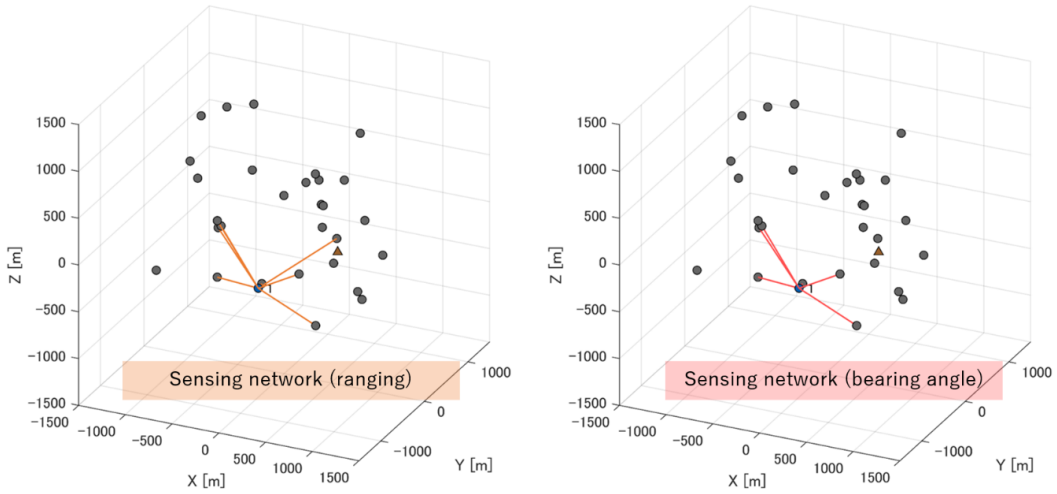


Figure 6.2: Sensing networks for the iFDF on the i-th spacecraft.

Internal Partially Decentralized Filtering (iPDF)

In the iPDF, to increase the information gain, the probability of the aggregated multi-spacecraft state is considered. The aggregated state vector managed by the i-th spacecraft

is defined as:

$$\boldsymbol{x}^{(\tilde{\mathcal{V}}^c_{i,k})} = \left[\left(\boldsymbol{x}_k^{(i)}\right)^T \quad \cdots \quad \left(\boldsymbol{x}_k^{(j)}\right)^T \quad \cdots \right]^T, \quad j \in \mathcal{V}^c_{i,k} \tag{6.6}$$

hence the corresponding estimated probability of the iPDF on the i-th spacecraft is represented with the canonical parameters as:

$$p\left(\boldsymbol{x}^{(\tilde{\mathcal{V}}^c_{i,k})}\right) = \mathcal{N}\left(\left(\boldsymbol{Z}^{(\tilde{\mathcal{V}}^c_{i,k})}\right)^{-1} \hat{\boldsymbol{z}}^{(\tilde{\mathcal{V}}^c_{i,k})},\ \left(\boldsymbol{Z}^{(\tilde{\mathcal{V}}^c_{i,k})}\right)^{-1}\right). \tag{6.7}$$

The prior probability of the i-th spacecraft state is computed on the i-th spacecraft itself; therefore, no communication is required. On the other hand, the prior probability of the adjacent spacecraft in $\mathcal{V}^c_{i,k}$ is transferred to the i-th spacecraft through the inter-spacecraft communication links.

The posterior probability expressed in the canonical parameterization is computed as:

$${}^p\boldsymbol{Z}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k} = \omega_i \left(\boldsymbol{Z}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k-1} + \boldsymbol{U}^{(\tilde{\mathcal{V}}^c_{i,k})}_k\right) + \sum_{j \in \tilde{\mathcal{M}}^c_{i,k}} \omega_j \boldsymbol{Z}^{(j)i}_{k|k-1} \tag{6.8}$$

$${}^p\hat{\boldsymbol{z}}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k} = \omega_i \left(\hat{\boldsymbol{z}}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k-1} + \boldsymbol{u}^{(\tilde{\mathcal{V}}^c_{i,k})}_k\right) + \sum_{j \in \tilde{\mathcal{M}}^c_{i,k}} \omega_j \hat{\boldsymbol{z}}^{(j)i}_{k|k-1} \tag{6.9}$$

where the left-superscript p represents the value of the iPDF and the information contribution is expressed as:

$$\boldsymbol{U}^{(\tilde{\mathcal{V}}^c_{i,k})i}_k = \sum_{s \in \{r,b\}} \left(\boldsymbol{H}^{(c)}_{s,k}\right)^T \left(\tilde{\boldsymbol{R}}^{(\tilde{\mathcal{V}}^c_{i,k})}_{s,k}\right)^{-1} \boldsymbol{H}^{(c)}_{s,k} \tag{6.10}$$

$$\boldsymbol{u}^{(\tilde{\mathcal{V}}^c_{i,k})i}_k = \sum_{s \in \{r,b\}} \left(\boldsymbol{H}^{(c)}_{s,k}\right)^T \left(\tilde{\boldsymbol{R}}^{(\tilde{\mathcal{V}}^c_{i,k})}_{s,k}\right)^{-1} \left(\boldsymbol{y}^{(c)}_{s,k} - \hat{\boldsymbol{y}}^{(c)}_{s,k|k-1} + \boldsymbol{H}^{(c)}_{s,k} \hat{\boldsymbol{x}}^{(c)c}_{k|k-1}\right). \tag{6.11}$$

The final output of the iPDF is the posterior probability of the i-spacecraft state. This is obtained by extracting the corresponding part of the i-th spacecraft, that is expressed as:

$${}^p\boldsymbol{Z}^{(i)i}_{k|k} = \Upsilon^{(i)}\left[\boldsymbol{Z}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k}\right] \tag{6.12}$$

$${}^p\hat{\boldsymbol{z}}^{(i)i}_{k|k} = \Upsilon^{(i)}\left[\hat{\boldsymbol{z}}^{(\tilde{\mathcal{V}}^c_{i,k})i}_{k|k}\right] \tag{6.13}$$

where $\Upsilon^{(i)}$ is an operator to extract the i-th spacecraft related part from an argument vector or matrix.

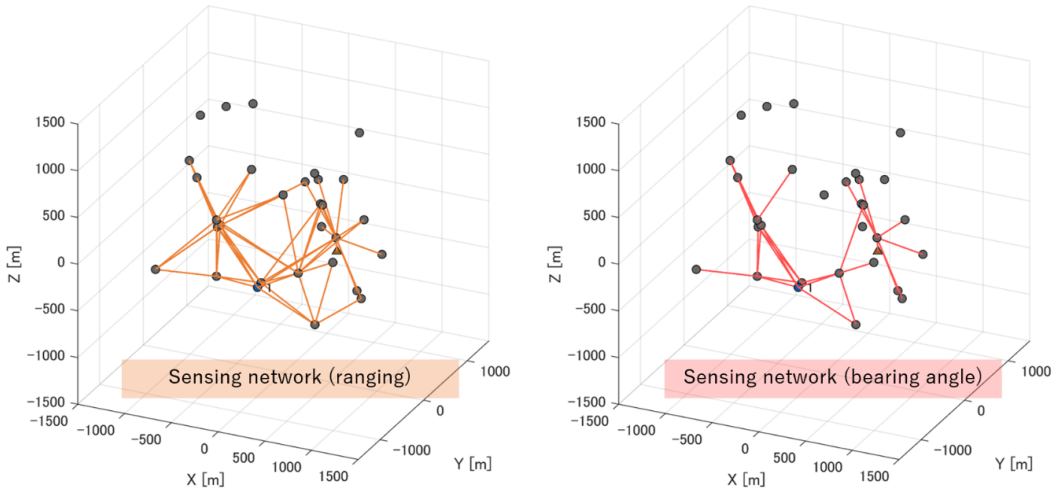


Figure 6.3: Sensing networks for the iPDF on the i-th spacecraft.

6.2.2 Fusing Phase

In the fusing phase succeeding the pre-filtering phase, the two probabilities from the iFDF and the iPDF are appropriately merged, and the posterior probability is generated as the final output of the ADF. The ADF has two fusion modes to balance the estimation accuracy and stability, the stability-oriented mode and the accuracy-oriented mode. In the early phase of the estimation process, the stability-oriented mode is activated. In this mode, the estimator only relies on the result of the iFDF to ensure that the estimation is undoubtedly converged.

$$
\begin{aligned}
{}^{a}\boldsymbol{Z}_{k|k}^{(i)i} &= {}^{f}\boldsymbol{Z}_{k|k}^{(i)i} && (6.14)\\
{}^{a}\hat{\boldsymbol{z}}_{k|k}^{(i)i} &= {}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i} && (6.15)
\end{aligned}
$$

Once satisfying the mode switching criteria, the stability-oriented mode is switched to the accuracy-oriented mode. As the switching criteria, we focus on the time variation of the posterior probability in the iFDF. If the following relationship is satisfied, the accuracy-oriented mode is activated.

$$
\sum_{l=1}^{\Delta k} D_{\mathrm{KL}}\left({}^{f}\mathcal{P}_{k}^{(i)i} \middle\| {}^{f}\mathcal{P}_{k-l}^{(i)i} \right) \leq \Delta P \tag{6.16}
$$

$$
{}^{f}\mathcal{P}_{k}^{(i)i} = \mathcal{N}\left({}^{f}\hat{\boldsymbol{x}}_{k}^{(i)i}, {}^{f}\boldsymbol{P}_{k}^{(i)i} \right) \tag{6.17}
$$

where the operator $\mathrm{D_{KL}}(\cdot)$ denotes the Kullback-Leibler divergence, which measures the difference between two probability distributions. The definition of the KL-divergence is

summarized in Appendix A.5. This assessment process of error covariances is visualized in Figure 6.4.

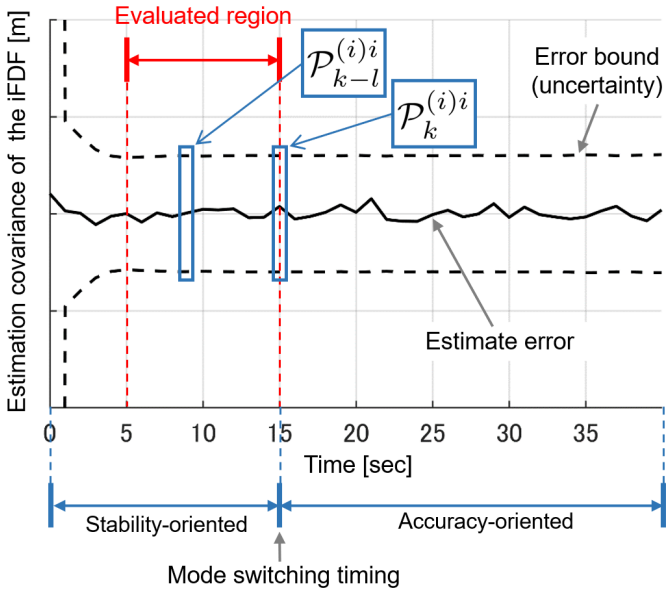


Figure 6.4: Error covariance assessment for the mode switching.

After the accuracy-oriented mode is activated, the fusion process of the iFDF and the iPDF starts. To fuse the two filtering results, we adopt the Covariance Intersection (CI) technique for the consistent merging of two estimates. [51] The CI-based fusion process of two estimation results is formulated as

$$
\begin{aligned}
{}^{a}\boldsymbol{Z}_{k}^{(i)i} &= \omega\left({}^{f}\boldsymbol{Z}_{k}^{(i)i}\right) + (1-\omega)\left({}^{p}\boldsymbol{Z}_{k}^{(i)i}\right), && (6.18)\\
{}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i} &= \omega\left({}^{f}\hat{\boldsymbol{z}}_{k}^{(i)i}\right) + (1-\omega)\left({}^{p}\hat{\boldsymbol{z}}_{k}^{(i)i}\right), && (6.19)\\
\omega &= \arg\min_{\omega}\left[\mathrm{trace}\left({}^{a}\boldsymbol{Z}_{k|k}^{(i)i}\right)^{-1}\right]. && (6.20)
\end{aligned}
$$

6.2.3 Summary

Algorithm 9 summarizes the computational flow of the ADF.

Algorithm 9: ADF: Adaptive Decentralized Filter

At the k-th time step on the i-th spacecraft;

(The following process runs on all the spacecraft.);

input : $\hat{\boldsymbol{z}}_{k-1}^{(i)i}, \boldsymbol{Z}_{k-1}^{(i)i}$

output: ${}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i}, {}^{a}\boldsymbol{Z}_{k}^{(i)i}$

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(i)i} = \left(\boldsymbol{F}^{(i)}\left(\boldsymbol{Z}_{k-1}^{(i)i}\right)^{-1}\left(\boldsymbol{F}^{(i)}\right)^{T} + \boldsymbol{Q}^{(i)}\right)^{-1};$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} = \boldsymbol{Z}_{k|k-1}^{(i)i}\left(\boldsymbol{F}^{(i)}\left(\boldsymbol{Z}_{k-1}^{(i)i}\right)^{-1}\hat{\boldsymbol{z}}_{k-1}^{(i)} + \boldsymbol{G}^{(i)}\boldsymbol{u}_{k}^{(i)}\right);$$

Inter-spacecraft communication step;

Collect the estimates and the related measurement data from the adjacent spacecraft.;

Internal Fully Decentralized Filtering (iFDF);

Compute the estimated positions by ;

Convert them to the information forms by ;

$${}^{f}\boldsymbol{Z}_{k|k}^{(i)i} = \omega^{(i)}\boldsymbol{Z}_{k|k-1}^{(i)i} + \textstyle\sum_{j\in\mathcal{V}_{k}^{(i)}}\omega^{(j)}\boldsymbol{Z}_{k|k-1}^{(i)j};$$

$${}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \omega^{(i)}\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} + \textstyle\sum_{j\in\mathcal{V}_{k}^{(i)}}\omega^{(j)}\hat{\boldsymbol{z}}_{k|k-1}^{(i)j};$$

where $\omega^{(i)} + \sum_{j\in\mathcal{V}_{k}^{(i)}}\omega^{(j)} = 1$ is satisfied.;

$$\mathcal{P}_{k}^{(i)i} = \mathcal{N}\left(\left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i}\right)^{-1}{}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i},\ \left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i}\right)^{-1}\right);$$

(... continues to the next page...)

Internal Partially Decentralized Filtering (iPDF);

Compute the adjusted error covariance based on the local sensing networks ;

$$\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} = \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^{s}\right), \quad s \in \{r, b\} \ ;$$

Compute the information contribution in the estimated subset ;

$$\boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \boldsymbol{H}_{s,k}^{(c)} \ ;$$

$$\boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c}\right) \ ;$$

Compute the information contribution of the support spacecraft $(l \in \tilde{\mathcal{M}}_{i,k}^{c}, j \in \tilde{\mathcal{E}}_{k}^{c})$;

$$\hat{\boldsymbol{p}}_{k|k-1}^{(j)l} = \boldsymbol{g}\left(\boldsymbol{p}_{k|k-1}^{(l)l}, \boldsymbol{y}_{k}^{(jl)l}\right) \ ;$$

$$\tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l} = \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_{k}^{(l)}}\right) \boldsymbol{P}_{p,k|k-1}^{(l)l} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_{k}^{(l)}}\right)^{T} + \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_{k}^{(jl)}}\right) \boldsymbol{R}_{k}^{(jl)} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_{k}^{(jl)}}\right)^{T} \ ;$$

Convert $\hat{\boldsymbol{p}}_{k|k-1}^{(j)l}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l}$ to $\hat{\boldsymbol{u}}_{k|k-1}^{(j)l}, \boldsymbol{U}_{k|k-1}^{(j)l}$;

Execute the covariance intersection ;

$$\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_i \left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^{c}} \omega_p \boldsymbol{U}_{k|k-1}^{(j)l} \ ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_i \left(\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^{c}} \omega_p \boldsymbol{u}_{k|k-1}^{(j)l} \ ;$$

Extract the i-th spacecraft related part;

$${}^{p}\boldsymbol{Z}_{k|k}^{(i)i} = \Upsilon^{(i)}\left[\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}\right], \qquad {}^{p}\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \Upsilon^{(i)}\left[\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}\right];$$

Fusing Phase;

if $\sum_{l=1}^{\Delta k} \mathrm{KL}\left({}^{f}\mathcal{P}_{k}^{(i)i}, {}^{f}\mathcal{P}_{k-l}^{(i)i}\right) \leq \Delta P$ **then**

$${}^{a}\boldsymbol{Z}_{k}^{(i)i} = \omega\left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i}\right) + (1-\omega)\left({}^{p}\boldsymbol{Z}_{k|k}^{(i)i}\right);$$

$${}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i} = \omega\left({}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i}\right) + (1-\omega)\left({}^{p}\hat{\boldsymbol{z}}_{k|k}^{(i)i}\right);$$

$$\omega = \arg\min_{\omega}\left[\mathrm{trace}\left({}^{a}\boldsymbol{Z}_{k|k}^{(i)i}\right)^{-1}\right];$$

else

$${}^{a}\boldsymbol{Z}_{k}^{(i)i} = {}^{f}\boldsymbol{Z}_{k|k}^{(i)i}, \quad {}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i} = {}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i};$$

end

Chapter 7

Observability-driven Decentralization

7.1 Overview of Observability-driven Decentralization

The adaptive decentralization introduced in Chapter 6 improves the estimation accuracy and stability. However, in the ADF, it is not guaranteed that the iPDF can execute measurement updates without divergence, even if the estimation results of the iFDF get stable. If the extended sensing networks considered in the iPDF do not have enough measurement information gain to estimate the probability of the aggregated spacecraft state, the measurement update, including the estimate of the partially decentralized filter, results in convergence to wrong estimates or divergence. Therefore, to improve the stability of the ADF, additional care for the estimate convergence would be necessary, especially when switching from the stability-oriented mode to the accuracy-oriented mode, as illustrated in Figure 7.1.

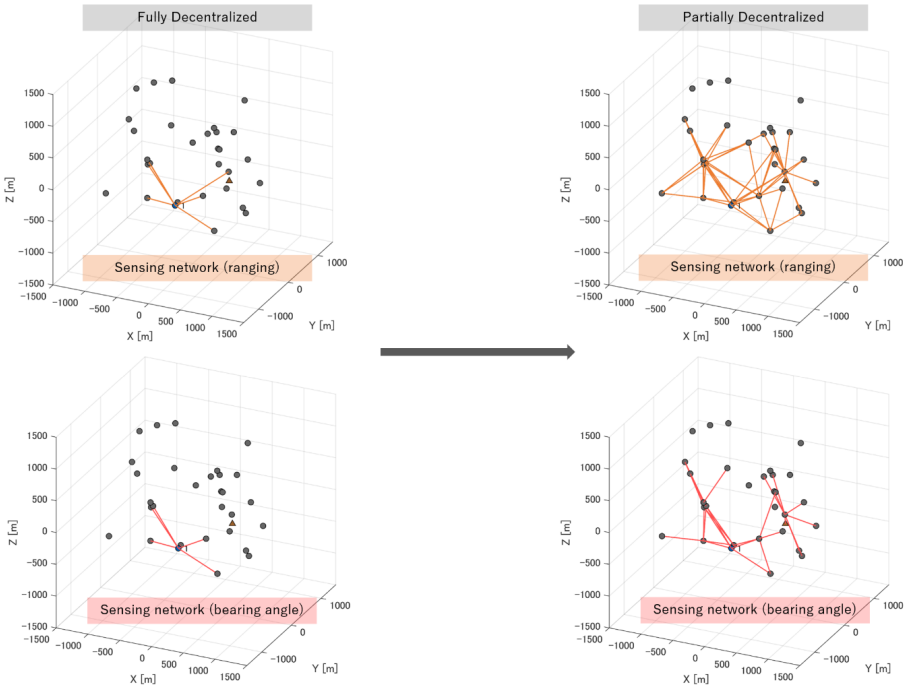


Figure 7.1: Observability change associated with switching from the iFDF to the iPDF.

As a quantitative parameter to evaluate the information sufficiency, we propose the evaluation of the observability of the sensing networks in the iPDF. For example, if the spacecraft is located in a low-observability configuration, the ADF should only rely on the estimate of the iFDF. Alternately, if the swarm is configured well in terms of observability, the ADF can safely start to fuse the estimate of the iPDF.

Figure 7.2 illustrates the concept of the observability-driven decentralization for the ADF. As a switching criterion between the stability-oriented mode and the accuracy-oriented mode, the observability assessment of the sensing network for the iPDF is introduced. The following items should be discussed and solved to implement this new functionality in the ADF.

- How to define the observability of the sensing networks? (discussed in Section 7.2)
- How to evaluate the observability with the onboard resources? (discussed in Section 7.3)
- How to decide the timing of switching decentralization based on the observability evaluation? (discussed in Section 7.4)

We discuss the above items in the following sections and propose the observability-driven adaptive decentralized filter (OD-ADF).

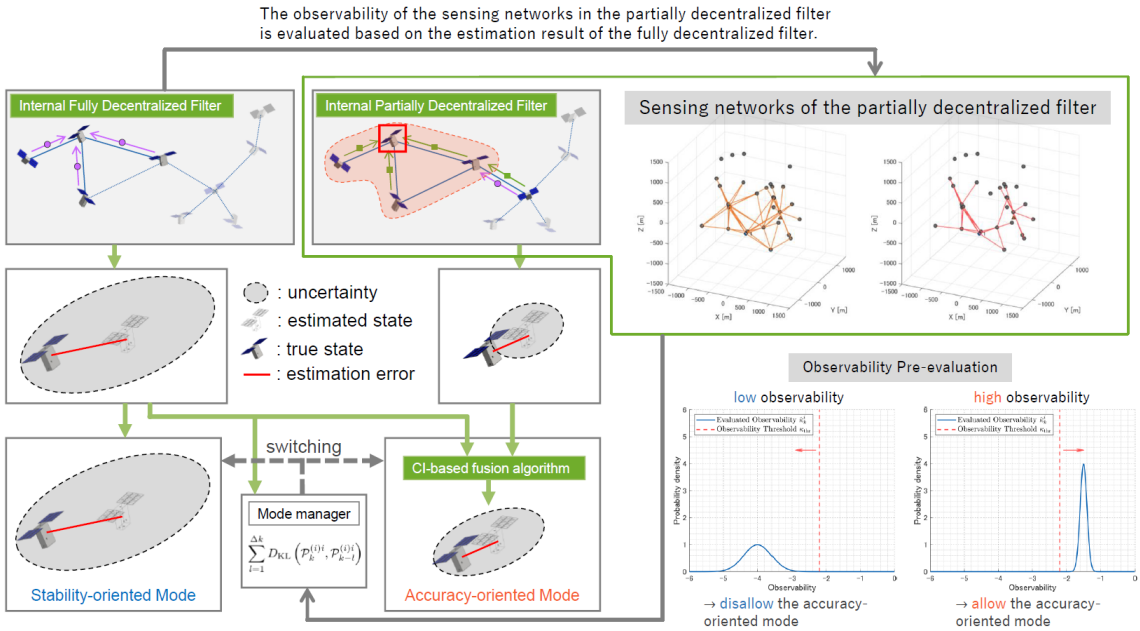


Figure 7.2: Observability-driven decentralization for the ADF.

7.2 Definition of Observability

State observability is defined as whether the system state can be properly estimated given a specific output. [52] In the context of spacecraft swarm navigation, the observability represents whether the spacecraft states in the swarm can be estimated with the measurement information of the onboard sensors. Although there are various observability metrics, we focus on the condition number of the observability Gramian.

Here we look into the meanings of the observability based on the condition number through the linear perturbed theory. We consider a simple linear system equation:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{7.1}$$

where $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is a square nonsingular matrix. If the system of Eq. (7.1) suffers from a certain perturbation, the equation is formulated as:

$$(\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b} + \Delta\boldsymbol{b}. \tag{7.2}$$

As a state estimation problem, $\boldsymbol{x}$ corresponds to the states to be estimated, $\boldsymbol{b}$ corresponds to measurement-related values, and $\boldsymbol{A}$ corresponds to the matrix representing the relationship between the states and the measurements. Here we are interested in how the perturbations of $\boldsymbol{A}$ and $\boldsymbol{b}$ affects the estimates of $\boldsymbol{x}$. Solving Eq. (7.2) for $\Delta\boldsymbol{x}$, one obtains:

$$\Delta\boldsymbol{x} = \left(\boldsymbol{I} + \boldsymbol{A}^{-1}\Delta\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{-1}\left(\Delta\boldsymbol{b} - \Delta\boldsymbol{A}\boldsymbol{x}\right) \tag{7.3}$$

where the relationships $\Delta \boldsymbol{A} \Delta \boldsymbol{x} \simeq \boldsymbol{0}$ and $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ are used. Considering the norms of the elements in Eq. (7.3), the following inequality equation is satisfied:

$$\|\Delta \boldsymbol{x}\| \leq \left\| \left(\boldsymbol{I} + \boldsymbol{A}^{-1} \Delta \boldsymbol{A} \right)^{-1} \right\| \cdot \left\| \boldsymbol{A}^{-1} \right\| \left(\|\Delta \boldsymbol{b}\| + \|\Delta \boldsymbol{A}\| \cdot \|\boldsymbol{x}\| \right). \tag{7.4}$$

When assuming $\boldsymbol{b} \neq \boldsymbol{0}$ and $\boldsymbol{x} \neq \boldsymbol{0}$, one obtains the inequality equation for the relative state error as:

$$\begin{aligned} \frac{\|\Delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} &\leq \left\| \left(\boldsymbol{I} + \boldsymbol{A}^{-1} \Delta \boldsymbol{A} \right)^{-1} \right\| \cdot \left\| \boldsymbol{A}^{-1} \right\| \left(\frac{\|\Delta \boldsymbol{b}\|}{\|\boldsymbol{x}\|} + \|\Delta \boldsymbol{A}\| \right) && (7.5) \\ &\leq \frac{\left\| \boldsymbol{A}^{-1} \right\| \cdot \|\boldsymbol{A}\|}{1 - \left\| \boldsymbol{A}^{-1} \right\| \cdot \|\boldsymbol{A}\| \frac{\|\Delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}} \left(\frac{\|\Delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|} + \frac{\|\Delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} \right). && (7.6) \end{aligned}$$

Therefore, the system and the measurement errors are amplified by $\|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{A}\|$ for small perturbation cases, in which the denominator is not further from 1. This amplification factor $\|\boldsymbol{A}^{-1}\| \cdot \|\boldsymbol{A}\|$ is defined as the condition number of the matrix $\boldsymbol{A}$. When considering the L^2 norm, the condition number is defined as:

$$\|\boldsymbol{A}^{-1}\|_2 \cdot \|\boldsymbol{A}\|_2 = \max \left[\sigma \left(\boldsymbol{A}^{-1} \right) \right] \cdot \max \left[\sigma \left(\boldsymbol{A} \right) \right] = \frac{\max \left[\sigma \left(\boldsymbol{A} \right) \right]}{\min \left[\sigma \left(\boldsymbol{A} \right) \right]} \tag{7.7}$$

where the operator $\sigma(\cdot)$ returns the singular value of a matrix, and the L^2 norm of the matrix is identical to the maximum singular value of the matrix.

7.3 Probabilistic Evaluation of Network Observability

To assess the observability of the sensing network, information on spacecraft location is necessary. However, considering an actual swarm operation, the spacecraft position is obtainable only by the onboard estimation. Therefore, in a practical operation, the observability needs to be evaluated with the estimated results of the spacecraft location. Figure 7.3 shows the observability assessed based on the true states and the estimated states. Though this is an exaggerated case, it is difficult to precisely evaluate the observability based on the estimated mean value of the spacecraft position, especially when the estimation uncertainty is relatively large. Therefore, it is demanded to assess the observability in a probabilistic manner considering the uncertainty of the estimated positions.

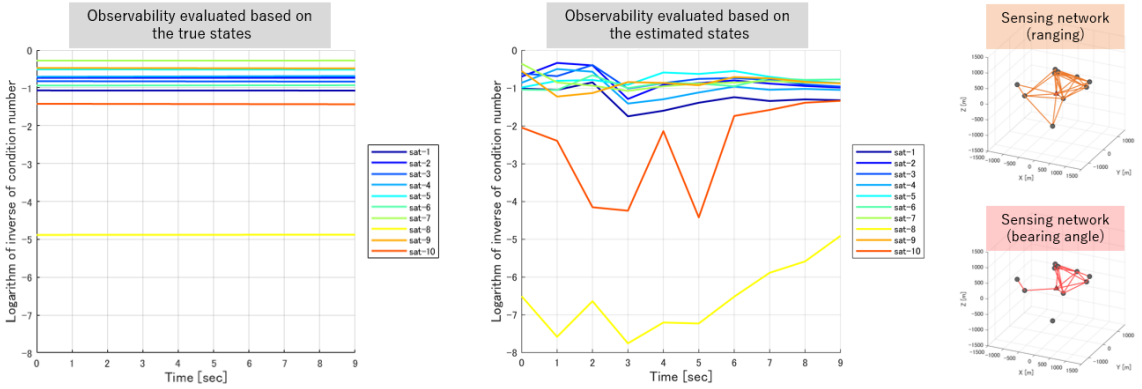


Figure 7.3: Observability assessment based on the true states or the estimated states.

To examine the observability of the sensing network for the iPDF, we consider the observability Gramian W defined as:

$$W\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right) = \sum_{s\in\{r,b\}} \left(\frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{s[c]}}}{\partial \boldsymbol{x}_k^{(c)}}\right)^T \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^s\right) \cdot \frac{\partial \boldsymbol{h}_{\mathcal{E}_k^{s[c]}}}{\partial \boldsymbol{x}_k^{(c)}}. \tag{7.8}$$

With the observability Gramian of Eq. (7.8), the observability metric of the iPDF on the i-th spacecraft κ_k^i is formulated as:

$$\kappa_k^i = f_{obs}\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right) = -\log_{10}\left[\frac{\max\left[W\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right)\right]}{\min\left[W\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right)\right]}\right]. \tag{7.9}$$

where we adopt the logarithmic value for the convenience of numerical processing. As mentioned in the beginning of this chapter, the spacecraft needs to compute Eq. (7.9) with the estimated value of $\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}$. To take the estimation uncertainty in consideration, we probabilistically evaluate Eq. (7.9). Let the estimated probability of the aggregated state as $p(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)})$, then the observability is evaluated as:

$$\kappa_k^i = \int f_{obs}\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right) p\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right) d\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}. \tag{7.10}$$

Figure 7.4 shows the result comparison of the following two cases. The left figure shows the result of evaluating Eq.(7.9) with the mean value of spacecraft states, and the right figure shows the probabilistically evaluated value by Eq.(7.10). When the Gaussian filter is applied as the navigation algorithm, the probability is approximated with the Gaussian distribution, hence:

$$\kappa_k^i \simeq \int f_{obs}\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}\right) \mathcal{N}\left(\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)} \middle| \hat{\boldsymbol{x}}_{k|k-1}^{(\mathcal{V}_{i,k}^c)i}, \boldsymbol{P}_{k|k-1}^{(\mathcal{V}_{i,k}^c)i}\right) d\boldsymbol{x}_k^{(\mathcal{V}_{i,k}^c)}. \tag{7.11}$$

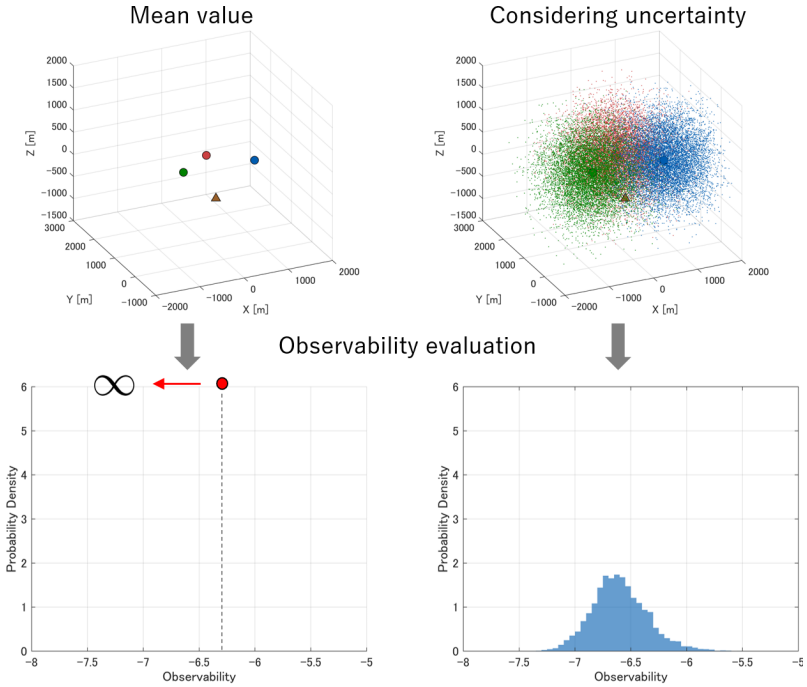


Figure 7.4: Observability computed only with the mean values and one computed considering the estimation uncertainty.

Additionally, we adopt the sigma point sampling to compute Eq. (7.11) as:

$$\hat{\kappa}_k^i \simeq \sum_s w_s f_{obs}\left({}^s\hat{\boldsymbol{x}}_{k|k-1}^{(\mathcal{V}_{i,k}^c)i}\right) \tag{7.12}$$

where we utilize the unscented transform introduced in Appendix A.3 for practical implementation. Figure 7.5 shows the result of probabilistic observability evaluation based on the sigma point sampling.

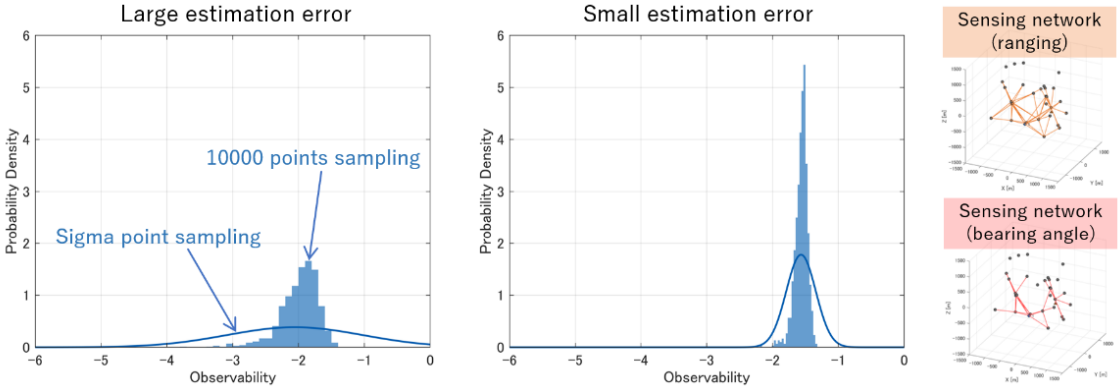


Figure 7.5: Observability evaluated by the sigma point sampling.

7.4 Threshold for Observability-Driven Decentralization

Each adaptive decentralized filter obtains the observability of the sensing network in the iPDF by computing Eq. (7.12). For the next step, we need to consider assessing the obtained observability. One solution is to define a threshold value, which is used for the decision to select the stability-oriented mode or the accuracy-oriented mode, and evaluate whether the computed observability exceeds the threshold or not.

The values of the observability defined by Eq. (7.9) highly depend on the system of an evaluated target, including the dimensions of an estimated variable, the types of sensor measurements, and so on. Therefore, it would be impractical to find some reasonable threshold values theoretically. Consequently, we consider deciding the threshold value by a heuristic approach.

To begin with, we evaluate the observability values of swarm configurations listed in Figure 7.6, which has 50 configurations for each number of spacecraft ranging from 2 to 9. The number of spacecraft whose states are evaluated in the iPDF can range from 2 to N; hence, we need to consider the cases with the different number of spacecraft. The observability values are computed based on the following conditions.

- Eq. (7.9) is used to calculate the observability. (This is the off-line evaluation; hence there is no need to assess the value in a probabilistic manner.)
- The sensing network for each swarm configuration is a complete graph.
- The inter-spacecraft range and bearing angle are considered as the sensor measurements.

The solid blue line in Figure 7.7 shows the evaluated values of the observability with respect to the swarm configurations in Figure 7.6. The circle represents the average value, and the edges of error bars mean the maximum and the minimum values. We can see that the observability values range from -4 to -8. Based on this result, we investigate the performances of the OD-ADF with different thresholds. With the spacecraft swarm simulator introduced in Chapter 4, we compare the performances of three estimation algorithms: the iFDF-only ADF, the iPDF-only ADF, and the OD-ADF with a constant threshold. The OD-ADF is evaluated with the different thresholds, $\kappa_{\mathrm{thr}} = -4$, -5, -6, -7, and -8, expressed with the dashed pink lines in Figure 7.7. The observability threshold should be designed to take advantages of the iFDF and the iPDF effectively, as shown in Figure 7.8.

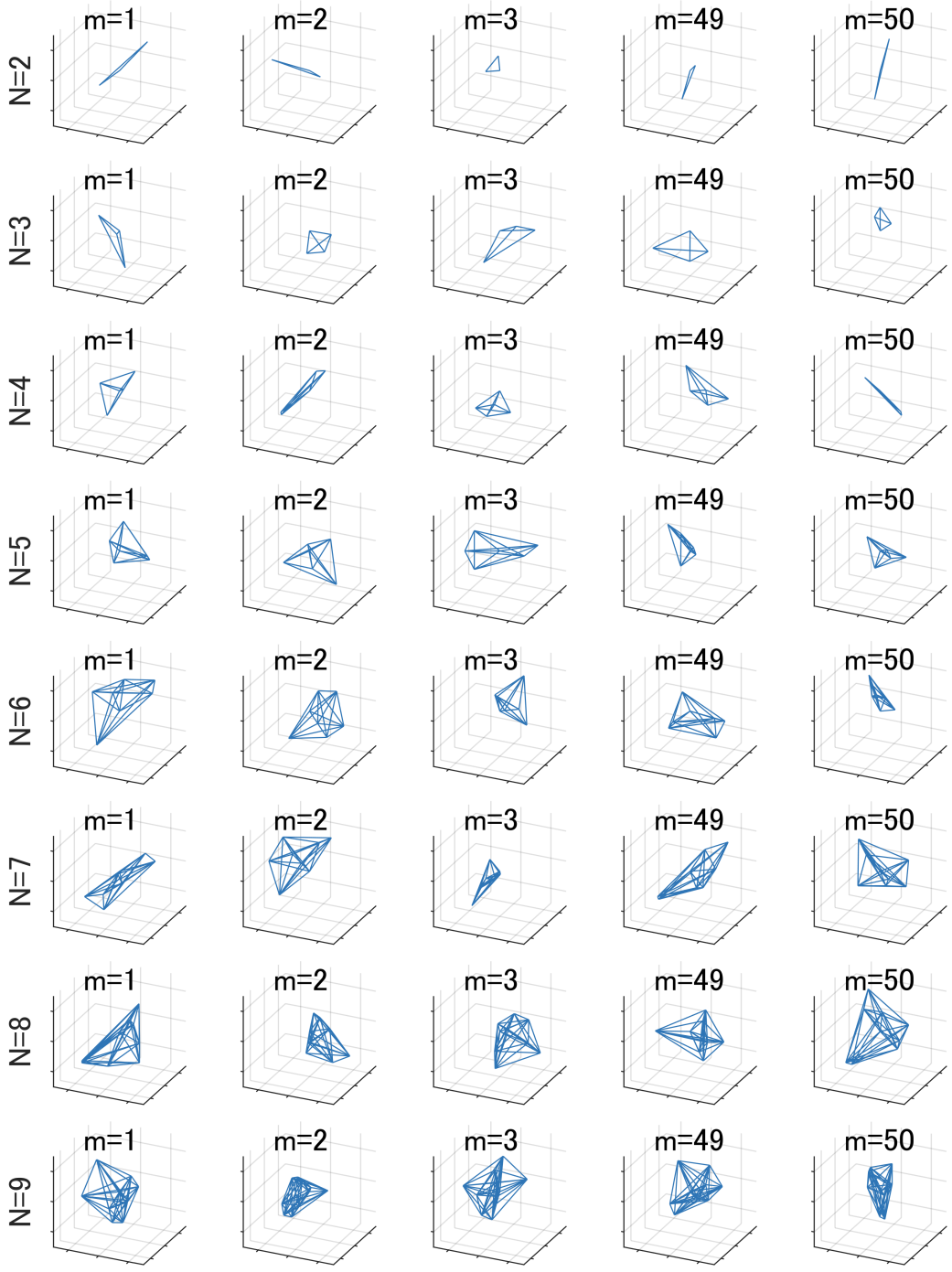


Figure 7.6: Swarm configurations and the networks to be evaluated for the observability threshold analysis.

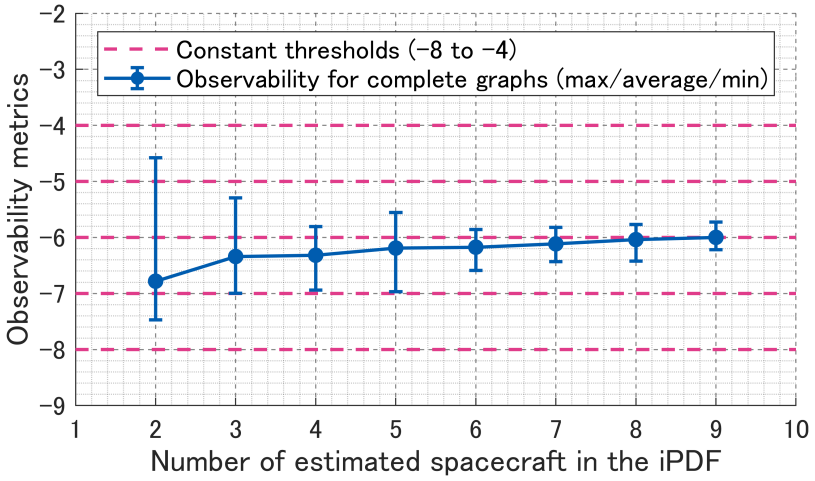


Figure 7.7: Observability pre-evaluation for the swarm configurations in Figure 7.6.

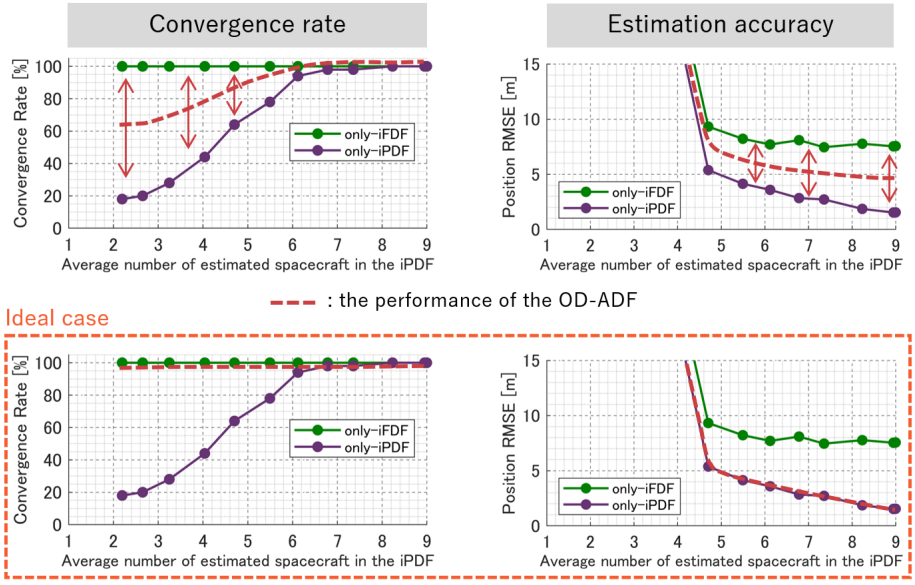
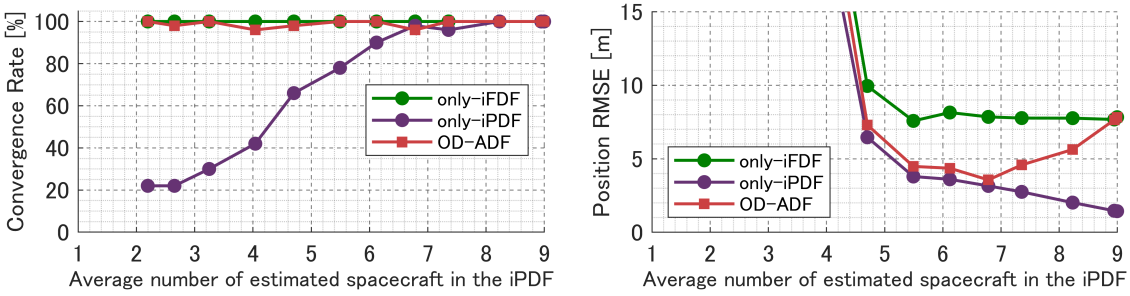


Figure 7.8: Requirements for the design of observability threshold.

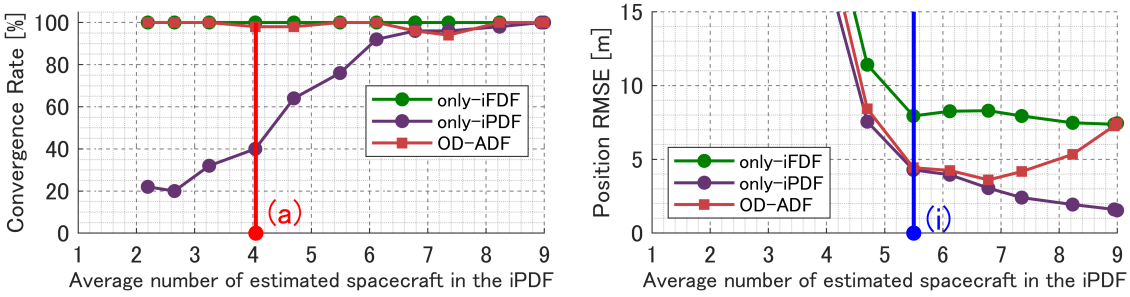
Figures 7.9 to 7.13 show the simulation results with the different observability thresholds for the OD-ADF. When the threshold is around -4 to -5, the OD-ADF achieves the 100% convergence rate. However, when the threshold is -6, the convergence rate of the OD-ADF is degraded where the number of estimated spacecraft in the iPDF $n(\tilde{\mathcal{V}}^c_{i,k})$ is around 2 to 4. This fact implies that the threshold should be higher than -5 when $n(\tilde{\mathcal{V}}^c_{i,k})$ is 4 to take advantage of the high stability of the iFDF. The solid red line in Figure 7.10 represents this requirement. When it comes to the position estimation error, we can find the upper bounds for the observability threshold. Firstly, when the threshold is -4, the estimation

accuracy of the OD-ADF is worse than that of the iPDF-only ADF regardless of $n(\tilde{\mathcal{V}}^c_{i,k})$. However, when the threshold is -5, the estimation error of the OD-ADF matches that of the iPDF-only ADF where $n(\tilde{\mathcal{V}}^c_{i,k})$ is around 5.5. This fact means that the threshold should be lower than -5 when $n(\tilde{\mathcal{V}}^c_{i,k})$ is 5.5 to take advantage of the high accuracy of the iPDF. The solid blue line in Figure 7.10 represents this requirement. We can find other requirements for convergence rate and estimation accuracy in Figures 7.11 and 7.12.



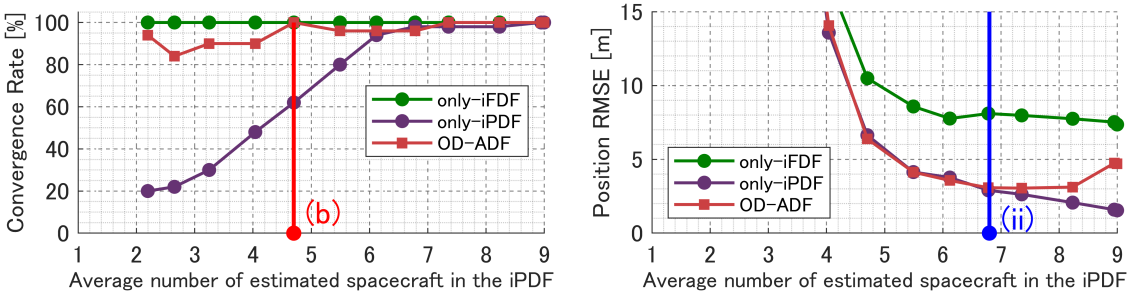
(a) Convergence rate **(b)** Position estimation error

Figure 7.9: Convergence rate and the position estimation error in the case of $\kappa_{\text{thr}} = -4$.



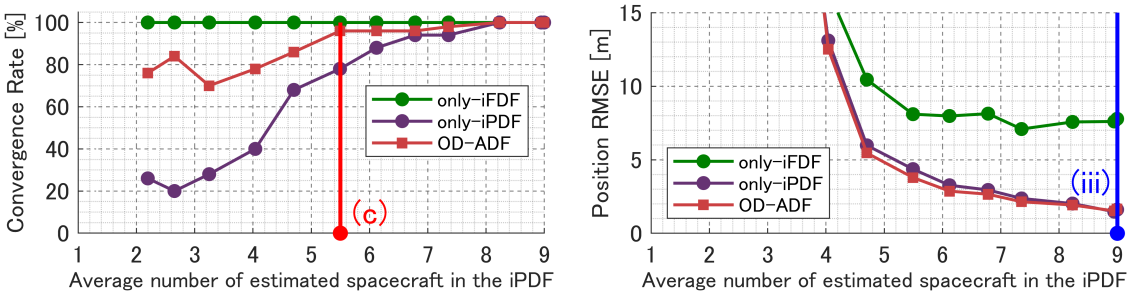
(a) Convergence rate **(b)** Position estimation error

Figure 7.10: Convergence rate and the position estimation error in the case of $\kappa_{\text{thr}} = -5$.



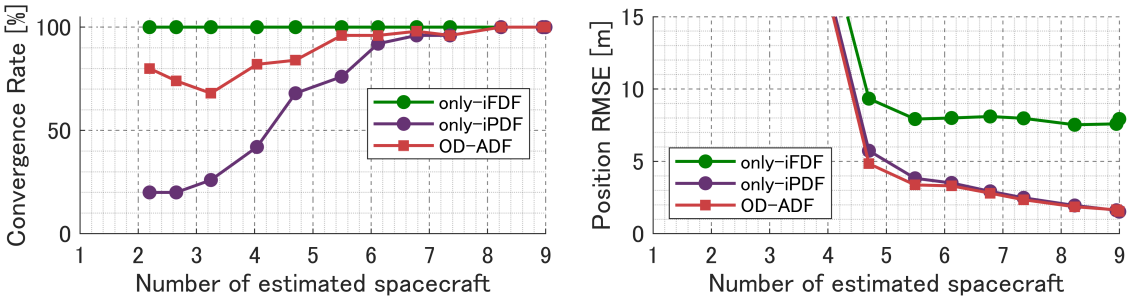
(a) Convergence rate **(b)** Position estimation error

Figure 7.11: Convergence rate and the position estimation error in the case of $\kappa_{\text{thr}} = -6$.



(a) Convergence rate **(b)** Position estimation error

Figure 7.12: Convergence rate and the position estimation error in the case of $\kappa_{\mathrm{thr}} = -7$.



(a) Convergence rate **(b)** Position estimation error

Figure 7.13: Convergence rate and the position estimation error in the case of $\kappa_{\mathrm{thr}} = -8$.

Based on the results of Figures 7.9 to 7.13 and the following discussions, we can draw the requirement for the observability threshold as Figure 7.14.

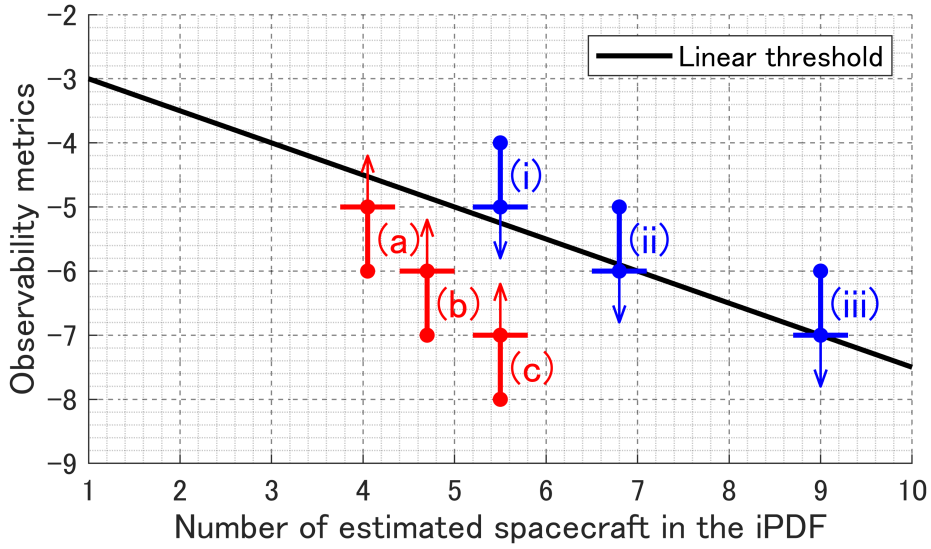
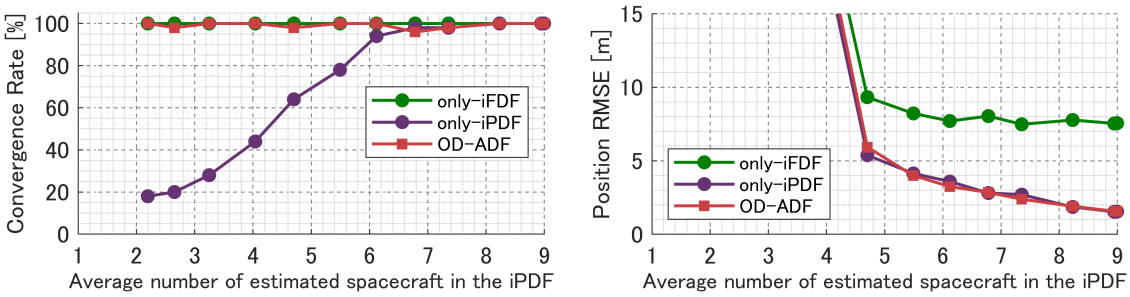


Figure 7.14: Requirements for the observability threshold and the designed linear threshold.

It is considered that the observability threshold should be adjusted according to the number of spacecraft estimated in the iPDF. In essential expression, the threshold should be adjusted by the dimensions of the variable estimated by the iPDF in the OD-ADF. We adopt the threshold of a linear function, which is expressed as:

$$\kappa_{\mathrm{thr}} = a \cdot n\left(\tilde{\mathcal{V}}^{c}_{i,k}\right) + b \tag{7.13}$$

where $a, b \in \mathbb{R}$ are constants for a linear function and $n(\tilde{\mathcal{V}}^{c}_{i,k})$ is the number of spacecraft estimated in the iPDF. We set these constants as $a = -0.5$ and $b = -2.5$, expressed as the black line in Figure 7.14. We evaluate the performance of the OD-ADF with the linear observability threshold of Eq. (7.13), whose result is shown in Figure 7.15. The OD-ADF with the tuned linear threshold successfully provides high accuracy and stability regardless of the number of spacecraft considered in the iPDF.



(a) Convergence rate

(b) Position estimation error

Figure 7.15: Convergence rate and the position estimation error in the case of the linear threshold.

7.5 Summary of OD-ADF

We summarize the algorithmic detail of the OD-ADF in Algorithm 10. The observability evaluation phase is added to the algorithm of the ADF in Algorithm 9. The switching between the iFDF and the iPDF is executed based on the observability evaluation in addition to the KL-divergence evaluation.

Algorithm 10: OD-ADF: Observability-Driven Adaptive Decentralized Filter

At the k-th time step on the i-th spacecraft;

(The following process runs on all the spacecraft.);

input : $\hat{\boldsymbol{z}}_{k-1}^{(i)i}, \boldsymbol{Z}_{k-1}^{(i)i}$

output: ${}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i}, {}^{a}\boldsymbol{Z}_{k}^{(i)i}$

Prediction step;

$$\boldsymbol{Z}_{k|k-1}^{(i)i} = \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \left(\boldsymbol{F}^{(i)} \right)^{T} + \boldsymbol{Q}^{(i)} \right)^{-1};$$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(i)i} = \boldsymbol{Z}_{k|k-1}^{(i)i} \left(\boldsymbol{F}^{(i)} \left(\boldsymbol{Z}_{k-1}^{(i)i} \right)^{-1} \hat{\boldsymbol{z}}_{k-1}^{(i)} + \boldsymbol{G}^{(i)} \boldsymbol{u}_{k}^{(i)} \right);$$

Inter-spacecraft communication step;

Collect the estimates and the related measurement data from the adjacent spacecraft.;

Internal Fully Decentralized Filtering (iFDF);

Compute the estimated positions by ;

Convert them to the information forms by ;

$${}^{f}\boldsymbol{Z}_{k|k}^{(i)i} = \omega^{(i)} \boldsymbol{Z}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{k}^{(i)}} \omega^{(j)} \boldsymbol{Z}_{k|k-1}^{(i)j};$$

$${}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \omega^{(i)} \hat{\boldsymbol{z}}_{k|k-1}^{(i)i} + \textstyle\sum_{j \in \mathcal{V}_{k}^{(i)}} \omega^{(j)} \hat{\boldsymbol{z}}_{k|k-1}^{(i)j};$$

where $\omega^{(i)} + \sum_{j \in \mathcal{V}_{k}^{(i)}} \omega^{(j)} = 1$ is satisfied.;

$$\mathcal{P}_{k}^{(i)i} = \mathcal{N}\left(\left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i} \right)^{-1} {}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i}, \ \left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i} \right)^{-1} \right);$$

(... continues to the next page...)

Internal Partially Decentralized Filtering (iPDF);

Compute the adjusted error covariance based on the local sensing networks ;

$$\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})} = \Gamma\left(\boldsymbol{R}_{s,k}^{(c)}, \mathcal{E}_{i,k}^{s}\right), \quad s \in \{r, b\} \ ;$$

Compute the information contribution in the estimated subset ;

$$\boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \boldsymbol{H}_{s,k}^{(c)} \ ;$$

$$\boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \sum_{s\in\{r,b\}} \left(\boldsymbol{H}_{s,k}^{(c)}\right)^{T} \left(\tilde{\boldsymbol{R}}_{s,k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right)^{-1} \left(\boldsymbol{y}_{s,k}^{(c)} - \hat{\boldsymbol{y}}_{s,k|k-1}^{(c)} + \boldsymbol{H}_{s,k}^{(c)} \hat{\boldsymbol{x}}_{k|k-1}^{(c)c}\right) \ ;$$

Compute the information contribution of the support spacecraft $(l \in \tilde{\mathcal{M}}_{i,k}^{c}, j \in \tilde{\mathcal{E}}_{k}^{c})$;

$$\hat{\boldsymbol{p}}_{k|k-1}^{(j)l} = \boldsymbol{g}\left(\boldsymbol{p}_{k|k-1}^{(l)l}, \boldsymbol{y}_{k}^{(jl)l}\right) \ ;$$

$$\tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l} = \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_{k}^{(l)}}\right) \boldsymbol{P}_{p,k|k-1}^{(l)l} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{p}_{k}^{(l)}}\right)^{T} + \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_{k}^{(jl)}}\right) \boldsymbol{R}_{k}^{(jl)} \left(\frac{\partial \boldsymbol{g}^{(jl)l}}{\partial \boldsymbol{y}_{k}^{(jl)}}\right)^{T} \ ;$$

Convert $\hat{\boldsymbol{p}}_{k|k-1}^{(j)l}, \tilde{\boldsymbol{P}}_{p,k|k-1}^{(j)l}$ to $\hat{\boldsymbol{u}}_{k|k-1}^{(j)l}, \boldsymbol{U}_{k|k-1}^{(j)l}$;

Execute the covariance intersection ;

$$\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_{i}\left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{U}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^{c}} \omega_{p} \boldsymbol{U}_{k|k-1}^{(j)l} \ ;$$

$$\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} = \omega_{i}\left(\hat{\boldsymbol{z}}_{k|k-1}^{(\tilde{\mathcal{V}}_{i,k}^{c})i} + \boldsymbol{u}_{k}^{(\tilde{\mathcal{V}}_{i,k}^{c})}\right) + \sum_{(j,l)\in\tilde{\mathcal{E}}_{i,k}^{c}} \omega_{p} \boldsymbol{u}_{k|k-1}^{(j)l} \ ;$$

Extract the i-th spacecraft related part;

$${}^{p}\boldsymbol{Z}_{k|k}^{(i)i} = \Upsilon^{(i)}\left[\boldsymbol{Z}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}\right], \qquad {}^{p}\hat{\boldsymbol{z}}_{k|k}^{(i)i} = \Upsilon^{(i)}\left[\hat{\boldsymbol{z}}_{k|k}^{(\tilde{\mathcal{V}}_{i,k}^{c})i}\right];$$

Observability Evaluation Phase;

$$\kappa_{\text{thr}} = a \cdot n\left(\tilde{\mathcal{V}}_{i,k}^{c}\right) + b;$$

$$\kappa_{k}^{i} = \int f_{obs}\left(\boldsymbol{x}_{k}^{(\mathcal{V}_{i,k}^{c})}\right) \mathcal{N}\left(\boldsymbol{x}_{k}^{(\mathcal{V}_{i,k}^{c})} \middle| \hat{\boldsymbol{x}}_{k|k-1}^{(\mathcal{V}_{i,k}^{c})i}, \boldsymbol{P}_{k|k-1}^{(\mathcal{V}_{i,k}^{c})i}\right) d\boldsymbol{x}_{k}^{(\mathcal{V}_{i,k}^{c})};$$

Fusing Phase;

if $\sum_{l=1}^{\Delta k} \text{KL}\left({}^{f}\mathcal{P}_{k}^{(i)i}, {}^{f}\mathcal{P}_{k-l}^{(i)i}\right) \leq \Delta P \bigwedge \kappa_{k}^{i} > \kappa_{\text{thr}}$ **then**

$${}^{a}\boldsymbol{Z}_{k}^{(i)i} = \omega\left({}^{f}\boldsymbol{Z}_{k|k}^{(i)i}\right) + (1-\omega)\left({}^{p}\boldsymbol{Z}_{k|k}^{(i)i}\right);$$

$${}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i} = \omega\left({}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i}\right) + (1-\omega)\left({}^{p}\hat{\boldsymbol{z}}_{k|k}^{(i)i}\right);$$

$$\omega = \arg\min_{\omega}\left[\text{trace}\left({}^{a}\boldsymbol{Z}_{k|k}^{(i)i}\right)^{-1}\right];$$

else

$${}^{a}\boldsymbol{Z}_{k}^{(i)i} = {}^{f}\boldsymbol{Z}_{k|k}^{(i)i}, \quad {}^{a}\hat{\boldsymbol{z}}_{k}^{(i)i} = {}^{f}\hat{\boldsymbol{z}}_{k|k}^{(i)i};$$

end

Chapter 8

Verification by Numerical Simulations

8.1 Overview of Simulation-based Verification

To compare the performance of the proposed algorithms with the existing estimation algorithms and clarify their effectiveness, we evaluate their performances by numerical simulations. Table 8.1 shows seven estimation algorithms whose performances will be evaluated in this chapter. The set of filters to be assessed is defined as:

$$X_f = \{\text{CF}, \text{FDF}, \text{R-FDF}, \text{PDF}, \text{R-PDF}, \text{ADF}, \text{OD-ADF}\}\,. \tag{8.1}$$

The simulation conditions in this chapter follow the parameters listed in Table 4.1. We assume that the network structures for one simulation scenario are identical, that is $\mathcal{G}^c = \mathcal{G}^r = \mathcal{G}^b$, regardless of the evaluation in Section 8.5.

8.2 Dependency on Swarm Configurations

The performance of relative navigation for spacecraft swarms is greatly affected by swarm configurations. The information gain of relative measurements is a function of spacecraft positions; hence, swarm configurations possibly have negative impacts on navigation algorithms. Furthermore, as shown in Figure 1.10, a swarm configuration would dynamically change with time, especially for reconfiguration maneuvers. Therefore, a navigation algorithm needs to handle various swarm configurations to ensure safe operations like this type of dynamical maneuvers.

Table 8.1: Evaluated estimation algorithms

No.	Acronym	Full name	Detail
(1)	CF	Centralized Filter	Algorithm 4
(2)	FDF	Fully Decentralized Filter	Algorithm 5
(3)	R-FDF	Robust Fully Decentralized Filter	Algorithm 6
(4)	PDF	Partially Decentralized Filter	Algorithm 7
(5)	R-PDF	Robust Partially Decentralized Filter	Algorithm 8
(6)	ADF	Adaptive Decentralized Filter	Algorithm 9
(7)	OD-ADF	Observability-Driven Adaptive Decentralized Filter	Algorithm 10

8.2.1 Randomly Generated Configurations

To validate whether the navigation algorithms can manage various swarm configurations, we ran the simulations on 50 types of configurations ($M = 50$) as illustrated in Figure 8.1. The values of initial states are generated from continuous uniform distributions, which are expressed as $\mathcal{U}_p(-\Delta\boldsymbol{p}^u, \Delta\boldsymbol{p}^u)$ for positions and $\mathcal{U}_v(-\Delta\boldsymbol{v}^u, \Delta\boldsymbol{v}^u)$ for velocities. When we define the probability density functions of $\mathcal{U}_p$ and $\mathcal{U}_v$ as $f(\boldsymbol{p})$ and $f(\boldsymbol{v})$ respectively, these functions are formulated as:

$$f(p_\alpha) = \begin{cases} \frac{1}{2\Delta p^u_\alpha} & \text{if } -\Delta p^u_\alpha \leq p_\alpha \leq \Delta p^u_\alpha \\ 0 & \text{otherwise} \end{cases} \tag{8.2}$$

$$f(v_\alpha) = \begin{cases} \frac{1}{2\Delta v^u_\alpha} & \text{if } -\Delta v^u_\alpha \leq v_\alpha \leq \Delta v^u_\alpha \\ 0 & \text{otherwise} \end{cases} \tag{8.3}$$

where $\alpha = \{x, y, z\}$ and $\Delta p^u_\alpha, \Delta v^u_\alpha > 0$. In the simulations, these parameters for the uniform distributions are set as $\Delta p^u_\alpha = 1000$[m] and $\Delta v^u_\alpha = 10$[m].

Estimation stability

Figure 8.2 shows the convergence rate of each estimation algorithm with respect to the average connection rate defined by Eq. (2.35). As a general tendency, the convergence rate

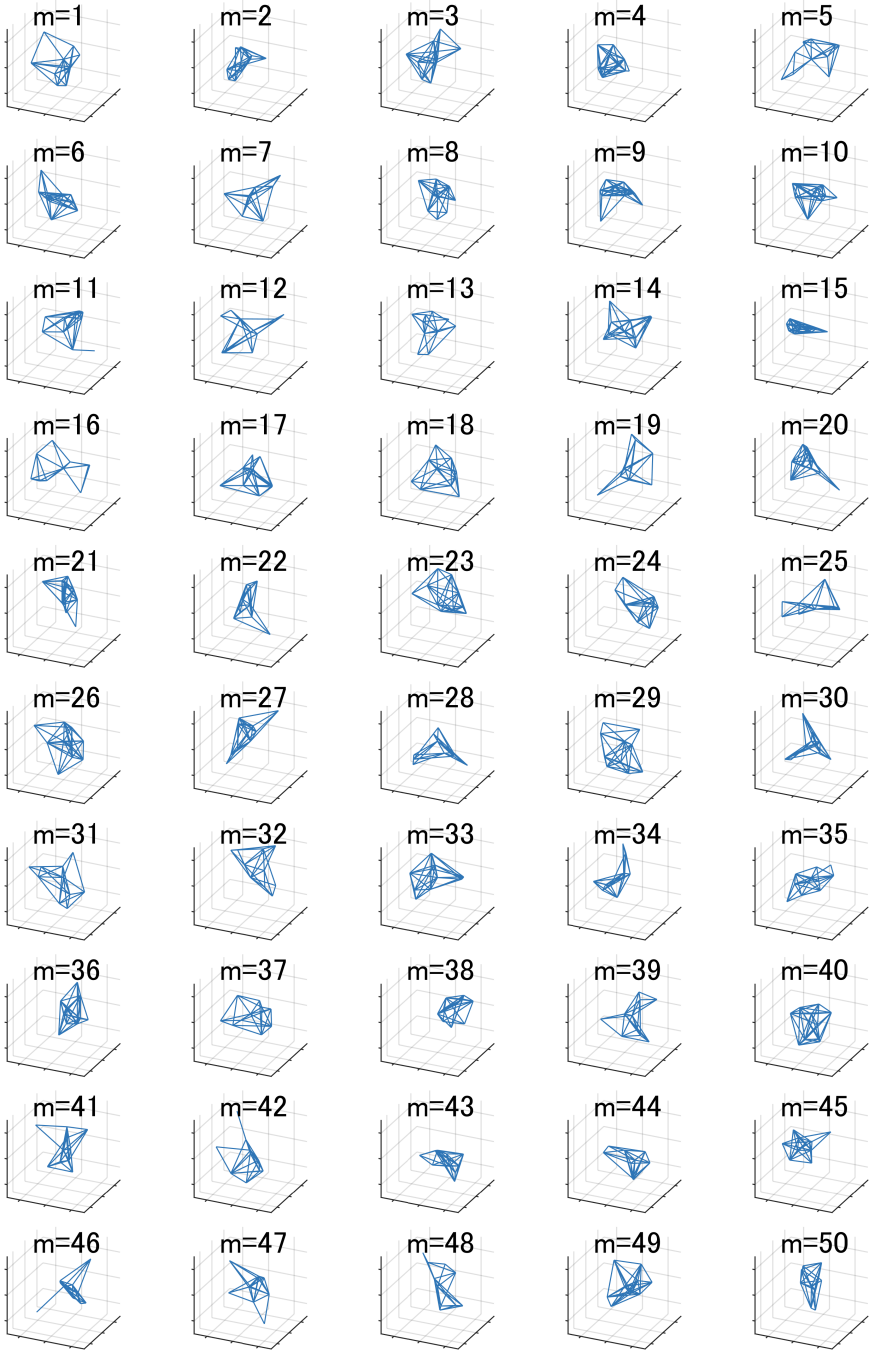


Figure 8.1: Swarm configurations and the corresponding networks ($\Delta R = 1500$m in this figure) evaluated in the numerical simulations.

gets lower when the network density is low. This is mainly because the observability tends to be insufficient when the sensing networks are sparse. The ADF provides meaningfully higher estimation performance than the CF, the FDF, the PDF, and the R-PDF. This result represents that the fusing process of the ADF significantly improves the estimation stability of the partially decentralized filters. However, the estimates of the ADF results in divergence in some cases when the sensing networks are sparse. On the other hand, the R-FDF runs its estimation process without divergence regardless of the network thresholds. By relying on the high stability of the R-FDF when the sensing networks do not hold sufficient observability, the OD-ADF achieves a higher convergence rate than the ADF.

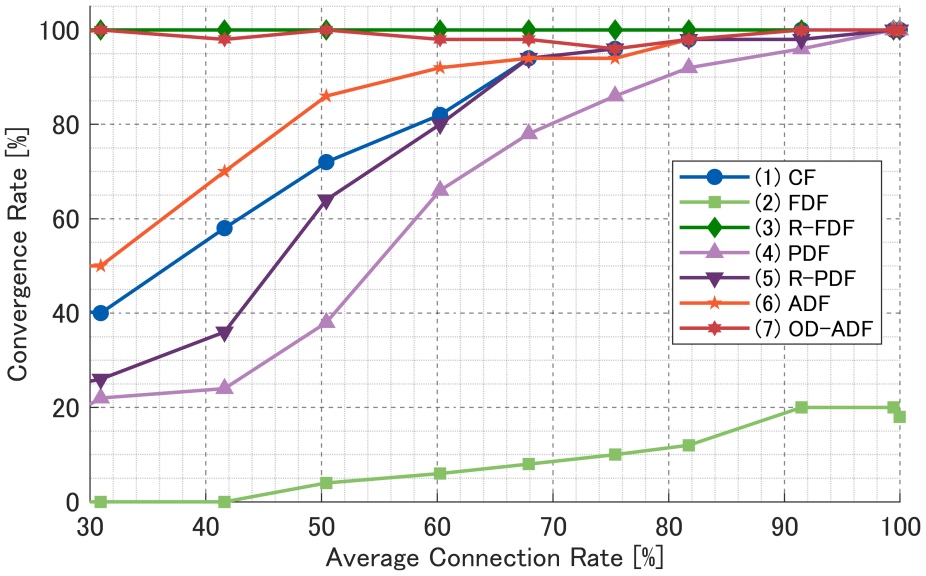


Figure 8.2: Convergence rate (the dependency on swarm configurations).

Figure 8.3 shows the average connection rate for each network threshold. When the network threshold is longer than 2500m, the network connection rate reaches 100%, meaning that the communication and sensing networks are complete graphs.

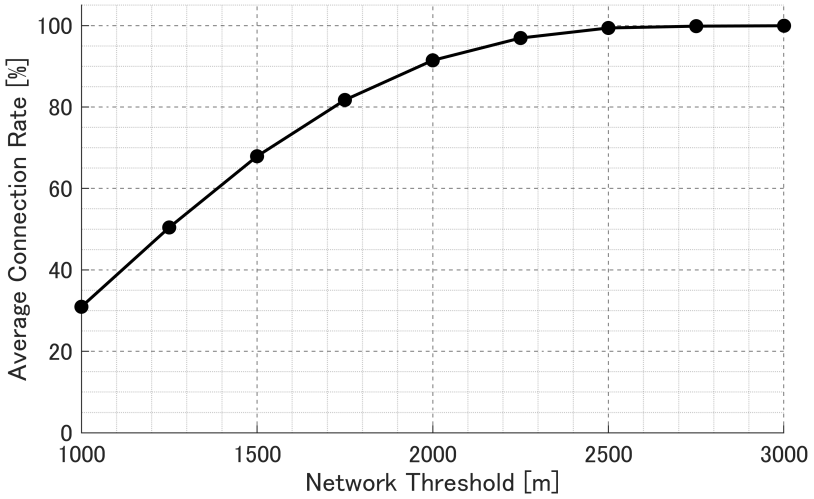


Figure 8.3: Network connection rate (the dependency on swarm configurations).

Estimation accuracy

Next, we investigate the estimation accuracy. To realize the meaningful comparison of the estimation error of each algorithm, it is essential to compare the results under the same conditions, including swarm configurations, network thresholds, and convergence. Figure 8.4 shows whether the estimation results of all the filters are converged or not for the specific network threshold and swarm configuration. The results of the FDF and the PDF are excluded since their convergence rate is remarkably lower than the other algorithms, as shown in Figure 8.2.

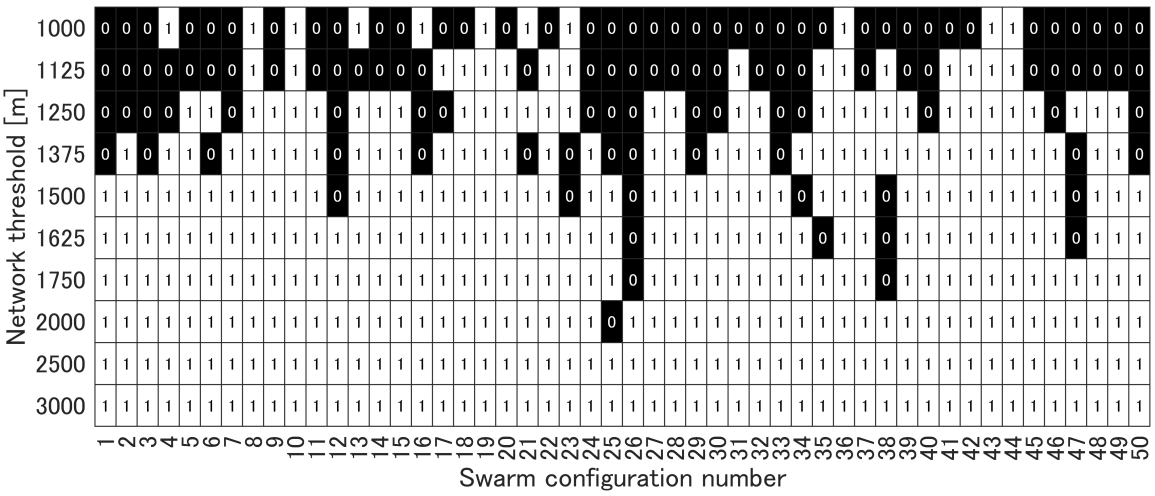
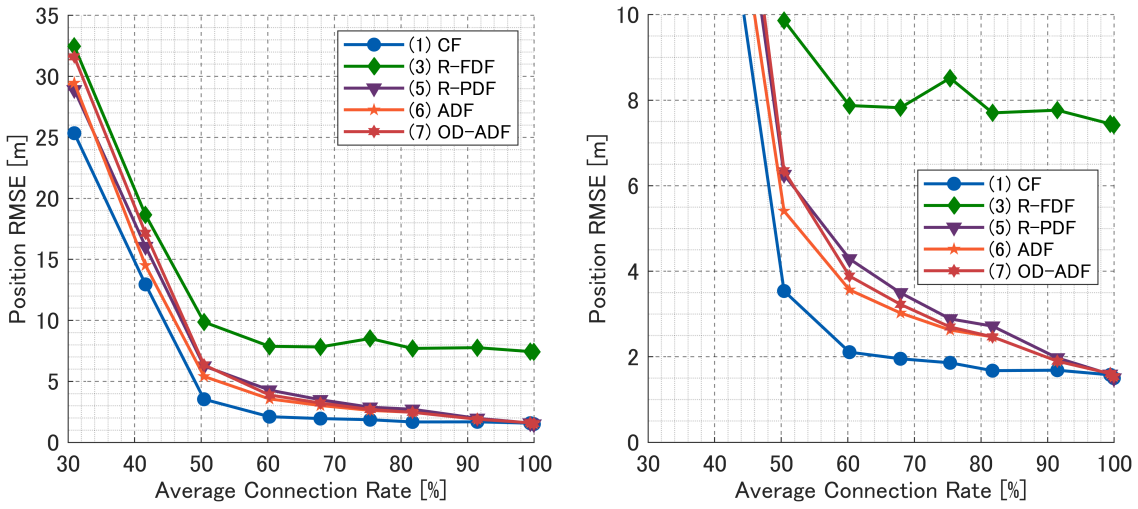


Figure 8.4: Gross convergence flags (the dependency on swarm configurations).

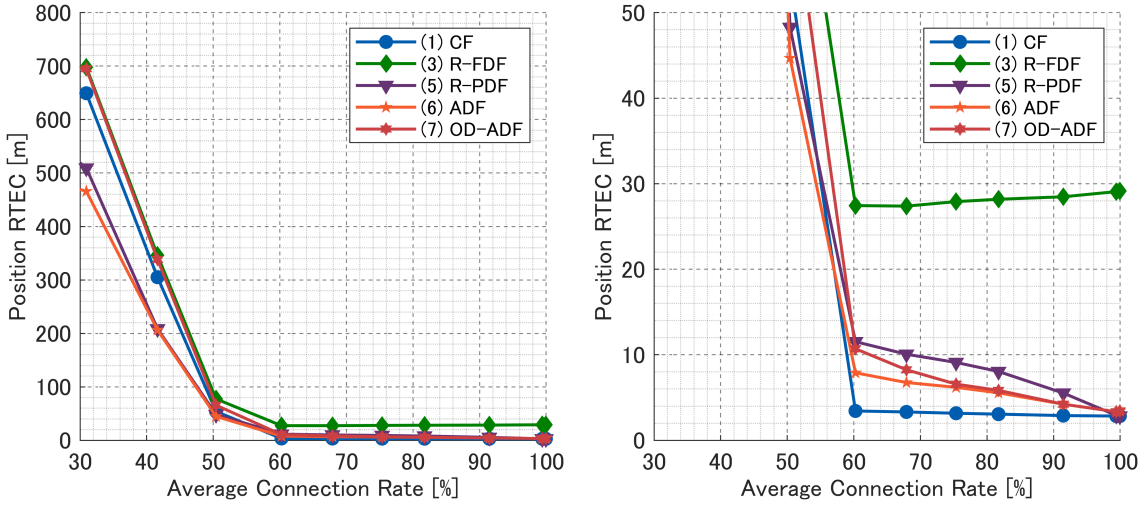
Figures 8.5 and 8.6 respectively show the average RMSE and the average RTEC of the

position estimation. The estimation accuracy of the CF is higher than those of the other filters since the CF can leverage all the information to update the state probability. On the other hand, the R-FDF provides the estimation results with remarkably lower estimation accuracy than the other filters. These facts follow the evaluation results in Section 5.4. The R-PDF, the ADF, and the OD-ADF generate slightly less accurate estimate results than the CF.



(a) Overall view **(b)** Extended view

Figure 8.5: Root mean square errors of the estimated positions (the dependency on swarm configurations).



(a) Overall view **(b)** Extended view

Figure 8.6: Root of the trace of the position error covariance matrix (the dependency on swarm configurations).

Figure 8.7 shows the RMSE of the estimated position of each spacecraft computed by all the algorithms. The result shows the case of the 43rd swarm configuration with the 1500m network threshold. The estimation error of each algorithm is defined by the formula listed in Table 8.2.

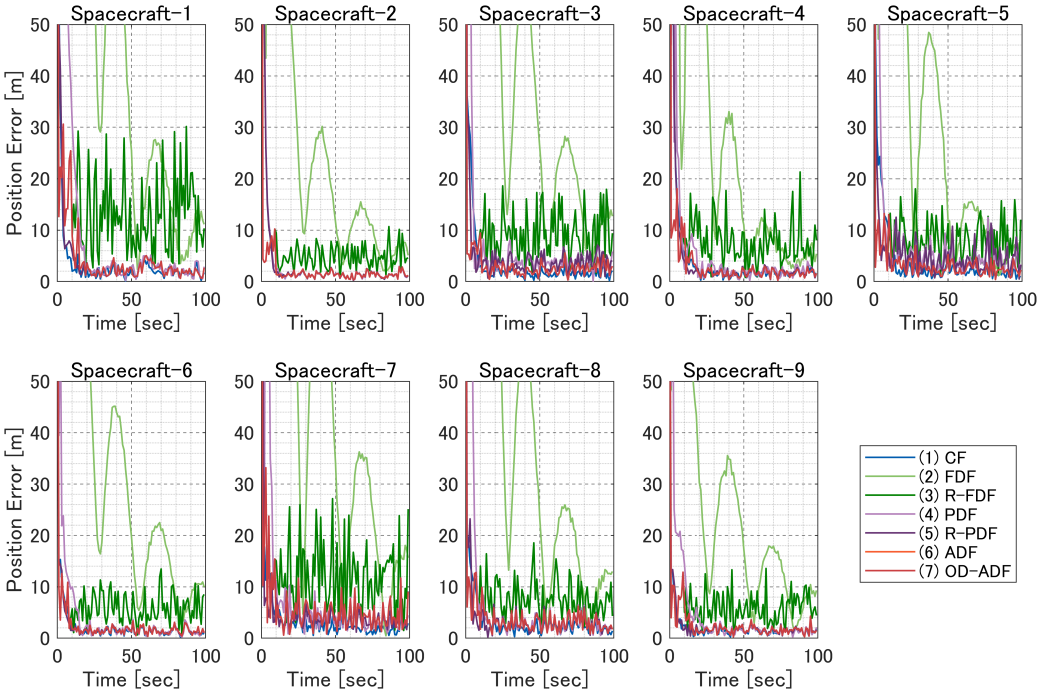


Figure 8.7: Position estimation error for every spacecraft in the 43rd swarm configuration with the 1500m network threshold (the dependency on swarm configurations).

Table 8.2: Definition of the position estimation error for each algorithm

Algorithm	Definition
(1) CF	$\left\| \Upsilon_p^{(i)} \left(\hat{\boldsymbol{x}}_k^{(c)} \right) - \boldsymbol{p}_k^{(i)} \right\|_2$
(2) FDF, (3) R-FDF	$\left\| \hat{\boldsymbol{p}}_k^{(i)} - \boldsymbol{p}_k^{(i)} \right\|_2$
(4) PDF, (5) R-PDF	$\left\| \Upsilon_p^{(i)} \left(\hat{\boldsymbol{x}}_k^{(\tilde{\mathcal{V}}_{i,k}^c)i} \right) - \boldsymbol{p}_k^{(i)} \right\|_2$
(6) ADF, (7) OD-ADF	$\left\| {}^a\hat{\boldsymbol{p}}_k^{(i)} - \boldsymbol{p}_k^{(i)} \right\|_2$

Figure 8.8 shows the position estimation error of the 1st spacecraft evaluated by the ADF on the 1st spacecraft. In this case, the ADF changes its estimation mode from the stability-oriented mode to the accuracy-oriented mode, which are defined in Section 6.2, at

$t = 11$[sec]. After the mode is switched, the estimation errors remarkably get smaller.

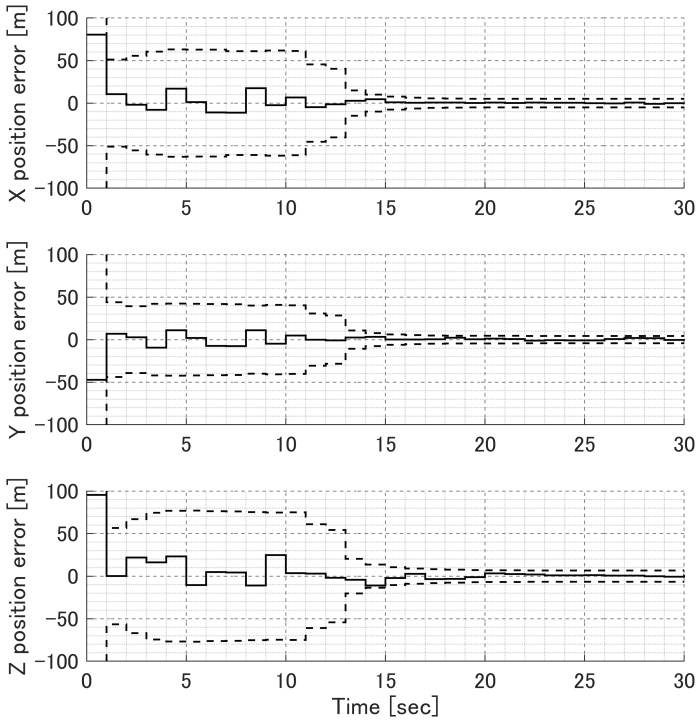


Figure 8.8: Position estimation error of the 1st spacecraft calculated by the ADF in the 43rd swarm configuration with the 1500m network threshold (solid line: the difference between the true and the estimated mean values, dashed line: 3-sigma error bound).

Communication load

Figures 8.9 and 8.10 show the average communication load of each estimation algorithm, which is defined by Eqs. (4.9) to (4.11). The CF requires a significantly higher communication load than the other decentralized filters since it needs the large-data transfer to perform the estimation computation on the leader spacecraft. The FDF takes much less communication load than the other filters because it does not send the covariance information. However, the FDF provides unacceptable estimation stability, as shown in Figure 8.2. In the same way, the R-PDF requires more communication resources than the PDF since an estimator receives the additional covariance information from support spacecraft in the R-PDF. Although the partially decentralized filters (the PDF, the R-PDF) take a higher communication load than the fully decentralized filters (the FDF, the R-FDF), the additional load for the partially decentralized filters is within the allowable range. As shown in Figure 8.10, the adaptive decentralized filters (the ADF, the OD-ADF) take the same amount of communication load as the R-PDF.

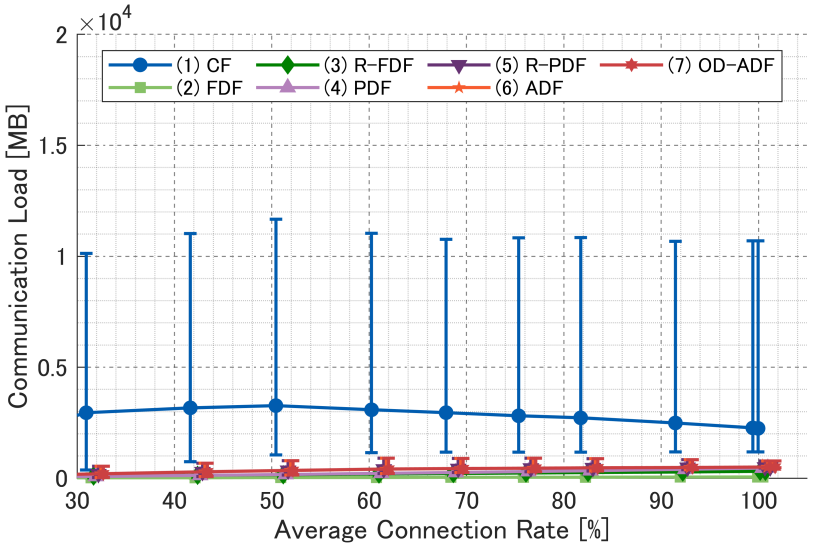


Figure 8.9: Communication load analysis (the top edges of the bars: maxCL, the bottom edges of the bars: minCL, and the circles on the lines: aveCL).

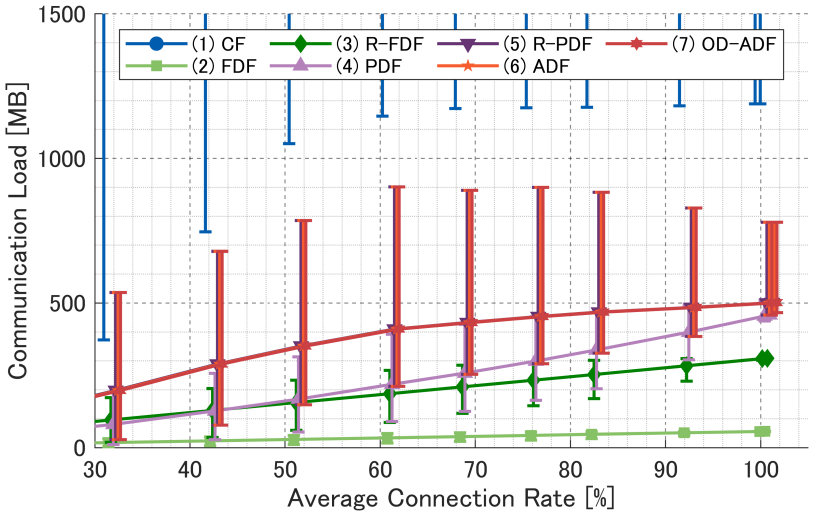


Figure 8.10: Communication load analysis: the extended view of Figure 8.9 (the top edges of the bars: maxCL, the bottom edges of the bars: minCL, and the circles on the lines: aveCL).

Discussion on the estimation results of the R-PDF and the ADF

Here we discuss the relationship of the estimation accuracy between the R-PDF and the ADF. When comparing the estimation accuracy of these filters, it is intuitively expected that the R-PDF provides more accurate estimates than the ADF for swarm configurations in which both filters are converged. That is because the ADF uses the estimation result

of the fully decentralized filter, which typically provides less accurate estimates than the partially decentralized filter. However, Figures 8.5 and 8.6 show that the estimated position of the ADF is more accurate than that of the R-PDF.

Figure 8.11 shows the RMSE of estimation position of each spacecraft for the 44th swarm configuration, which is one of the configurations whose estimation results are included in Figures 8.5 and 8.6. For the 1st to 7th spacecraft, the R-PDF and the ADF provide the estimated values with almost the same accuracy. However, for the 8th and 9th spacecraft, the estimation accuracy of the R-PDF is remarkably worse than that of the ADF.

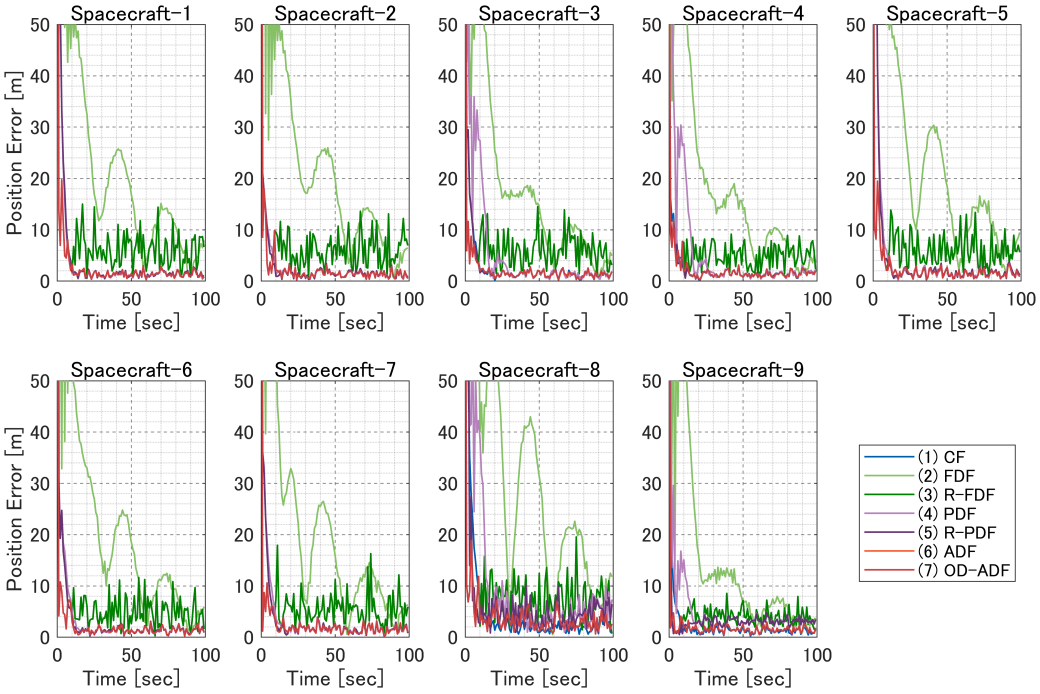
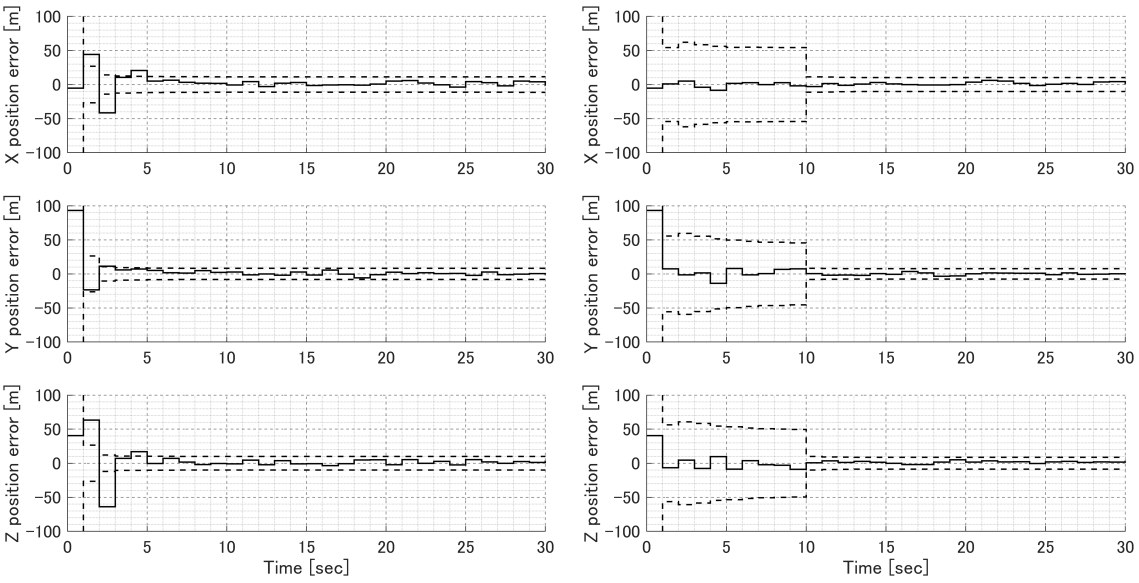


Figure 8.11: Position estimate error for every spacecraft in the 44th swarm configuration with the 1500m network threshold.

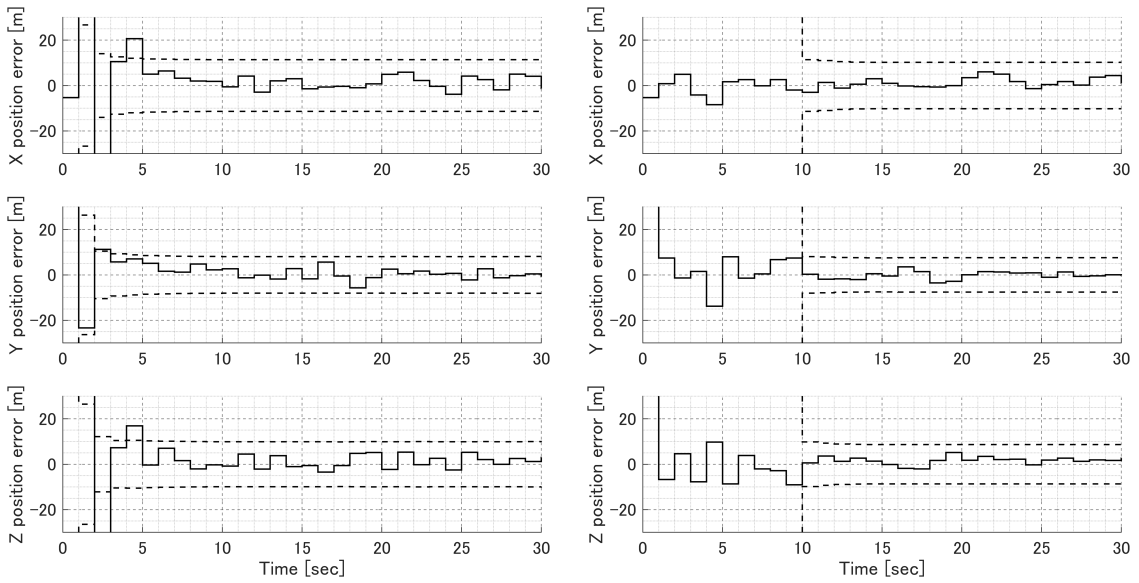
Figures 8.12 and 8.13 show the position estimation errors of the 8th spacecraft, computed by the R-PDF and the ADF. As seen in Figure 8.12, the R-PDF executes the measurement update of the partially decentralized estimation at the beginning. Due to the approximation of the nonlinear measurement equation, the likelihood with non-negligible error would be used for the measurement update in the early estimation phase. On the other hand, the ADF starts the estimation from the fully decentralized estimation and then switches to the partially decentralized estimation. In other words, the ADF executes the conservative measurement update when the evaluated likelihood possibly has large errors and starts more aggressive measurement updates after the accurate evaluation of

likelihood is available. Therefore, it is considered that the ADF enables more accurate estimation than the R-PDF by preventing the likelihood evaluation that possibly generates linearization errors, especially for the early estimation phase.



(a) R-PDF **(b)** ADF

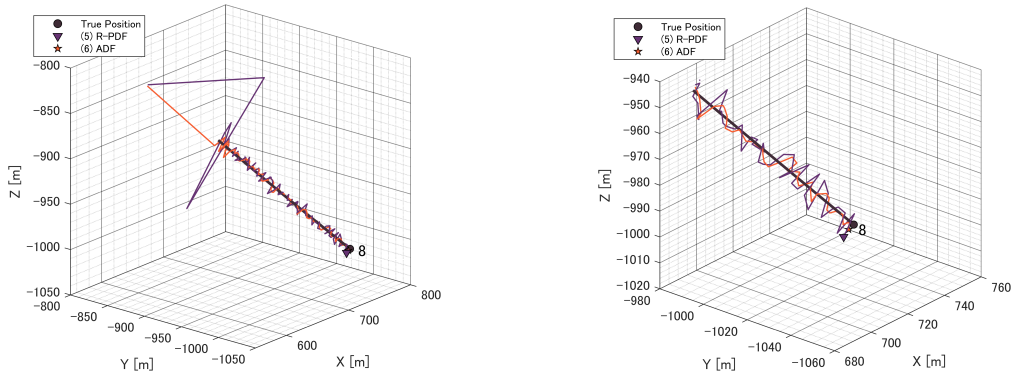
Figure 8.12: Position estimation errors of the 8th spacecraft calculated by (a) R-PDF and (b) ADF in the 44th swarm configuration with the 1500m network threshold (solid line: the RMSE, dashed line: 3-sigma error bound).



(a) R-PDF **(b)** ADF

Figure 8.13: Extended view of Figure 8.12.

Figure 8.14 plots the trajectories of the 8th spacecraft estimated by the R-PDF and the ADF. As we discussed above, it can be seen that the estimation result of the R-PDF is converged to the position with more estimation errors than that of the ADF.



(a) Overall view **(b)** Extended view

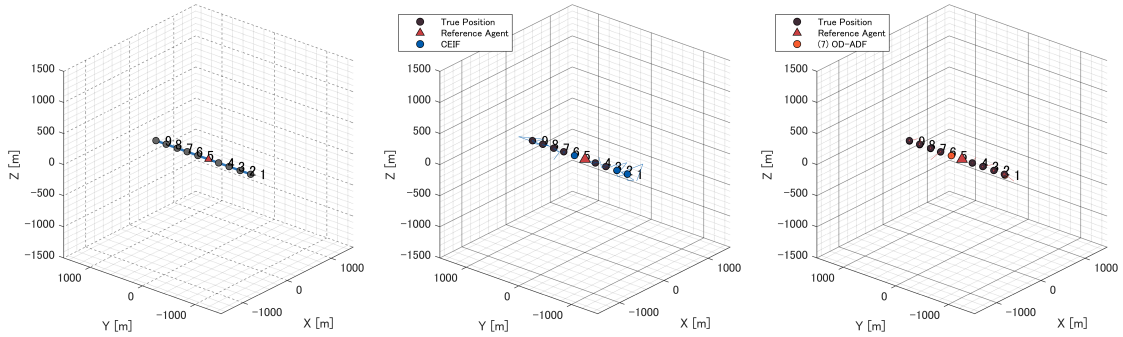
Figure 8.14: Trajectory of 8th spacecraft estimated by the R-PDF and the ADF.

8.2.2 Coordinated Configurations

For the practical applications of spacecraft swarms, spacecraft are typically controlled to form coordinated configurations for mission purposes. Although the main targets of this dissertation are complex configurations during the reconfiguration maneuvers such as Figure 1.10, we also investigate the performances of the estimation algorithms for representative configurations, including the in-train geometry and the general circular orbit.

In-train geometry

Firstly, we consider an in-train configuration as shown in Figure 8.15-(a). Each spacecraft is connected by a single communication/sensing link with the adjacent spacecraft. Figures 8.15-(b) and (c) illustrate the three-dimensional visualization of the estimated positions by the CF and the OD-ADF, respectively. Figure 8.16 shows the RMSE of all the filters. Except for the FDF and the PDF, all the filters provide reasonable estimation results without divergence.



(a) In-train geometry **(b)** Estimates by the CF **(c)** Estimates by the OD-ADF

Figure 8.15: In-train configuration and the estimated positions by the CF and the OD-ADF.

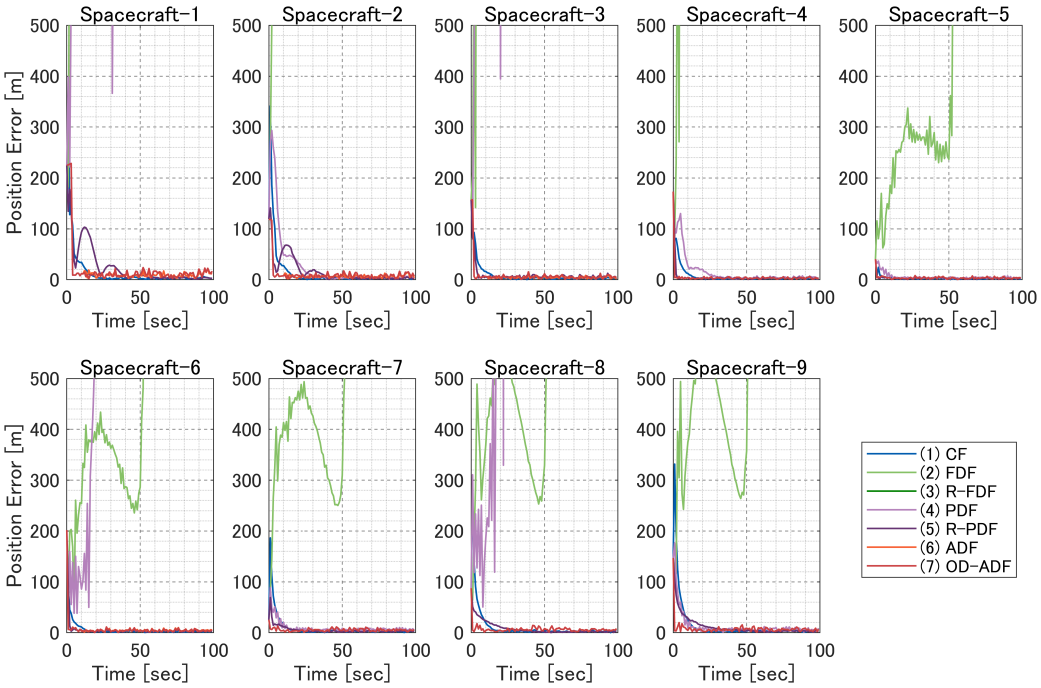
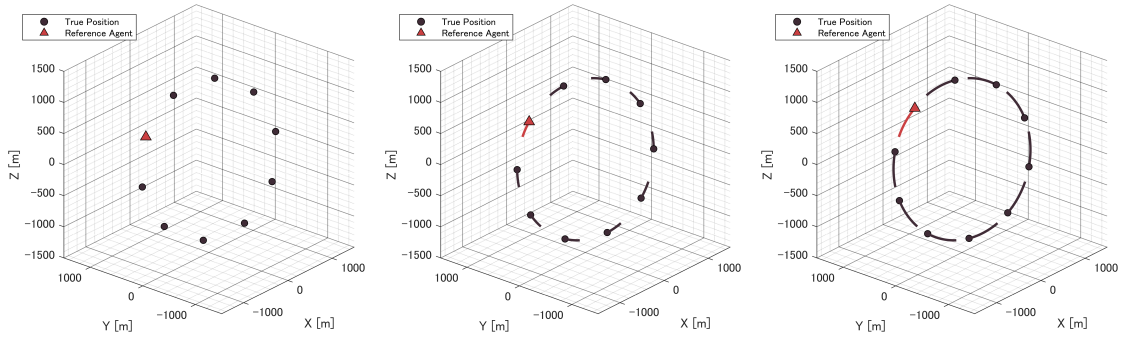


Figure 8.16: Position estimation errors for the in-train configuration.

General Circular Orbit

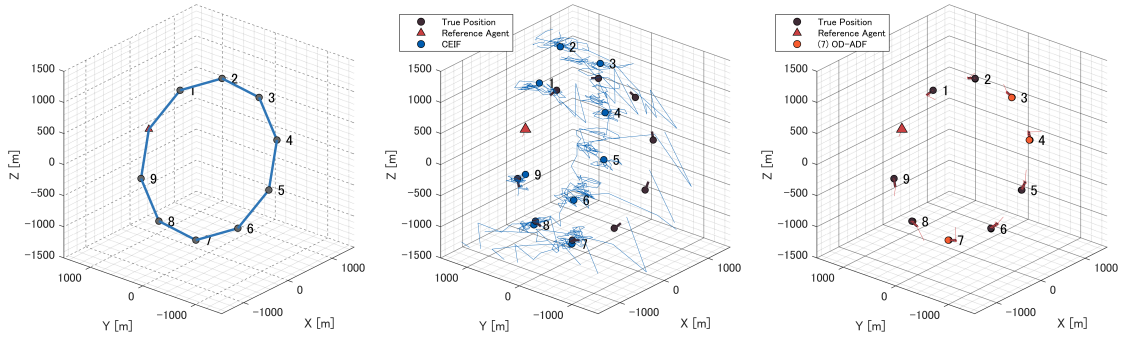
We consider the General Circular Orbit (GCO) as the second example of coordinated configurations [44]. As shown in Figure 8.17, all the spacecraft are located on the three-dimensional circular orbit. The spacecraft arrangement on the GCO is typically applied for pseudo large apertures, as illustrated in Figure 1.2.



(a) $t = 0$[sec] **(b)** $t = 200$[sec] **(c)** $t = 400$[sec]

Figure 8.17: Spacecraft on the General Circular Orbit (GCO).

Firstly, we consider the ring-shaped network of Figure 8.18-(a), in which each spacecraft is connected to the adjacent spacecraft with the network threshold $\Delta R = 1000$[m]. Figure 8.19 shows the RMSE of estimated positions. The estimation results of the CF converge to incorrect values as visualized in Figure 8.18-(b); on the other hand, the OD-ADF successfully provides reasonable estimated positions as visualized in Figure 8.18-(c).



(a) ring-shaped network **(b)** Estimates by the CF **(c)** Estimates by the OD-ADF

Figure 8.18: ring-shaped network for the GCO ($\Delta R = 1000$[m]), and the estimated positions by the CF and the OD-ADF.

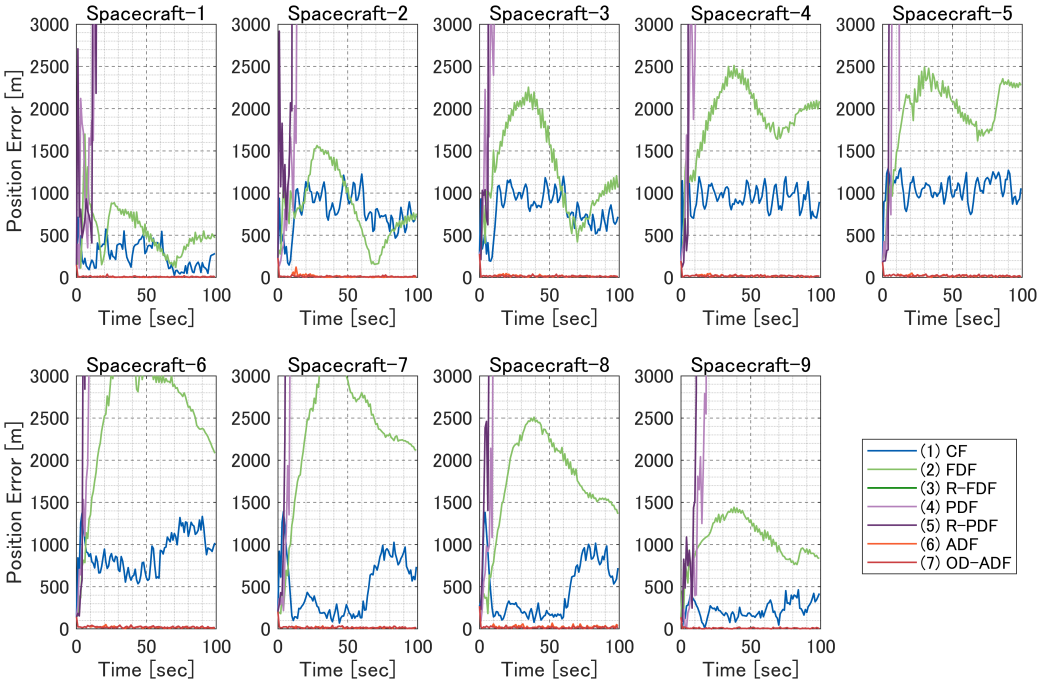
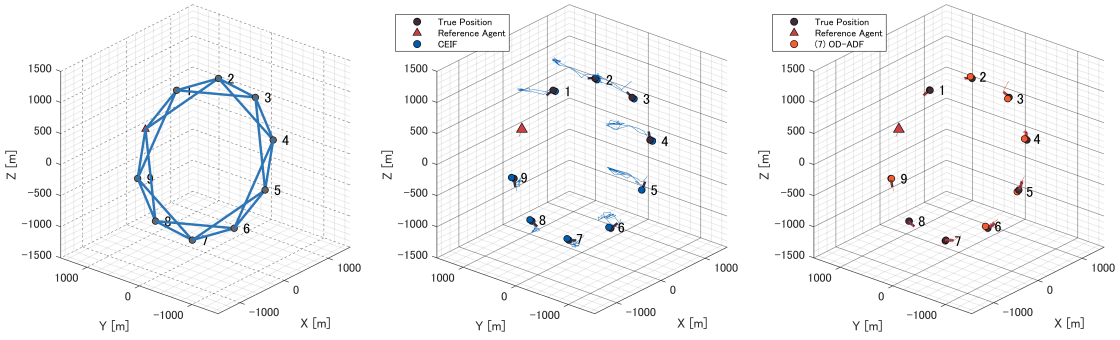


Figure 8.19: Position estimation errors for the GCO configuration ($\Delta R = 1000$[m]).

As shown in Figure 8.20-(a), we consider adding more communication and sensing links. In this case, the CF provides reasonable estimates since it can compute more accurate likelihood with additional sensing information, though it still has non-negligible offset error as shown in Figure 8.20-(b) and Figure 8.21. On the other hand, the OD-ADF successfully estimates the positions as visualized in Figure 8.20-(c), similar to the sparse ring-shaped network of Figure 8.18-(a).



(a) ring-shaped network **(b)** Estimates by the CF **(c)** Estimates by the OD-ADF

Figure 8.20: ring-shaped network for the GCO ($\Delta R = 1500$[m]), and the position estimates by the CF and the OD-ADF..

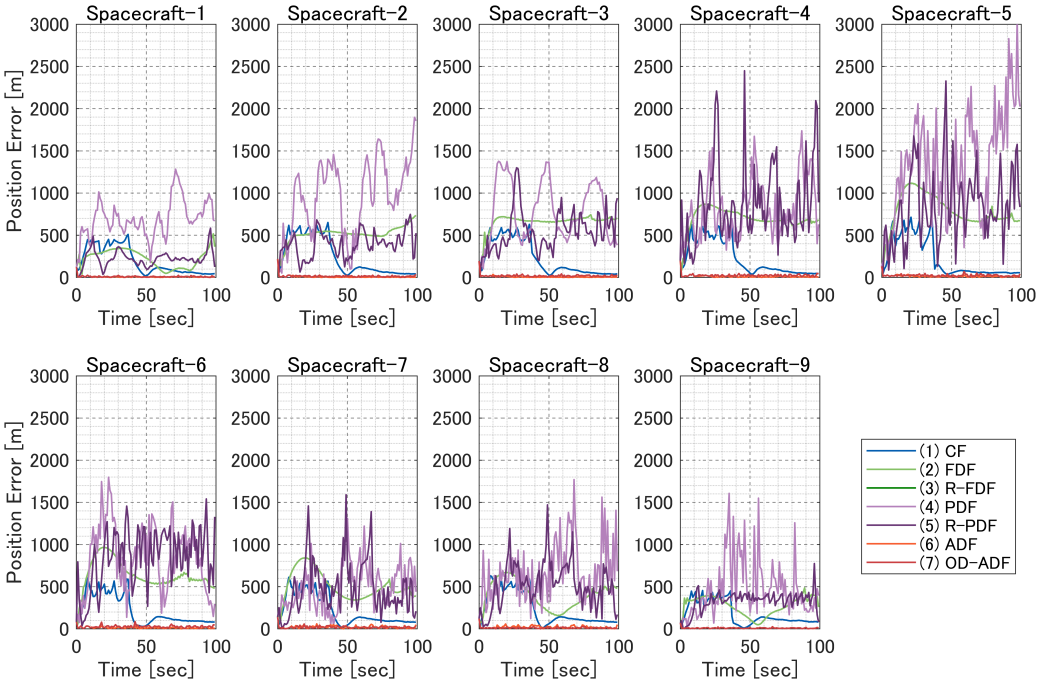


Figure 8.21: Position estimation errors for the GCO configuration ($\Delta R = 1500$[m]).

8.3 Sensitivity to Initial Estimation Errors

Initial estimation errors also have a significant influence on estimation performances, especially when dealing with nonlinear propagation and measurement update. We consider several scenarios with different amounts of initial estimation errors to evaluate adaptability for a wide range of initial estimation errors, as illustrated in Figure 8.22.

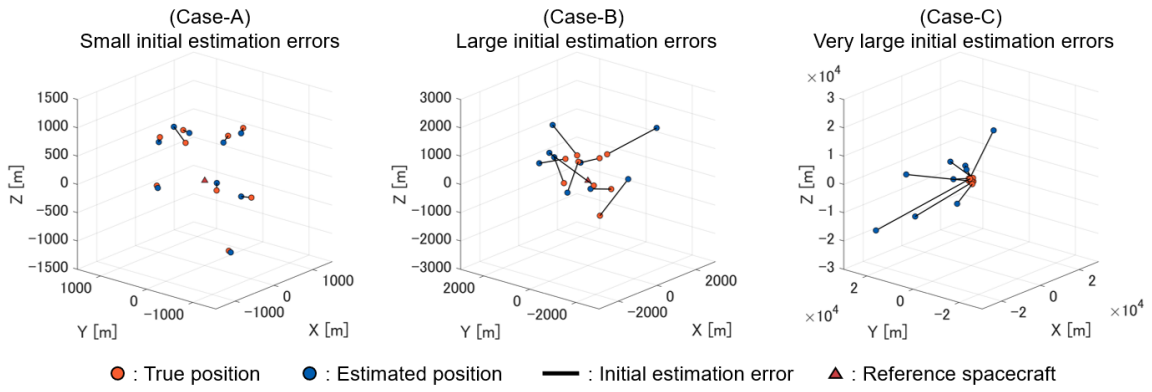


Figure 8.22: Initial position estimation errors of Case-A, Case-B, and Case-C, showing the case of the 1st swarm configuration.

The covariance parameters for the initial estimation error $\sigma_{p_{init}}$ and $\sigma_{v_{init}}$, which are

defined in Section 4.2.2, are set as the values summarized in Table 8.3.

Table 8.3: Parameters to generate initial estimation errors

Parameter	Value	Units
Case-A: Small initial estimation errors		
position error ($\sigma_{p_{init}}$)	100.0	[m]
velocity error ($\sigma_{v_{init}}$)	1.0	[m/s]
Case-B: Large initial estimation errors		
position error ($\sigma_{p_{init}}$)	1000.0	[m]
velocity error ($\sigma_{v_{init}}$)	10.0	[m/s]
Case-C: Very large initial estimation errors		
position error ($\sigma_{p_{init}}$)	10000.0	[m]
velocity error ($\sigma_{v_{init}}$)	100.0	[m/s]

For the values of the initial estimation errors listed in Table 8.3, we have evaluated the convergence rate of all the estimation algorithms in Table 8.1. The parameter settings are the same as the parameters stated in Section 4.2. For the swarm configurations, the analysis in this section also considers various swarm configurations as Figure 8.1. Figures 8.23, 8.24, and 8.25 show the convergence rates for the initial estimation errors: Case-A, Case-B, and Case-C, respectively. When the initial estimation error is small, the existing algorithms like the CF, the PDF, and the R-PDF provide relatively high estimation stability, as shown in Figure 8.23. However, when the initial estimation errors get large, these algorithms no longer provide a sufficiently high convergence rate, as shown in Figures 8.24 and 8.25. On the other hand, the proposed algorithms, the ADF and the OD-ADF, keep high estimation stability even with large initial estimation errors.

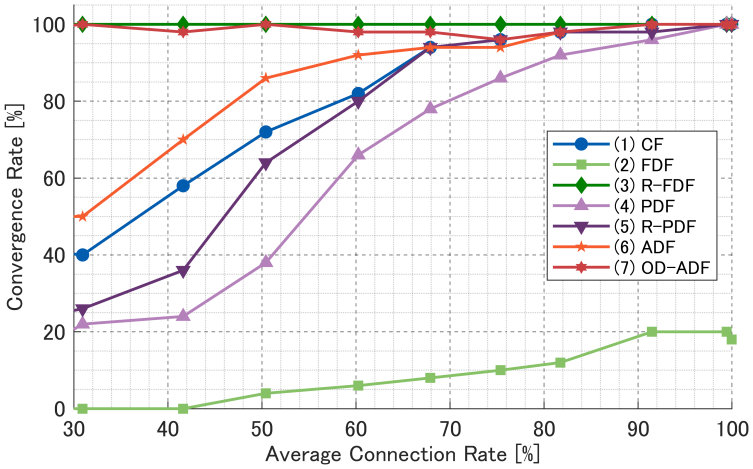


Figure 8.23: Convergence rate of the Case-A: small initial estimation errors.

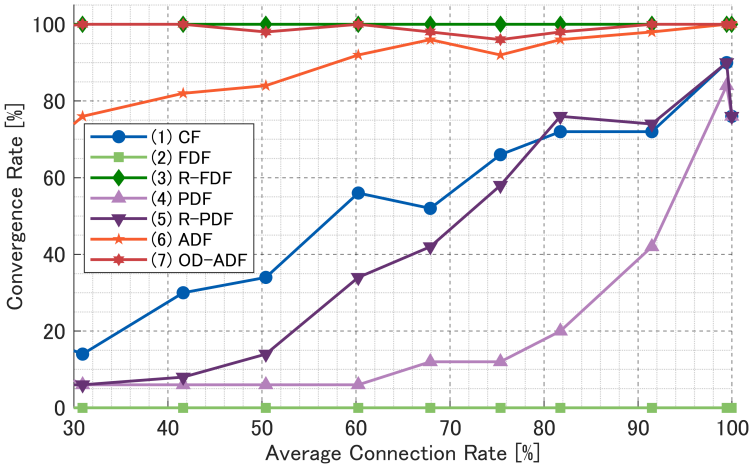


Figure 8.24: Convergence rate of the Case-B: large initial estimation errors.

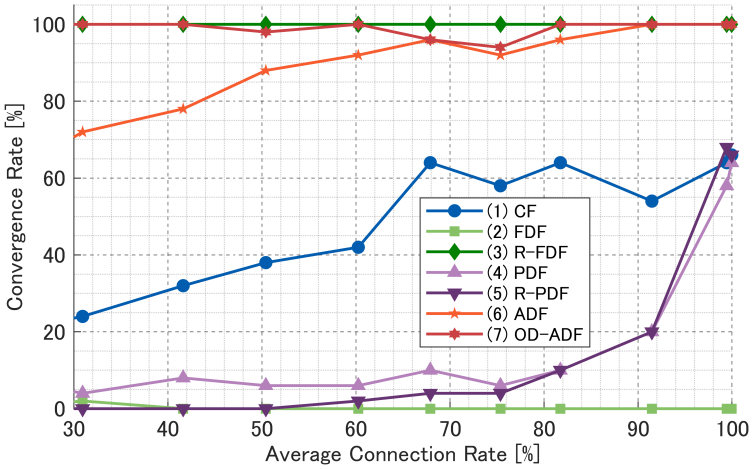


Figure 8.25: Convergence rate of the Case-C: very large initial estimation errors.

Figures 8.26, 8.27, and 8.28 show the gross convergence flags for Case-A, Case-B, and Case-C, respectively. In these figures, the results of the FDF and the PDF are excluded.

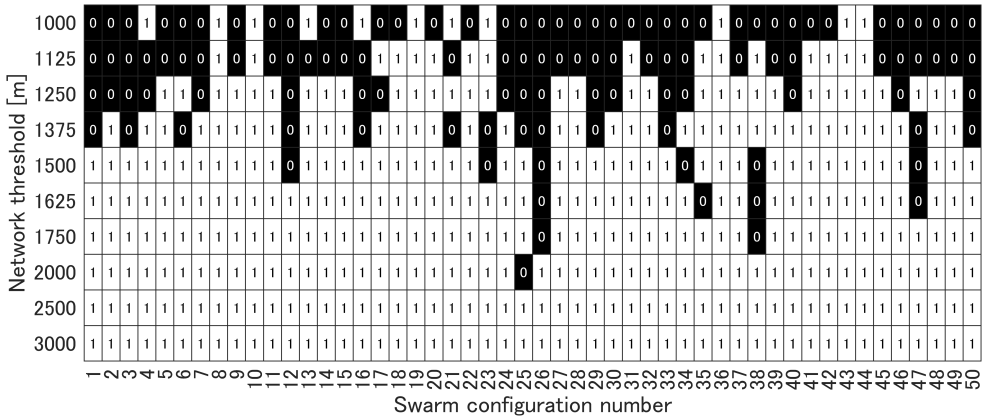


Figure 8.26: Gross convergence flags (Case-A: small initial estimation errors).

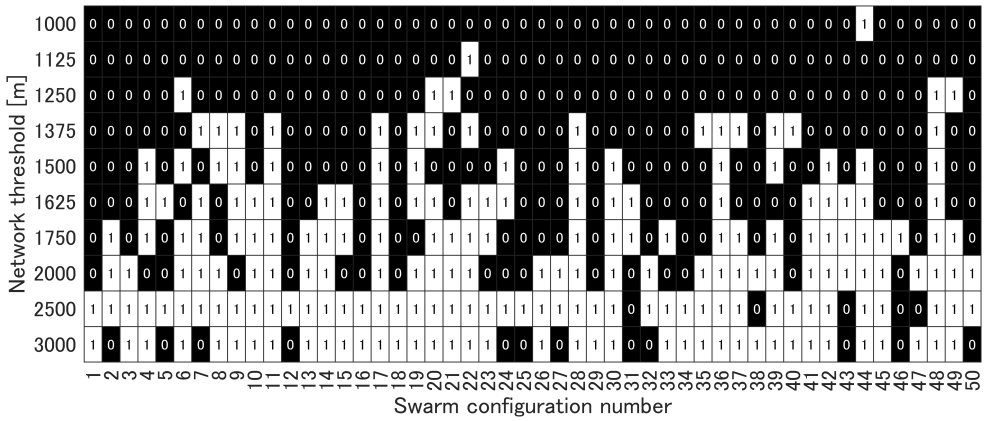


Figure 8.27: Gross convergence flags (Case-B: large initial estimation errors).

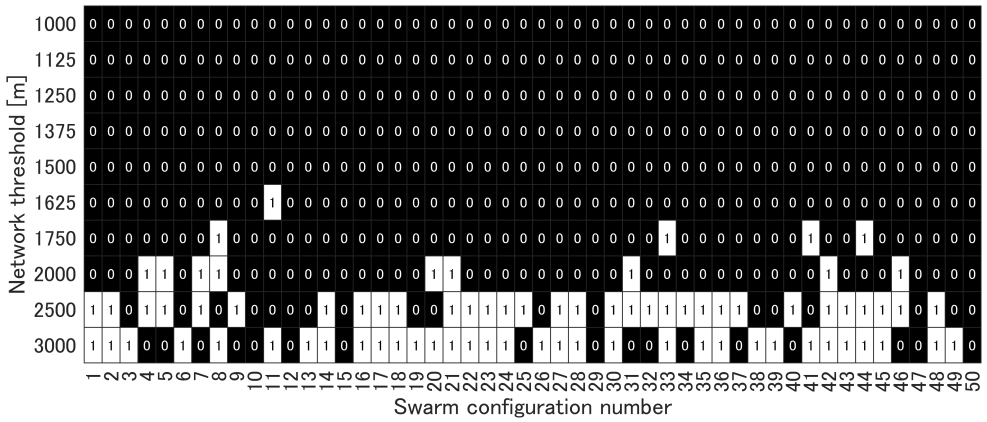


Figure 8.28: Gross convergence flags (Case-C: very large initial estimation errors).

8.4 Fault Tolerance Performance

In a practical situation, a swarm system would experience the malfunction of one of the member spacecraft or its onboard sensors. Therefore, the fault tolerance of an estimation algorithm is another essential performance metric for safe and stable operation. In this section, we evaluate the durability of estimation algorithms against onboard failures by simulating possible malfunctions.

8.4.1 Node Fault Case

As illustrated in Figure 8.29, we consider the case where one of the member spacecraft faces some operational problems. Here we define this type of malfunctioning as the node fault. We assume that the malfunctioning spacecraft under the node fault cannot communicate with the other spacecraft or provide the inter-spacecraft sensing.

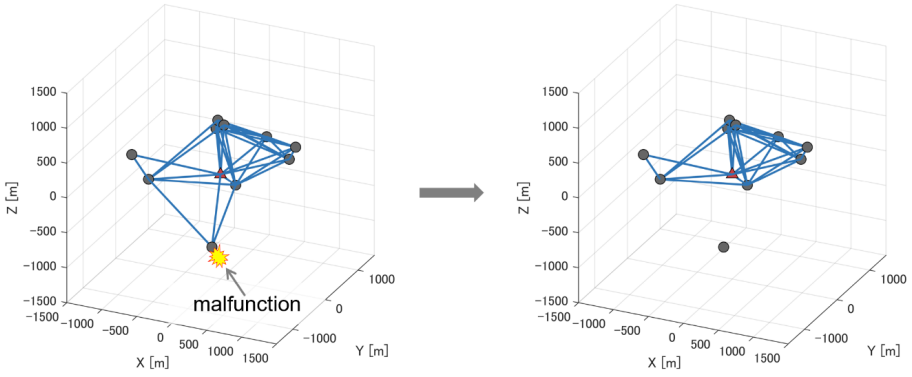


Figure 8.29: Simulation of a failure mode: the malfunction of a single spacecraft (node fault).

The node fault is mathematically expressed by modifying the adjacency matrix of the communication and sensing networks. We assume that the j_f-th spacecraft ($j_f \in \mathcal{V}$) get malfunctioning at the k_f-th time step. Then, the adjacency matrix is defined as:

$$\mathcal{A} = \begin{cases} \mathcal{A}\left(\mathcal{G}_k^x\right) & (k_0 \leq k < k_f) \\ \mathcal{A}\left(\breve{\mathcal{G}}_k^x\right) & (k_f \leq k \leq k_e) \end{cases} \tag{8.4}$$

where $x \in \{c, r, b\}$, $\mathcal{G}_k^x = (\mathcal{V}, \mathcal{E}_k^x)$, and $\breve{\mathcal{G}}_k^x = (\mathcal{V}, \breve{\mathcal{E}}_k^x)$. The edge set $\breve{\mathcal{E}}_k^x$ is generated by extracting all the edges connected to the j_f-th spacecraft from $\mathcal{E}_k^x$. The time steps k_0 and k_e represent the initial and the end time of the simulation, respectively. Table 8.4 summarizes the parameters used in the simulation for the node fault case.

Table 8.4: Parameters for simulating the node fault.

Parameter	Value	Units
Initial estimation error (identical to Case-A in Table 8.3)		
position error ($\sigma_{p_{init}}$)	100.0	[m]
velocity error ($\sigma_{v_{init}}$)	1.0	[m/s]
Timing of the failure mode (t_f: corresponding to the k_f-th time step)	50	[s]
Malfunctioning spacecraft (the j_f-th spacecraft)	1	[-]
Other parameters	Table 4.1	[-]

Figure 8.30 shows the convergence rate with a single node fault. Figure 8.31 shows whether the estimation results of all the filters are converged or not for the specific network threshold and the swarm configuration. When comparing the convergence rate with the result of the non-fault case (Figure 8.23), the estimation stability is slightly worse than the non-fault case. When a spacecraft is completely separated from the other spacecraft, its state is not modified in the measurement update step. Therefore, it is considered that the degradation of the convergence rate is not so high when a node fault occurs.

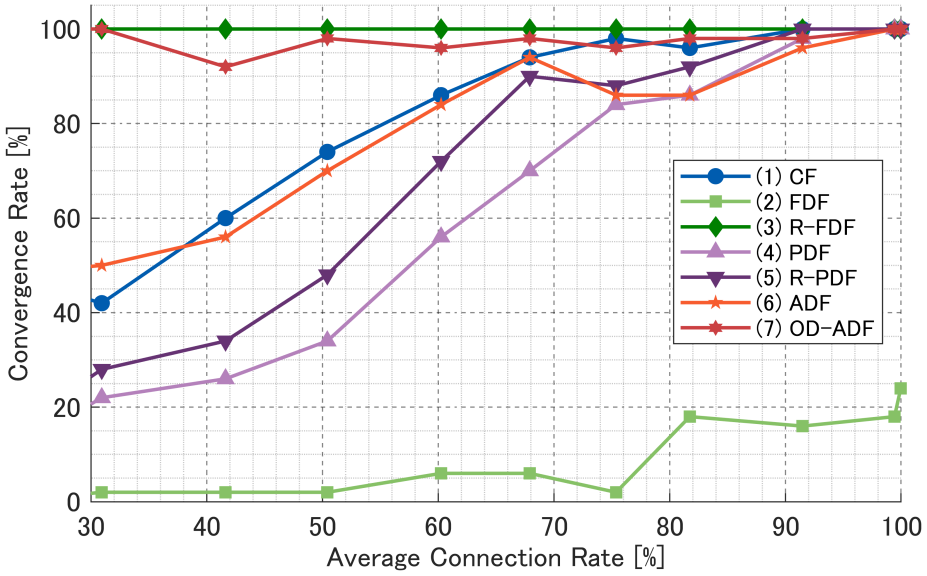


Figure 8.30: Convergence rate (the node fault case).

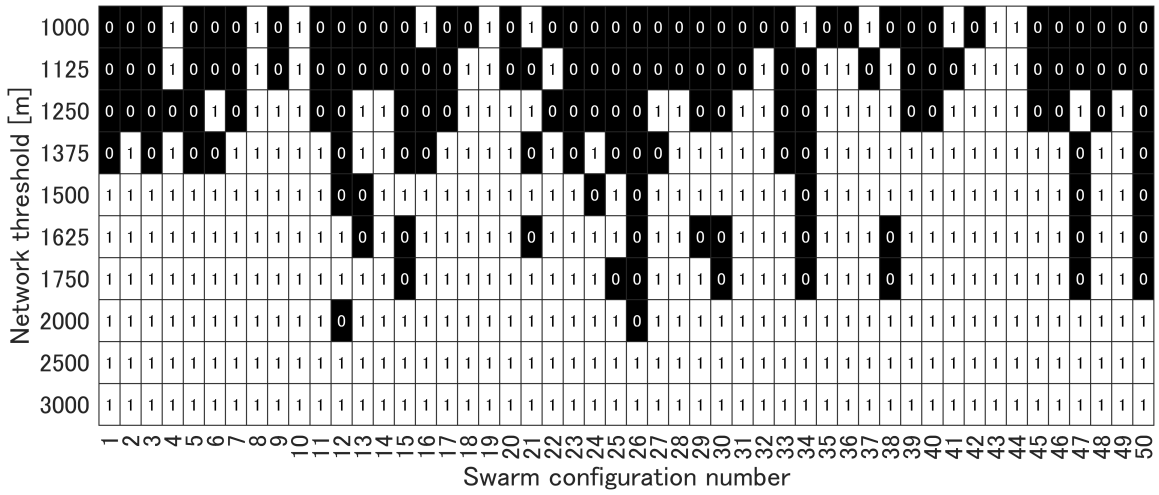
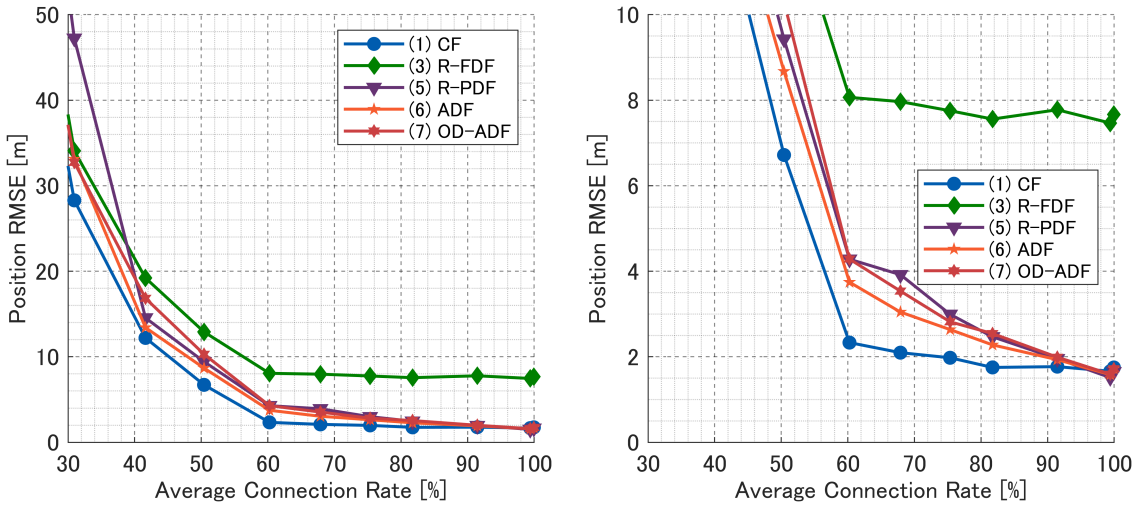


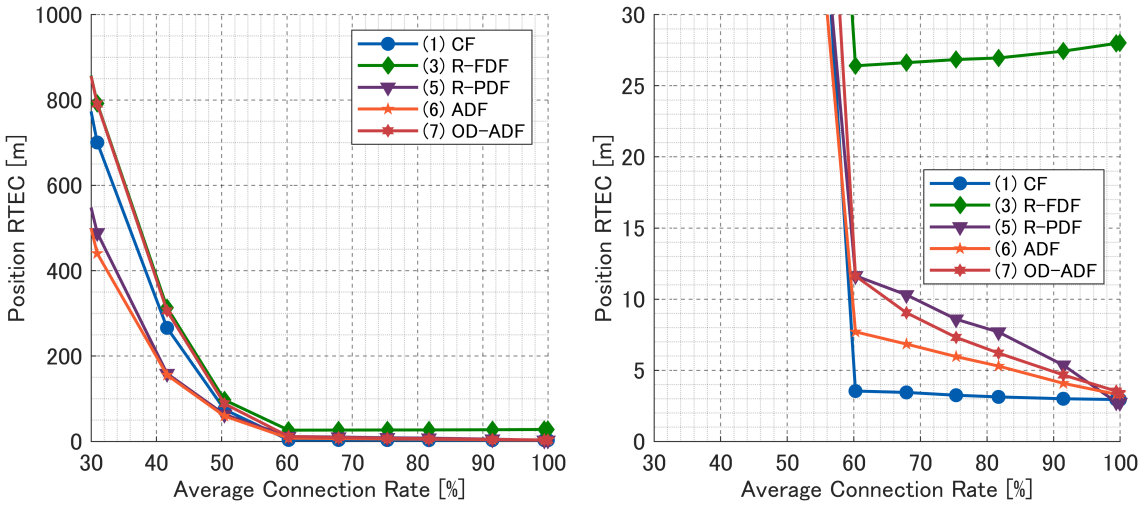
Figure 8.31: Gross convergence flags (the node fault case).

Figures 8.32 and 8.33 respectively show the average RMSE and the average RTEC of the estimated positions. The figures plot the average values removing the estimation result of the malfunctioning spacecraft. Even with the node fault, all the filters can continue to provide the estimated positions with almost the same accuracy as the non-fault cases of Figures 8.5 and 8.6.



(a) Overall view **(b)** Extended view

Figure 8.32: Root mean square errors of the estimated positions (the node fault case: the estimation result of the malfunctioning spacecraft $i = 1$ is excluded).



(a) Overall view **(b)** Extended view

Figure 8.33: Root of the trace of the position error covariance matrix (the node fault case: the estimation result of the malfunctioning spacecraft $i = 1$ is excluded).

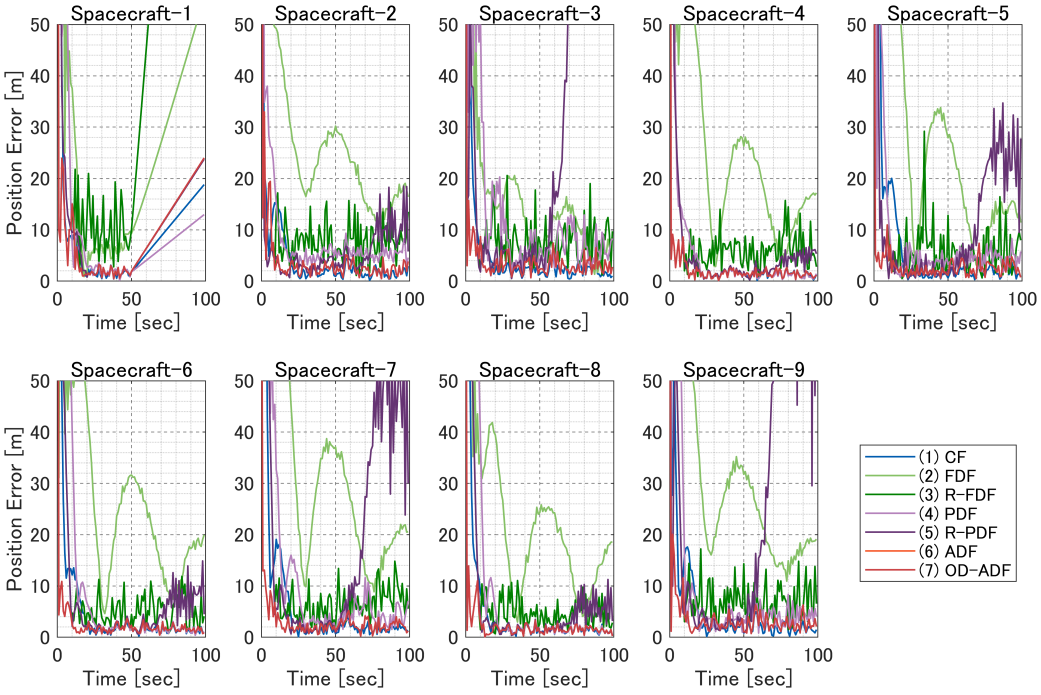
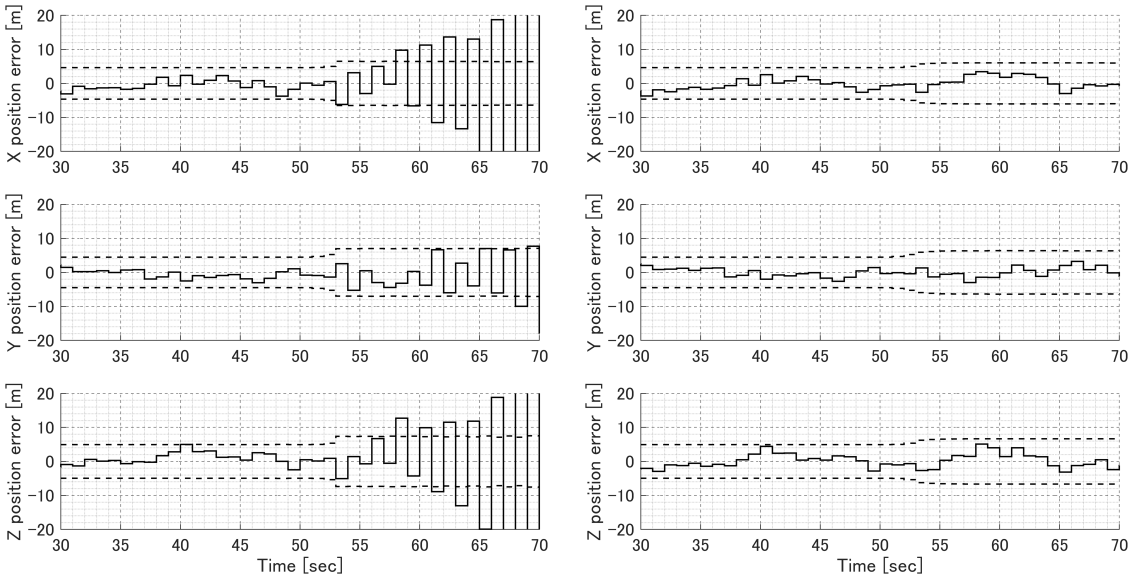


Figure 8.34: Position estimation error for every spacecraft for the 13th swarm configuration with the 1500m network threshold (the node fault case).

Figure 8.34 shows the RMSE of the estimated positions for the 13th swarm configuration with the 1500m network threshold. Since the 1st spacecraft get malfunctioning, the measurement update of its state stops at 50sec and the estimation error increases with

time. The estimation results imply that the R-PDF falls into divergence since its RMSE of the estimation positions rapidly increase from 50sec. Figure 8.35 shows the RMSE of the estimation position of the 9th spacecraft, which the R-PDF and the OD-ADF compute.



(a) R-PDF **(b)** OD-ADF

Figure 8.35: Position estimation error of the 9th spacecraft calculated by (a) R-PDF and (b) OD-ADF for the 13th swarm configuration with the 1500m network threshold (solid line: the RMSE of the estimated positions, dashed line: 3-sigma error bound).

8.4.2 Edge Fault Case

In a practical situation, a spacecraft would experience the malfunction of the onboard sensors or communication devices. Alternatively, inter-spacecraft communication and sensing get unavailable due to the temporary attenuation of sensor sensitivity. We define this type of malfunctioning as the edge fault, as illustrated in Figure 8.36.

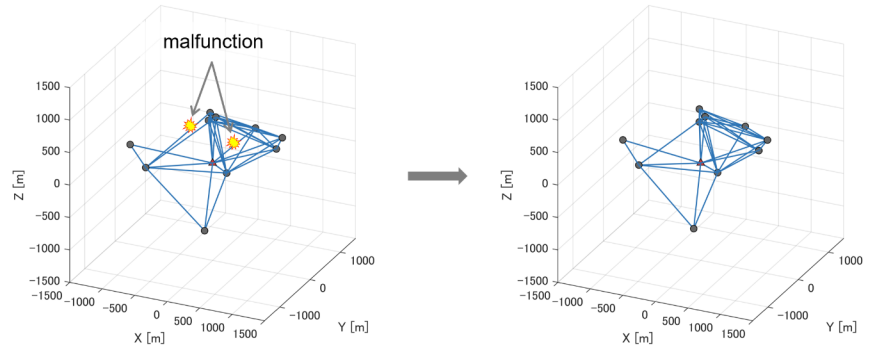


Figure 8.36: Simulation of a failure mode: the malfunction of inter-spacecraft links (edge fault).

The edge fault is mathematically expressed by modifying the adjacency matrices as with the node fault. We assume that N_f links between spacecraft, which are randomly selected from the edge set, get malfunctioning at the k_f-th time step.

$$\mathcal{A} = \begin{cases} \mathcal{A}\left(\mathcal{G}_k^x\right) & (k_0 \leq k < k_f) \\ \mathcal{A}\left(\breve{\mathcal{G}}_k^x\right) & (k_f \leq k \leq k_e) \end{cases} \tag{8.5}$$

where $x \in \{c, r, b\}$, $\mathcal{G}_k^x = (\mathcal{V}, \mathcal{E}_k^x)$, and $\breve{\mathcal{G}}_k^x = (\mathcal{V}, \breve{\mathcal{E}}_k^x)$. The edge set $\breve{\mathcal{E}}_k^x$ is generated by extracting randomly selected N_f edges from $\mathcal{E}_k^x$. Table 8.5 summarizes the parameters used in the simulation for the edge fault case.

Table 8.5: Parameters for simulating the edge fault.

Parameter	Value	Units
Initial estimation error (identical to Case-A in Table 8.3)		
position error ($\sigma_{p_{init}}$)	100.0	[m]
velocity error ($\sigma_{v_{init}}$)	1.0	[m/s]
Timing of the failure mode (t_f: corresponding to the k_f-th time step)	50	[s]
Number of malfunctioning edges (N_f)	3	[-]
Other parameters	Table 4.1	[-]

Figure 8.37 shows the convergence rate with three edge faults, and Figure 8.38 shows whether the estimation results of all the filters are converged or not for the specific network

threshold and the swarm configuration. When comparing the convergence rate of the edge fault case with that of the node fault case, the edge fault case results in lower convergence rate. In the node fault case, the state of the malfunctioning spacecraft is not estimated. Therefore, the estimation algorithms are likely to keep sufficient measurement information for the estimated states even if several measurements linked with the malfunctioning spacecraft get unavailable. On the other hand, in the edge fault case, the inter-spacecraft measurements are purely lost while keeping the number of estimated spacecraft. Hence, the edge fault can be more severe fault modes for the estimation algorithms, especially for the FDF, the PDF, and the R-PDF.

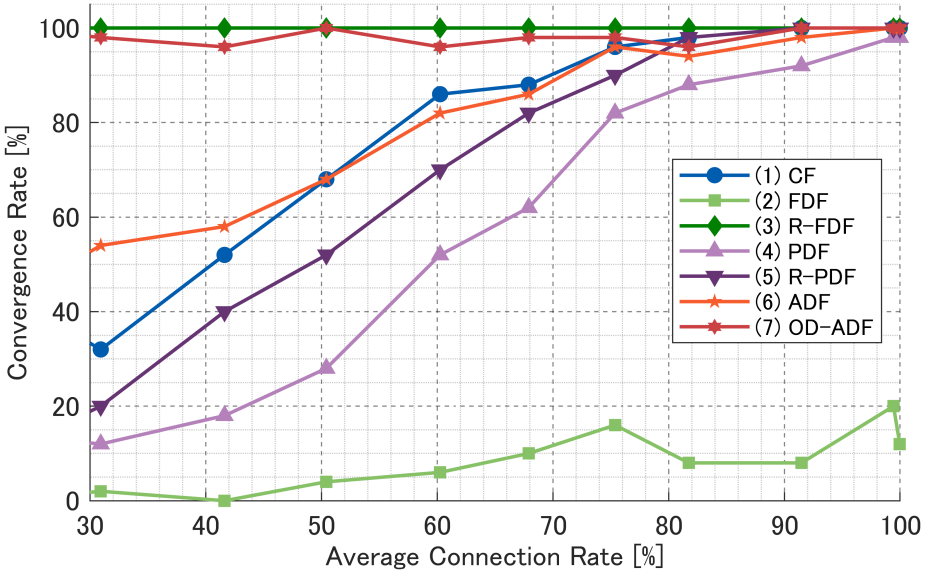


Figure 8.37: Convergence rate (the edge fault case).

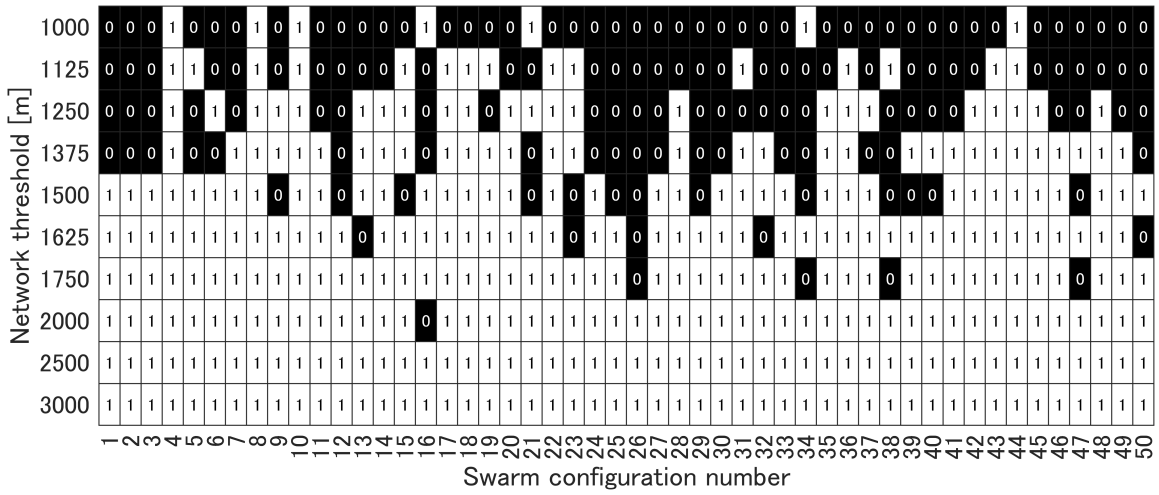
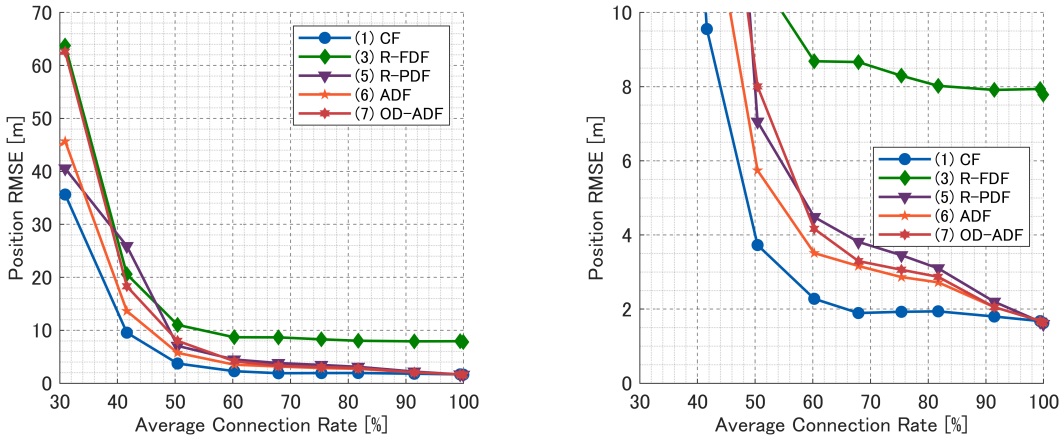


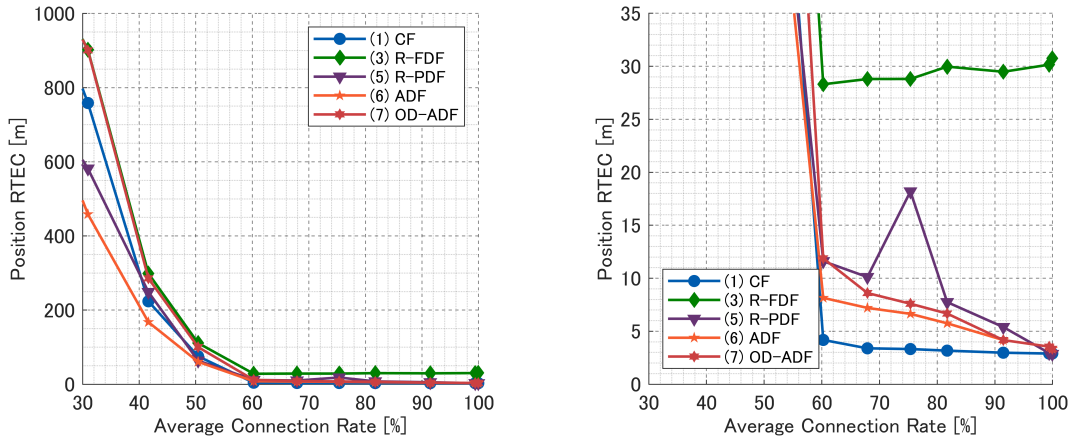
Figure 8.38: Gross convergence flags (the edge fault case).

Figures 8.39 and 8.40 show the average RMSE and the average RTEC of the estimated positions. As with the node fault case, all the filters can keep almost the same estimation accuracy as the non-fault cases in the converged scenarios.



(a) Overall **(b)** Extended view

Figure 8.39: Root mean square errors of the estimated positions (the edge fault case).



(a) Overall **(b)** Extended view

Figure 8.40: Root of the trace of the position error covariance matrix (the edge fault case).

8.5 Sensitivity to Network Structures

For the simulations in Sections 8.2 to 8.4, it is assumed that the three networks are identical: $\mathcal{G}^c = \mathcal{G}^r = \mathcal{G}^b$. Practically, this is realized by setting the distance thresholds as $\Delta R^c = \Delta R^r = \Delta R^b$. However, the network structures can be easily different due to some implementation constraints like sensors' field of view and the limited number of

simultaneous communication links.

In this section, we evaluate the performances of the estimation algorithms under conditions with different types of network structures. Figures 8.41, 8.42, and 8.43 show three variations of network structures to be evaluated, in which the case of the 4th swarm configuration is illustrated. The simulation condition follows Table 8.6.

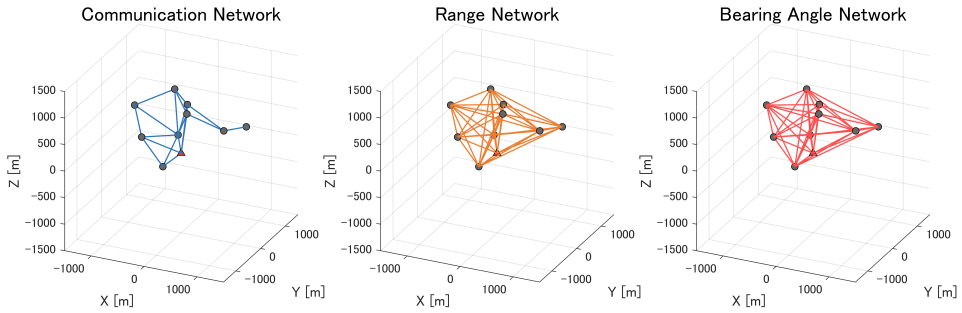


Figure 8.41: Network structures (A) ($m = 4$, $\Delta R^c = 1000[m]$, $\Delta R^r = 2000[m]$, $\Delta R^b = 2000[m]$).

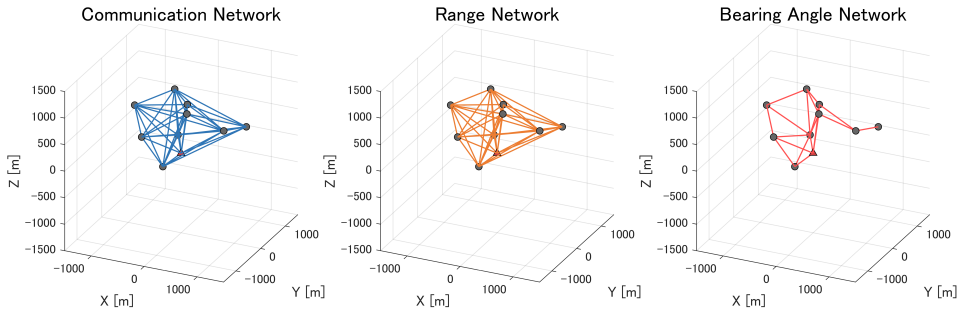


Figure 8.42: Network structures (B) ($m = 4$, $\Delta R^c = 2000[m]$, $\Delta R^r = 2000[m]$, $\Delta R^b = 1000[m]$).

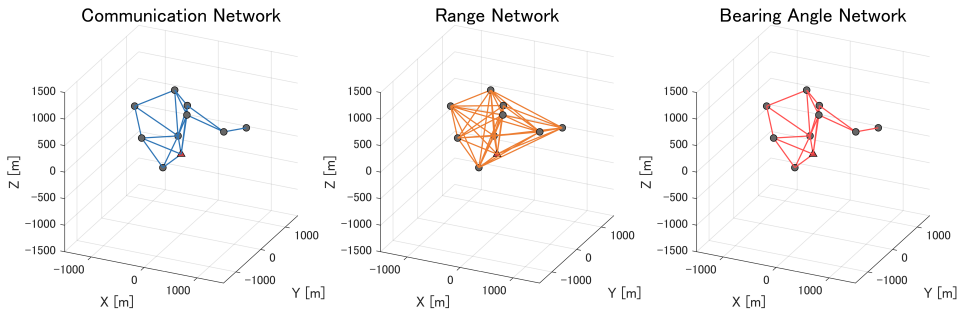


Figure 8.43: Network structures (C) ($m = 4$, $\Delta R^c = 1000[m]$, $\Delta R^r = 2000[m]$, $\Delta R^b = 1000[m]$).

Table 8.6: Parameters for the evaluation of the sensitivity to network structures.

Parameter	Value	Units
Initial estimation error (identical to Case-A in Table 8.3)		
position error ($\sigma_{p_{init}}$)	100.0	[m]
velocity error ($\sigma_{v_{init}}$)	1.0	[m/s]
Other parameters	Table 4.1	[-]

Table 8.7 summarizes the convergence rate of each filter for the three types of network structures. The values in the table show the result assessed with respect to the 50 swarm configurations of Figure 8.1. The first thing to notice is that the CF provides almost 100% estimation stability. In the simulation results of Figure 8.2 with the uniform network structures, the CF achieves a 100% convergence rate when the network threshold is above 2000m. Therefore, the result of Table 8.7 implies that the CF still be able to perform a stable estimation with the sparse bearing angle network (corresponding to $\Delta R^b = 1000[m]$).

In the partially decentralized filters, especially for the PDF, the R-FDF, and the ADF, the number of estimated states is decided by the communication network. Therefore, their convergence rates of the case (A) are higher than the other two cases. However, the estimation stability of these filters is remarkably lower than the CF and the R-FDF.

Finally, the OD-ADF achieves high stability with an almost 100% convergence rate as with the CF and the R-FDF. Although the non-identical network structures degrade the estimation stability of the ADF, the OD-ADF is able to provide highly stable estimation by evaluating the observability of the non-identical sensing networks.

Table 8.7: Convergence rate for the different network structures

	Thresholds	CF	FDF	R-FDF	PDF	R-PDF	ADF	OD-ADF
(A)	$\Delta R^c = 1000[m]$ $\Delta R^r = 2000[m]$ $\Delta R^b = 2000[m]$	100%	12%	100%	18%	40%	76%	100%
(B)	$\Delta R^c = 2000[m]$ $\Delta R^r = 2000[m]$ $\Delta R^b = 1000[m]$	100%	2%	100%	2%	38%	58%	96%
(C)	$\Delta R^c = 1000[m]$ $\Delta R^r = 2000[m]$ $\Delta R^b = 1000[m]$	98%	2%	100%	4%	34%	62%	100%

8.6 Summary of Performance Evaluation

This section summarizes the main results and discussions of the numerical simulations. Based on the review of the related work in Table 1.2, we evaluate the performance of the proposed algorithms by comparing them with the other existing methods while considering the influential factors: swarm configurations, initial estimation errors, fault tolerance, and network structures.

Firstly, we evaluated the performances of whether each algorithm successfully handles a variety of swarm configurations in Section 8.2. It is proved that the proposed methods, the ADF and the OD-ADF, provide the high estimation stability than the other partially decentralized filters while keeping reasonably high estimation accuracy. Secondly, we investigate the influence of initial estimation errors in Section 8.3. The simulation results show that the initial estimation errors significantly impact the estimation stability. Almost all the filters result in a low convergence rate when the initial estimation errors are substantially large. However, the OD-ADF and the R-FDF keep nearly 100% convergence rate regardless of initial estimation errors. Section 8.4 investigates whether each estimation

Table 8.8: Correspondence between the evaluation conditions and the sections

Evaluation condition	Section
Swarm configurations	Section 8.2
Initial estimation errors	Section 8.3
Fault tolerance	Section 8.4
Network structures	Section 8.5
Comparison with other methods	Table 8.1

algorithm can handle malfunctioning scenarios, including the node fault and the edge fault. The simulation results show that the OD-ADF can keep an almost 100% convergence rate even with malfunctioning and sparse networks. Finally, we evaluate the performance sensitivity of the estimation algorithms to the network structures. Although the estimation stability of several filters (the PDF, the R-FDF, and the ADF) are governed by the most sparse network, the OD-ADF keeps almost the 100% convergence rate. As with the uniform network structure cases, the OD-ADF can evaluate the observability of the non-identical network structures and judge whether the integration of the iPDF estimation is safe or not.

Chapter 9

Conclusions and Future Work

9.1 Conclusions

In this dissertation, we proposed the new estimation algorithms, the adaptive decentralized filter (ADF) and the observability-driven adaptive decentralized filter (OD-ADF), for the relative navigation of spacecraft swarms. The ADF achieves the estimation with high accuracy and high stability by adjusting the decentralization levels of estimation algorithms. Additionally, the OD-ADF introduces the observability evaluation for the adjustment process of decentralization in the ADF and realizes a more stable estimation than the ADF. The performances of these algorithms are evaluated for a variety of conditions, including multiple swarm configurations, initial estimation errors, spacecraft malfunctioning, and inter-spacecraft network structures. The numerical simulations show that the proposed algorithms provide well-balanced performances in terms of estimation accuracy, estimation stability, and communication load.

In Chapter 1, we presented the background of spacecraft swarms and reviewed various relative navigation algorithms. We clarify the problems to be solved in this dissertation by analyzing the related work.

In Chapter 2, we presented the problem formulation for the relative navigation of spacecraft swarms. We gave the definitions of the estimated states, the dynamics models, and the inter-spacecraft measurement models. We also stated the mathematical treatment of communication and sensing networks.

In Chapter 3, we summarized the basic ideas and formulations of the recursive Bayesian filtering, which is the core estimation approach in this dissertation. We also mentioned the canonical parameterization of estimation algorithms, represented by the information vector

and matrix. It is an appropriate representation of probability distributions for decentralized estimation algorithms, specifically for numerical implementation.

In Chapter 4, we presented the overview of the spacecraft swarm simulator and its relevant parameters and performance metrics. The simulator is widely used to evaluate estimation algorithms introduced in this dissertation.

In Chapter 5, we presented the representative estimation algorithms compared with the proposed algorithms. The algorithms introduced in this chapter include the centralized filter (CF), the fully decentralized filter (FDF, R-FDF), and the partially decentralized filter (PDF). We also proposed a new algorithm called the robust partially decentralized filter (R-PDF) by modifying the PDF to improve the estimation stability. We executed the performance analysis for three algorithms (the CF, the R-FDF, and the R-PDF) and clarified that these algorithms have challenges to balance high accuracy and stability when the communication and sensing networks are sparse.

In Chapter 6, we presented the algorithmic details of the proposed ADF. The ADF is roughly divided into three sections, the internal fully decentralized filter (iFDF), the internal partially decentralized filter (iPDF), and the fusion process of the estimation results of the iFDF and the iPDF.

In Chapter 7, we presented the algorithmic details of the proposed OD-ADF. We stated the key elements to enable the observability-driven decentralization, including the definition of the observability for sensing networks, the probabilistic on-board evaluation of the observability, and the decision process of the observability threshold.

In Chapter 8, we presented the performance verification of the proposed algorithms by the numerical simulations. For the comprehensive analysis, we evaluated the algorithms under various conditions: multiple swarm configurations, initial estimation errors, spacecraft malfunctioning, and inter-spacecraft network structures. The numerical simulations show that the proposed algorithms provide well-balanced performances in terms of estimation accuracy, estimation stability, and communication load.

9.2 Future Work

The proposed algorithms achieve high performances; however, they have the potentials to be improved and need further considerations for the practical implementation on real spacecraft swarms. The following problems are possible research targets, which would be addressed in the future.

- **Theoretical observability threshold**: We decide the observability threshold of the OD-ADF by the heuristic approach as stated in Section 7.4. However, the heuristic approach should be reapplied when the OD-ADF is used for swarm systems with different implementations, such as sensor devices or the number of spacecraft. Therefore, it is highly demanded to develop a theoretical approach to decide the observability threshold.
- **Active sensing and the information-based decentralization**: We can consider the following aspects to enhance the effectiveness of the proposed navigation algorithms. The preliminary discussion on this topic is summarized in Appendix B.
  - How to select the target spacecraft which each spacecraft communicates with or senses to, in decentralized estimation architectures
  - How to decide the size of the aggregated state vector considered in the partially/adaptive decentralized filter
- **Communication delay and time synchronization**: Although we consider the data availability limited by the communication network, we assume that the spacecraft experience no communication delay and their onboard timers are perfectly synchronized. However, in practical implementations, the above assumptions are not realistic, and navigation algorithms need to handle the communication delay and consider the quality of time synchronization.
- **Hardware demonstration**: The proposed algorithms and the other existing algorithms are evaluated in the full software simulator in this dissertation. Hardware-based performance evaluations would be beneficial to find out significant factors for the navigation algorithms, for instance, constraints on computational resources, a real anomaly for fault tolerance evaluation, etc.

Acknowledgement

I would like to express my gratitude to my supervisor, Shinichi Nakasuka, for his insightful advice and consistent support. I would also like to acknowledge my thesis committee members, Ryu Funase, Yuichi Tsuda, Yasuhiro Kawakatsu, and Takehisa Yairi, for their very helpful and constructive advice. I am also grateful to Satoshi Ikari for his dedicated support, and Toshiro Shimizu, Takashi Shimizu, and Kakuto Hori for their helpful advice on communication technology. Finally, I would like to thank my parents and sister for their continuous support.

Appendix A

Probability

A.1 Derivation of Gaussian Filters

A.1.1 Derivation of the Measurement Update in the Gaussian Filter

We summarize the derivation of the measurement update in the Gaussian filter. The exponential term of $p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})$ in Eq. (3.4), which is defined as X, can be represented as:

$$X = (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T P_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + (\boldsymbol{y}_k - \boldsymbol{h}(\boldsymbol{x}))^T R^{-1}(\boldsymbol{y}_k - \boldsymbol{h}(\boldsymbol{x})) \tag{A.1}$$

If we assume that $\boldsymbol{h}(\boldsymbol{x})$ follows the Gaussian distribution with the mean $\hat{\boldsymbol{y}}_{k|k-1}$ and the covariance $P_{k|k-1}^{yy}$, X can be reformulated as:

$$\begin{aligned}
X &= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T P_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\
&\quad +(\boldsymbol{y}_k - P_{k|k-1}^{-1}P_{k|k-1}^{xy}\boldsymbol{x}_k)^T \left(P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1} \\
&\qquad\qquad \times(\boldsymbol{y}_k - P_{k|k-1}^{-1}P_{k|k-1}^{xy}\boldsymbol{x}_k) \\
&= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T P_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\
&\quad + \left\{\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})\right\}^T \\
&\qquad \times \left(P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1} \\
&\qquad\quad \times \left\{\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})\right\}
\end{aligned}$$

and then,

$$\begin{aligned}
X &= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \Big\{ P^{-1}_{k|k-1} + P^{-1}_{k|k-1} P^{xy}_{k|k-1} \\
&\quad \times \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} \Big\} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\
&+ (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&- (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \\
&\quad \times \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&- (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} \\
&\quad \times (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})
\end{aligned}$$

Due to the following matrix relationship (Woodbury matrix identity) [53]:

$$(\boldsymbol{A} + \boldsymbol{BDC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D}^{-1} + \boldsymbol{CA}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{CA}^{-1} \tag{A.2}$$

Then, X is reformulated as:

$$\begin{aligned}
X &= (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \left(P_{k|k-1} - P^{xy}_{k|k-1} (P^{yy}_{k|k-1})^{-1} (P^{xy}_{k|k-1})^T \right)^{-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) \\
&+ (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&- (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&- (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T \left(P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1} \right)^{-1} \\
&\times (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})
\end{aligned} \tag{A.3}$$

By defining $P_{k|k} \triangleq P_{k|k-1} - P^{xy}_{k|k-1}(P^{yy}_{k|k-1})^{-1}(P^{xy}_{k|k-1})^T$, $\Sigma_{k|k} \triangleq P^{yy}_{k|k-1} - (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P^{xy}_{k|k-1}$, X can be expressed as:

$$\begin{aligned}
X &= \left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + P_{k|k} P^{-1}_{k|k-1} P^{xy}_{k|k-1} \Sigma^{-1}_{k|k} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \right]^T P^{-1}_{k|k} \\
&\times \left[(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + P_{k|k} P^{-1}_{k|k-1} P^{xy}_{k|k-1} \Sigma^{-1}_{k|k} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \right] \\
&+ (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T \left(\Sigma^{-1}_{k|k} - \Sigma^{-1}_{k|k} (P^{xy}_{k|k-1})^T P^{-1}_{k|k-1} P_{k|k} P^{-1}_{k|k-1} P^{xy}_{k|k-1} \Sigma^{-1}_{k|k} \right) (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})
\end{aligned}$$

Therefore, the mean of the posterior probability is:

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + P_{k|k}P_{k|k-1}^{-1}P_{k|k-1}^{xy}\Sigma_{k|k}^{-1}(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{x}}_{k|k-1} + \left(P_{k|k-1} - P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T\right)P_{k|k-1}^{-1}P_{k|k-1}^{xy} \\
&\quad \times \left(P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1}(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{x}}_{k|k-1} + \left(\boldsymbol{I} - P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}\right)P_{k|k-1}^{xy} \\
&\quad \times \left(\boldsymbol{I} - (P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1}(P_{k|k-1}^{yy})^{-1}(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{x}}_{k|k-1} + P_{k|k-1}^{xy}\left(\boldsymbol{I} - (P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right) \\
&\quad \times \left(\boldsymbol{I} - (P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1}(P_{k|k-1}^{yy})^{-1}(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{x}}_{k|k-1} + P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1}(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \qquad \text{(A.4)}
\end{aligned}$$

Finally, the measurement update step of the Gaussian filter can be summarized as:

$$K_k \triangleq P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1} \qquad \text{(A.5)}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + K_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \qquad \text{(A.6)}$$

$$P_{k|k} = P_{k|k-1} - K_k P_{k|k-1}^{yy} K_k^T \qquad \text{(A.7)}$$

where K_k is the Kalman gain at the k-th time step.

A.1.2 Derivation of the Measurement Update in the Gaussian Information Filter

We summarize the derivation of the measurement update in the Gaussian information filter. Firstly, we summarize the derivation of the posterior information matrix. By Eq. (3.18) and Woodbury matrix identity formula of Eq. (A.2):

$$\begin{aligned}
\boldsymbol{Z}_k &= P_k^{-1} = \left(P_{k|k-1} - P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1}(P_{k|k-1}^{xy})^T\right)^{-1} \\
&= P_{k|k-1}^{-1} + P_{k|k-1}^{-1}P_{k|k-1}^{xy}\left(P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1}(P_{k|k-1}^{xy})^T P_{k|k-1}^{-1} \\
&= \boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T\left(P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1}P_{k|k-1}^{xy}\right)^{-1}\mathcal{H}_k \qquad \text{(A.8)}
\end{aligned}$$

where a pseudo measurement matrix is difined as $\mathcal{H}_k^T \triangleq P_{k|k-1}^{-1} P_{k|k-1}^{xy}$. The term $P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1} P_{k|k-1}^{xy}$ in Eq. (A.8) is calculated as:

$$P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1} P_{k|k-1}^{xy} = \int_{\mathbb{R}^{n_x}} A\mathcal{N}(\boldsymbol{x}_k|\hat{\boldsymbol{x}}_{k|k-1}, P_{k|k-1}) d\boldsymbol{x}_k + R \quad \text{(A.9)}$$

$$A = (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T - (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T \times \frac{1}{(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})^T} (\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})^T = \mathbf{0} \quad \text{(A.10)}$$

$$P_{k|k-1}^{yy} - (P_{k|k-1}^{xy})^T P_{k|k-1}^{-1} P_{k|k-1}^{xy} = R \quad \text{(A.11)}$$

Therefore, the measurement update of the information matrix can be expressed in the following simple form.

$$\boldsymbol{Z}_k = \boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k \quad \text{(A.12)}$$

Secondly, we summarize the derivation of the posterior information vector. Eq. (3.22) is reformulated as:

$$\begin{aligned}
\hat{\boldsymbol{z}}_k &= \boldsymbol{Z}_k \hat{\boldsymbol{x}}_k \\
&= \{\boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k\}\{\hat{\boldsymbol{x}}_{k|k-1} + K_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})\} \\
&= \boldsymbol{Z}_{k|k-1}\hat{\boldsymbol{x}}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k \hat{\boldsymbol{x}}_{k|k-1} + (\boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k) K_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{z}}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k \hat{\boldsymbol{x}}_{k|k-1} + (\boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k) K_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1})
\end{aligned}$$

where the third term, which is defined as Y here, can be formulated as:

$$\begin{aligned}
Y &= (\boldsymbol{Z}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k) K_k(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= (P_{k|k-1}^{-1} + \mathcal{H}_k^T R^{-1} P_{k|k-1}^{xy} P_{k|k-1}^{-1}) P_{k|k-1}^{xy} (P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= P_{k|k-1}^{-1} P_{k|k-1}^{xy} (P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&\quad + \mathcal{H}_k^T R^{-1} P_{k|k-1}^{xy} P_{k|k-1}^{-1} P_{k|k-1}^{xy} (P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \mathcal{H}_k^T (P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) + \mathcal{H}_k^T R^{-1} (P_{k|k-1}^{yy} - R)(P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \mathcal{H}_k^T \left(I + R^{-1} P_{k|k-1}^{yy} - R^{-1} R \right) (P_{k|k-1}^{yy})^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \mathcal{H}_k^T R^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \quad \text{(A.13)}
\end{aligned}$$

Therefore, the measurement update of the information vector is formulated as:

$$\begin{aligned}
\hat{\boldsymbol{z}}_k &= \hat{\boldsymbol{z}}_{k|k-1} + \mathcal{H}_k^T R^{-1} \mathcal{H}_k \hat{\boldsymbol{x}}_{k|k-1} + \mathcal{H}^T R^{-1} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}) \\
&= \hat{\boldsymbol{z}}_{k|k-1} + \mathcal{H}_k^T R^{-1} \left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} + \mathcal{H}_k \hat{\boldsymbol{x}}_{k|k-1} \right) \quad \text{(A.14)}
\end{aligned}$$

A.2 Linear Approximation for Probability Integrals

In this section, we summarize how to calculate the following integral of probability distributions with reference to [54].

$$\int_{\mathbb{R}^{n_x}} \boldsymbol{F}(\boldsymbol{x}) \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \exp\left(\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^T P^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right) d\boldsymbol{x} \tag{A.15}$$

$$\int_{\mathbb{R}^{n_x}} \left(\boldsymbol{F}_1(\boldsymbol{x}) - I[\tilde{\boldsymbol{F}}_1]\right)\left(\boldsymbol{F}_2(\boldsymbol{x}) - I[\tilde{\boldsymbol{F}}_2]\right)^T \times \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \exp\left(\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^T P^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right) d\boldsymbol{x} \tag{A.16}$$

where $\boldsymbol{F}(\boldsymbol{x})$ is a possibly nonlinear function. Let $P = S^T S$, and substitute $\boldsymbol{x} = \hat{\boldsymbol{x}} + S^T\boldsymbol{\xi}$, and Eq. (A.15) is expressed as:

$$\begin{aligned}
&\int_{\mathbb{R}^{n_x}} \boldsymbol{F}(\boldsymbol{x}) \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \exp\left(\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^T P^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right) d\boldsymbol{x} \\
&= \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \int_{\mathbb{R}^{n_x}} \boldsymbol{F}(\hat{\boldsymbol{x}} + S^T\boldsymbol{\xi}) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T S(S^T S)^{-1} S^T \boldsymbol{\xi}\right) d\boldsymbol{\xi} \cdot S^T \\
&= \frac{1}{(2\pi)^{n_x/2}} \int_{\mathbb{R}^{n_x}} \boldsymbol{F}(\hat{\boldsymbol{x}} + S^T\boldsymbol{\xi}) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right) d\boldsymbol{\xi} \\
&= \frac{1}{(2\pi)^{n_x/2}} \int_{\mathbb{R}^{n_x}} \tilde{\boldsymbol{F}}(\boldsymbol{\xi}) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right) d\boldsymbol{\xi} \\
&= \int_{\mathbb{R}^{n_x}} \tilde{\boldsymbol{F}}(\boldsymbol{\xi}) \mathcal{N}(\boldsymbol{\xi}|\boldsymbol{0}, \boldsymbol{I}) d\boldsymbol{\xi} \triangleq I[\tilde{\boldsymbol{F}}]
\end{aligned} \tag{A.17}$$

and Eq. (A.16) is expressed as:

$$\begin{aligned}
&\int_{\mathbb{R}^{n_x}} \left(\boldsymbol{F}_1(\boldsymbol{x}) - I[\tilde{\boldsymbol{F}}_1]\right)\left(\boldsymbol{F}_2(\boldsymbol{x}) - I[\tilde{\boldsymbol{F}}_2]\right)^T \\
&\qquad \times \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \exp\left(\frac{1}{2}(\boldsymbol{x}-\hat{\boldsymbol{x}})^T P^{-1}(\boldsymbol{x}-\hat{\boldsymbol{x}})\right) d\boldsymbol{x} \\
&= \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \int_{\mathbb{R}^{n_x}} \boldsymbol{F}_1(\hat{\boldsymbol{x}} + S^T\boldsymbol{\xi}) \boldsymbol{F}_2^T(\hat{\boldsymbol{x}} + S^T\boldsymbol{\xi}) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right) d\boldsymbol{\xi} \cdot S^T \\
&\quad - I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T - I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T + I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T \\
&= \frac{1}{(2\pi)^{n_x/2}|P|^{1/2}} \int_{\mathbb{R}^{n_x}} \tilde{\boldsymbol{F}}_1(\boldsymbol{\xi}) \tilde{\boldsymbol{F}}_2^T(\boldsymbol{\xi}) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right) d\boldsymbol{\xi} - I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T \\
&= \int_{\mathbb{R}^{n_x}} \tilde{\boldsymbol{F}}_1(\boldsymbol{\xi}) \tilde{\boldsymbol{F}}_2(\boldsymbol{\xi}) \mathcal{N}(\boldsymbol{\xi}|\boldsymbol{0}, \boldsymbol{I}) d\boldsymbol{\xi} - I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T \\
&= I[\tilde{\boldsymbol{F}}_1 \tilde{\boldsymbol{F}}_2^T] - I[\tilde{\boldsymbol{F}}_1] I[\tilde{\boldsymbol{F}}_2]^T \triangleq J[\tilde{\boldsymbol{F}}_1, \tilde{\boldsymbol{F}}_2^T]
\end{aligned} \tag{A.18}$$

Eqs. (A.17) and (A.18) cannot be calculated analytically when $\tilde{\boldsymbol{F}}(\boldsymbol{x})$ is a nonlinear function. In the linear approximation, $\boldsymbol{F}(\boldsymbol{x})$ is approximated by the first-order Taylor polynomial as:

$$\tilde{\boldsymbol{F}}(\boldsymbol{\xi}) \simeq \tilde{\boldsymbol{F}}(\boldsymbol{0}) + \tilde{\boldsymbol{F}}'(\boldsymbol{0})\boldsymbol{\xi}. \tag{A.19}$$

With this linear approximation, Eq. (A.17) is reformulated as:

$$\begin{aligned} I[\tilde{\boldsymbol{F}}] &= \frac{1}{(2\pi)^{n_x/2}} \int_{\mathbb{R}^{n_x}} \left(\tilde{\boldsymbol{F}}(0) + \tilde{\boldsymbol{F}}'(0)\boldsymbol{\xi} \right) \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi} \right) d\boldsymbol{\xi} \\ &= \tilde{\boldsymbol{F}}(0) = \boldsymbol{F}(\hat{\boldsymbol{x}}), \end{aligned} \tag{A.20}$$

and Eq. (A.18) is also reformulated as:

$$\begin{aligned} J[\tilde{\boldsymbol{F}}_1, \tilde{\boldsymbol{F}}_2] &= \frac{1}{(2\pi)^{n_x/2}} \int_{\mathbb{R}^{n_x}} \left(\tilde{\boldsymbol{F}}_1(0) + \tilde{\boldsymbol{F}}'_1(0)\boldsymbol{\xi} \right) \left(\tilde{\boldsymbol{F}}_2(0) + \tilde{\boldsymbol{F}}'_2(0)\boldsymbol{\xi} \right)^T \exp\left(\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi} \right) d\boldsymbol{\xi} \\ &\quad -\tilde{\boldsymbol{F}}_1(0)\tilde{\boldsymbol{F}}_2(0) \\ &= \tilde{\boldsymbol{F}}'_1(0)\tilde{\boldsymbol{F}}'^T_2(0) = \frac{\partial}{\partial\boldsymbol{\xi}}\tilde{\boldsymbol{F}}_1(\boldsymbol{\xi})\bigg|_{\boldsymbol{\xi}=\mathbf{0}} \frac{\partial}{\partial\boldsymbol{\xi}}\tilde{\boldsymbol{F}}_2^T(\boldsymbol{\xi})\bigg|_{\boldsymbol{\xi}=\mathbf{0}} \\ &= \frac{\partial}{\partial\boldsymbol{x}}\boldsymbol{F}_1(\boldsymbol{x})\bigg|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\xi}} \left(\frac{\partial}{\partial\boldsymbol{x}}\boldsymbol{F}_2(\boldsymbol{x})\bigg|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \frac{\partial\boldsymbol{x}}{\partial\boldsymbol{\xi}} \right)^T \\ &= \frac{\partial}{\partial\boldsymbol{x}}\boldsymbol{F}_1(\boldsymbol{x})\bigg|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} S^T S \left(\frac{\partial}{\partial\boldsymbol{x}}\boldsymbol{F}_2(\boldsymbol{x})\bigg|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} \right)^T \\ &= \boldsymbol{F}'_1(\hat{\boldsymbol{x}})P\boldsymbol{F}'^T_2(\hat{\boldsymbol{x}}). \end{aligned} \tag{A.21}$$

A.3 Spherical Cubature Integration

In this section, we summarize the overview of the spherical cubature integration with reference to [55]. For the mathematical simplicity, we consider the following integral of the normal Gaussian distribution as:

$$\int \tilde{\boldsymbol{F}}(\boldsymbol{\xi})\mathcal{N}\left(\boldsymbol{\xi}|\mathbf{0}, \boldsymbol{I}\right) d\boldsymbol{\xi}. \tag{A.22}$$

As an example, we consider the unscented transform [51]. The integral calculation is approximated as:

$$\int_{\mathbb{R}^n} \tilde{\boldsymbol{F}}(\boldsymbol{\xi})\mathcal{N}\left(\boldsymbol{\xi}|\mathbf{0}, \boldsymbol{I}\right) d\boldsymbol{\xi} \simeq W_0\tilde{\boldsymbol{F}}(\mathbf{0}) + W\sum_{i=1}^{2n} \tilde{\boldsymbol{F}}(c\boldsymbol{u}^{(i)}) \tag{A.23}$$

where W is a weight coefficient, c is a parameter yet to be determined, and points $\boldsymbol{u}^{(i)}$ belong to the symmetric set $[\mathbf{1}]$ defined as:

$$[\mathbf{1}] = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots \right\}. \tag{A.24}$$

In order that this rule is exact for all monomials of degree up to three, it requires that the rule is exact for all monomials for which $\sum_{i=1}^{n} d_i = 0, 2$. To find the unknow parameters, it suffices to consider monomials $\tilde{F}_j(\boldsymbol{\xi}) = 1$, and $\tilde{F}_j(\boldsymbol{\xi}) = \xi_j^2$ because of the fully symetric cubature rule.

$$\int_{\mathbb{R}^n} 1 \cdot \mathcal{N}\left(\boldsymbol{\xi}|\mathbf{0}, \boldsymbol{I}\right) d\boldsymbol{\xi} = 1 \tag{A.25}$$

$$\int_{\mathbb{R}^n} \xi_j^2 \cdot \mathcal{N}\left(\boldsymbol{\xi}|\mathbf{0}, \boldsymbol{I}\right) d\boldsymbol{\xi} = 1 \tag{A.26}$$

Considering the approximation of Eq.(A.23), the weight coefficients should sarisfy the following conditions:

$$W_0 \cdot 1 + W \cdot \sum_{i=1}^{2n} 1 = W_0 + 2nW = 1, \tag{A.27}$$

$$W_0 \cdot 0 + W \cdot \sum_{i=1}^{2n} \left[c u_j^{(i)} \right]^2 = 2c^2 W = 1. \tag{A.28}$$

By solving Eqs.(A.27) and (A.28), we obtain

$$W_0 = \frac{\kappa}{n+\kappa} \tag{A.29}$$

$$W = \frac{1}{2(n+\kappa)} \tag{A.30}$$

$$c = \sqrt{n+\kappa} \tag{A.31}$$

where κ is a free parameter. Finally, with the approximation based on the unscented transform, the integral is analytically computed as:

$$\int_{\mathbb{R}^{n_x}} \boldsymbol{F}(\boldsymbol{x}) \mathcal{N}\left(\boldsymbol{x}|\hat{\boldsymbol{x}}, P\right) d\boldsymbol{x} \simeq \frac{\kappa}{n+\kappa} \boldsymbol{F}(\hat{\boldsymbol{x}}) + \frac{1}{2(n+\kappa)} \sum_{i=1}^{2n} \boldsymbol{F}\left(\hat{\boldsymbol{x}} + \sqrt{P}\boldsymbol{\xi}^{(i)}\right). \tag{A.32}$$

A.4 Covariance Intersection

The Covariance Intersection algorithm is a data fusion algorithm that takes a convex combination of the means and covariances in the information space. [51] We consider two probabilities $\mathcal{N}(\boldsymbol{\mu}_a, \Sigma_{aa})$ and $\mathcal{N}(\boldsymbol{\mu}_b, \Sigma_{bb})$, which are assumed to follow the Gaussian distributions. The problem is to fuse these two distributions and to yield a new probability $\mathcal{N}(\boldsymbol{\mu}_c, \Sigma_{cc})$. The Covariance Intersection algorithm is formulated as

$$\Sigma_{cc}^{-1} = \omega\Sigma_{aa}^{-1} + (1-\omega)\Sigma_{bb}^{-1} \tag{A.33}$$

$$\Sigma_{cc}^{-1}\boldsymbol{\mu}_c = \omega\Sigma_{aa}^{-1}\boldsymbol{\mu}_a + (1-\omega)\Sigma_{bb}^{-1}\boldsymbol{\mu}_b \tag{A.34}$$

where $\omega \in [0, 1]$. The free parameter ω can be selected with different performance criteria such as minimizing the trace or the determinant of Σ_{cc}.

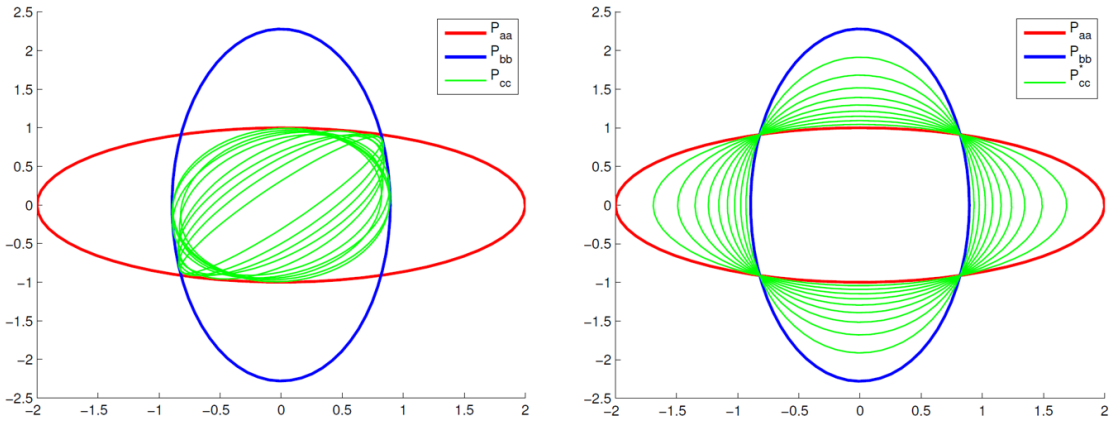


Figure A.1: Covariance ellipsoids Σ_{aa}, Σ_{bb}, and Σ_{cc} (left: the updated Σ_{cc} with different values of Σ_{ab}, right: the CI updated Σ_{cc}) [10]

The consistency of Eqs. (A.33) and (A.34) are proved by the following discussion, that is cited from [51]. The actual error in the estimate by the CI is represented as:

$$\boldsymbol{\mu}_c = \Sigma_{cc}\left\{\omega\Sigma_{aa}^{-1}\boldsymbol{\mu}_a + (1-\omega)\Sigma_{bb}^{-1}\boldsymbol{\mu}_b\right\} \tag{A.35}$$

Then, the actual mean squared error is calculated by taking expectations as:

$$\begin{aligned}\boldsymbol{E}\left[\boldsymbol{\mu}_c\boldsymbol{\mu}_c^T\right] = \Sigma_{cc}\{&\omega^2\Sigma_{aa}^{-1}\bar{\Sigma}_{aa}\Sigma_{aa}^{-1} + \omega(1-\omega)\Sigma_{aa}^{-1}\bar{\Sigma}_{ab}\Sigma_{bb}^{-1} \\ &+\omega(1-\omega)\Sigma_{bb}^{-1}\bar{\Sigma}_{ba}\Sigma_{aa}^{-1} + (1-\omega)^2\Sigma_{bb}^{-1}\bar{\Sigma}_{bb}\Sigma_{bb}^{-1}\}\end{aligned} \tag{A.36}$$

where $\bar{\Sigma}_{pq} \triangleq \mathrm{E}[\tilde{\boldsymbol{\mu}}_p\tilde{\boldsymbol{\mu}}_q^T]$ and $\tilde{\boldsymbol{\mu}}_p \triangleq \boldsymbol{x}_p - \boldsymbol{\mu}_p$ $(p, q \in \{a, b\})$. Here, it is assumed that we do not have information on $\bar{\Sigma}_{ab}$, hence, we consider evaluating the upper bound of this quantity.

The consistency condition is formulated as:

$$\begin{aligned}\Sigma_{cc}^{-1} - \omega^2\Sigma_{aa}^{-1}\bar{\Sigma}_{aa}\Sigma_{aa}^{-1} - \omega(1-\omega)\Sigma_{aa}^{-1}\bar{\Sigma}_{ab}\Sigma_{bb}^{-1} \\ -\omega(1-\omega)\Sigma_{bb}^{-1}\bar{\Sigma}_{ba}\Sigma_{aa}^{-1} - (1-\omega)^2\Sigma_{bb}^{-1}\bar{\Sigma}_{bb}\Sigma_{bb}^{-1} \geq \mathbf{0}\end{aligned} \tag{A.37}$$

From the consistency condition of $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$:

$$\Sigma_{aa} - \bar{\Sigma}_{aa} \geq \mathbf{0}, \quad \Sigma_{bb} - \bar{\Sigma}_{bb} \geq \mathbf{0} \tag{A.38}$$

Then, we have another inequality conditions of Σ_{cc} as:

$$\begin{aligned}\Sigma_{cc}^{-1} &= \omega\Sigma_{aa}^{-1} + (1-\omega)\Sigma_{bb}^{-1} \\ &\geq \omega\Sigma_{aa}^{-1}\bar{\Sigma}_{aa}\Sigma_{aa}^{-1} + (1-\omega)\Sigma_{bb}^{-1}\bar{\Sigma}_{bb}\Sigma_{bb}^{-1}\end{aligned} \tag{A.39}$$

By substituting this lower bound to Eq. (A.37),

$$\omega(1-\omega)\left(\Sigma_{aa}^{-1}\bar{\Sigma}_{aa}\Sigma_{aa}^{-1} + \Sigma_{aa}^{-1}\bar{\Sigma}_{ab}\Sigma_{bb}^{-1} + \Sigma_{bb}^{-1}\bar{\Sigma}_{ba}\Sigma_{aa}^{-1} + \Sigma_{bb}^{-1}\bar{\Sigma}_{bb}\Sigma_{bb}^{-1}\right) \geq \mathbf{0} \tag{A.40}$$

then we finally get:

$$\omega(1-\omega)\mathrm{E}\left[\left\{\Sigma_{aa}^{-1}\tilde{\boldsymbol{\mu}}_a - \Sigma_{bb}^{-1}\tilde{\boldsymbol{\mu}}_b\right\}\left\{\Sigma_{aa}^{-1}\tilde{\boldsymbol{\mu}}_a - \Sigma_{bb}^{-1}\tilde{\boldsymbol{\mu}}_b\right\}^T\right] \geq \mathbf{0} \tag{A.41}$$

As obviously known from the above discussion, the consistency holds for any choices of $\bar{\Sigma}_{ab}$ and $\omega \in [0, 1]$.

A.5 Kullback-Leibler Divergence of Gaussian Distributions

The Kullback-Leibler (KL) divergence is the statistical distance between two probability distributions, how one probability distribution is different from the other probability distribution [56, 57]. The KL divergence for the continuous probability distributions is defined as:

$$D_{KL}(p||q) = \int_{\mathbb{R}^{n_x}} p(\boldsymbol{x}) \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} d\boldsymbol{x} \tag{A.42}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$. When we assume that the probability distributions follow the multivariate Guassian distributions, $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are represented as:

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n_x/2}\left|\Sigma_p\right|^{1/2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)^T\Sigma_p^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)\right) \tag{A.43}$$

$$q(\boldsymbol{x}) = \frac{1}{(2\pi)^{n_x/2}\left|\Sigma_q\right|^{1/2}} \exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_q\right)^T\Sigma_q^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_q\right)\right) \tag{A.44}$$

where $p(\boldsymbol{x}) = \mathcal{N}\left(\boldsymbol{x}|\boldsymbol{\mu}_p, \Sigma_p\right)$ and $q(\boldsymbol{x}) = \mathcal{N}\left(\boldsymbol{x}|\boldsymbol{\mu}_q, \Sigma_q\right)$. Eq. (A.42) is reformulated with Eqs. (A.43) and (A.44) as:

$$\begin{aligned} D_{KL}(p||q) &= \mathrm{E}\left[\log p(\boldsymbol{x}) - \log q(\boldsymbol{x})\right] \\ &= \frac{1}{2}\log\frac{|\Sigma_q|}{|\Sigma_p|} - \frac{1}{2}\mathrm{E}\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)^T \Sigma_p^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)\right] \\ &\quad + \frac{1}{2}\mathrm{E}\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_q\right)^T \Sigma_q^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_q\right)\right] \end{aligned} \tag{A.45}$$

The second term of Eq. (A.45) is rewritten with the matrix trace as:

$$\begin{aligned} (\text{the 2nd term}) &= \frac{1}{2}\mathrm{E}\left[\mathrm{trace}\left\{\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)^T \Sigma_p^{-1}\right\}\right] \\ &= \frac{1}{2}\mathrm{trace}\left\{\mathrm{E}\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)^T \Sigma_p^{-1}\right]\right\} \\ &= \frac{1}{2}\mathrm{trace}\left\{\mathrm{E}\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)\left(\boldsymbol{x}-\boldsymbol{\mu}_p\right)^T\right]\Sigma_p^{-1}\right\} \\ &= \frac{1}{2}\mathrm{trace}\left\{\Sigma_p\Sigma_p^{-1}\right\} = \frac{n_x}{2} \end{aligned} \tag{A.46}$$

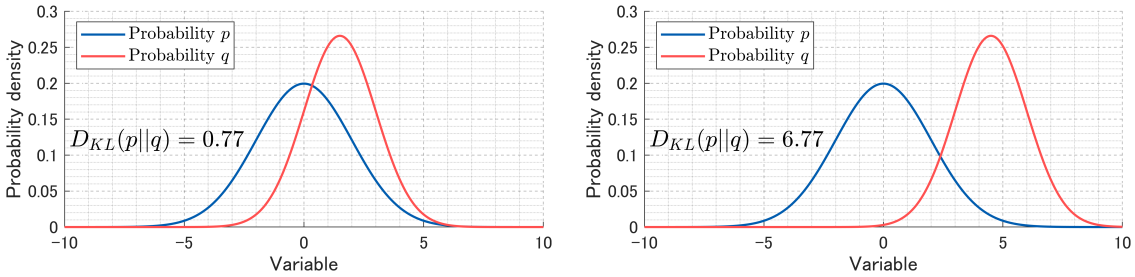
where the relationship $\mathrm{trace}(\boldsymbol{ABC}) = \mathrm{trace}(\boldsymbol{BCA}) = \mathrm{trace}(\boldsymbol{CAB})$ is used for the reformulation. The third term of Eq. (A.45) is simplified as:

$$(\text{the 3rd term}) = \left(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q\right)^T \Sigma_q^{-1}\left(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q\right) + \mathrm{trace}\left\{\Sigma_q^{-1}\Sigma_p\right\} \tag{A.47}$$

Finally, the KL-divergence of the multivariate Gaussian distributions is derived as:

$$D_{\mathrm{KL}}(p||q) = \frac{1}{2}\left[\log\frac{|\Sigma_q|}{|\Sigma_p|} - n_x + \left(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q\right)^T \Sigma_q^{-1}\left(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q\right) + \mathrm{trace}\left\{\Sigma_q^{-1}\Sigma_p\right\}\right] \tag{A.48}$$

Figure A.2 shows the examples of the KL-divergence.



(a) Small difference **(b)** Large difference

Figure A.2: Examples of the KL-divergence.

Appendix B

Evaluation of Information Gain

B.1 Measurement Information Gain

To compute the belief of the estimated states based on the measurements from other spacecraft, it is required to pay costs for communicating that information. This fact is one of the primary motivations to make the estimation process decentralized to reduce the amount of data transmitted among spacecraft. In this subsection, we consider how to evaluate the values of inter-spacecraft measurements.

B.1.1 Information Gain of Inter-spacecraft Measurements

The evaluation metrics of measurement information gain can be roughly divided into two categories: the evaluation based on covariance and the evaluation based on information matrix, that are summarized in Appendix B.2. Firstly, we consider the information contribution of the range and bearing angle measurements for the i-th spacecraft are represented as:

$$\text{trace}\left(\boldsymbol{U}_k^{r(ij)}\right) = \frac{1}{\sigma_r^2} \tag{B.1}$$

$$\text{trace}\left(\boldsymbol{U}_k^{\theta(ij)}\right) = \frac{1}{\sigma_b^2} \cdot \frac{1}{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2} \tag{B.2}$$

$$\text{trace}\left(\boldsymbol{U}_k^{\phi(ij)}\right) = \frac{1}{\sigma_b^2} \cdot \frac{1}{\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2 + \left(p_z^{(j)} - p_z^{(i)}\right)^2} \tag{B.3}$$

where the measurement noise variances are assumed to be identical for the same type of sensors. $\boldsymbol{U}_k^{r(ij)}$, $\boldsymbol{U}_k^{\theta(ij)}$, and $\boldsymbol{U}_k^{\phi(ij)}$ represent the contribution information matrices of the range, azimuth angle, and elevation angle measurements respectively. The derivations of

Eqs.(B.1) to (B.3) are provided in Appendix B.3.

If assuming that the information matrix of the prior probability is $\boldsymbol{Z}_{k|k-1}^{(i)i}$, the information utility metric is computed as:

$$\begin{aligned}\Psi^{(i)}\left(\boldsymbol{y}_k^{(ij)}\right) &= \text{trace}\left(\boldsymbol{Z}_{k|k-1}^{(i)i} + \sum_{s\in\{r,\theta,\phi\}} \boldsymbol{U}_k^{s(ij)}\right) - \text{trace}\left(\boldsymbol{Z}_{k|k-1}^{(i)i}\right) \\ &= \sum_{s\in\{r,\theta,\phi\}} \text{trace}\left(\boldsymbol{U}_k^{s(ij)}\right). \end{aligned} \tag{B.4}$$

where $\Psi^{(i)}(\boldsymbol{y}_k^{(ij)})$ represents the information gain of $\boldsymbol{y}_k^{(ij)}$ with respect to the state estimate of the i-th spacecraft. As obviously seen in Eq.(B.4), the measurement information gain evaluated in the canonical parameterization is assessed regardless of the prior probability. This evaluation is effective for conditions where information contribution values vary, such as heterogeneous sensor networks. However, in the swarm navigation relying solely on relative sensing, the information contribution is highly uniform as represented in Eqs.(B.1) to (B.3). There is difficulty exploiting the evaluation in the canonical parameterization since it would provide a meaningless assessment.

As an alternative way, we can consider the covariance-based evaluation. The information utility metric is formulated as:

$$\Psi^{(i)}\left(\boldsymbol{y}_k^{(ij)}\right) = -\text{trace}\left(\boldsymbol{Z}_{k|k-1}^{(i)i} + \sum_{s\in\{r,\theta,\phi\}} \boldsymbol{U}_k^{s(ij)}\right)^{-1} + \text{trace}\left(\boldsymbol{Z}_{k|k-1}^{(i)i}\right)^{-1}. \tag{B.5}$$

As mentioned in Appendix B.2, the trace of a covariance matrix is proportional to the circumference of the rectangular region enclosing the covariance ellipsoid. Therefore, Eq.(B.5) represents the difference of circumferences between the posterior and the prior covariance ellipsoids.

B.1.2 Example of Information Gain Evaluation

For evaluating information gain, we consider a small swarm system as shown in Figure B.1. In this scenario, the 1st spacecraft estimates its own state with the states of the adjacent spacecraft (the 2nd and the 3rd spacecraft). At the k-th time step, the 1st spacecraft already have measurement data expressed as the green edges in Figure B.1. We consider adding another inter-spacecraft measurement and evaluating its information gain based on this condition.

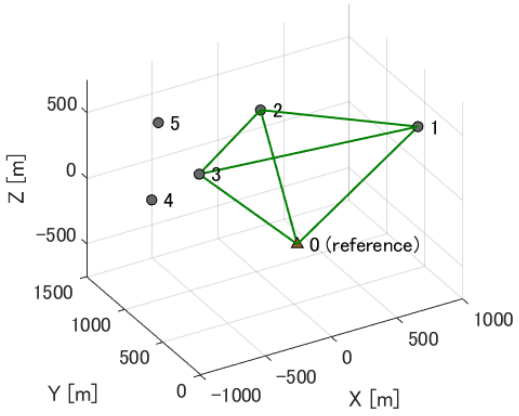


Figure B.1: Problem setting for the information gain analysis

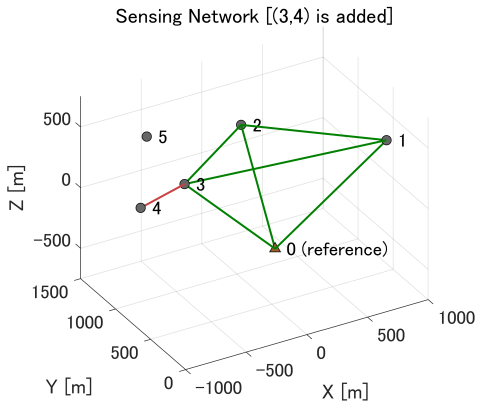
The aggregated state vector is represented as:

$$\boldsymbol{x}^{(\tilde{\mathcal{V}}_1)} = \begin{bmatrix} \left(\boldsymbol{x}^{(1)}\right)^T & \left(\boldsymbol{x}^{(2)}\right)^T & \left(\boldsymbol{x}^{(3)}\right)^T \end{bmatrix}^T \tag{B.6}$$

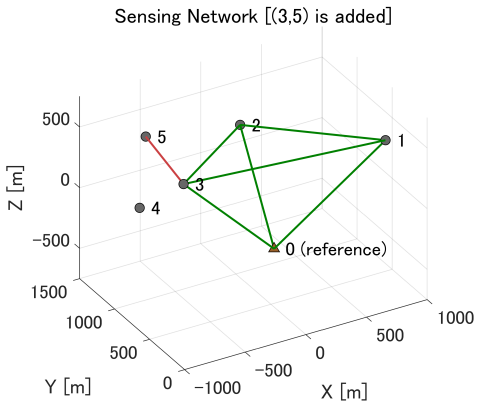
where the aggregated information vector and information matrix are $\hat{\boldsymbol{z}}^{(\tilde{\mathcal{V}}_1)}$ and $\boldsymbol{Z}^{(\tilde{\mathcal{V}}_1)}$ respectively. As shown in Figure B.2, we consider adding the measurement (3,4) or the measurement (3,5). The information gains of these measurements for the i-th spacecraft are evaluated as:

$$\begin{aligned} \Psi^{(i)}\left(\boldsymbol{y}^{(lm)}\right) = & \\ & - \mathrm{trace}\left(\Upsilon_p^{(i)}\left[\left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_1)i} + \sum_{s\in\{r,\theta,\phi\}}\sum_{i,j\in\tilde{\mathcal{V}}_1, i\neq j}\boldsymbol{U}_k^{s(ij)} \right.\right.\right. \\ & \qquad \left.\left.\left. + \sum_{s\in\{r,\theta,\phi\}}\boldsymbol{U}_k^{s(lm)}\right)^{-1}\right]\right) \\ & + \mathrm{trace}\left(\Upsilon_p^{(i)}\left[\left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_1)i} + \sum_{s\in\{r,\theta,\phi\}}\sum_{i,j\in\tilde{\mathcal{V}}_1, i\neq j}\boldsymbol{U}_k^{s(ij)}\right)^{-1}\right]\right) \end{aligned} \tag{B.7}$$

where $\tilde{\mathcal{V}}_1 = \{1,2,3\}$, $l \in \{3\} \subseteq \tilde{\mathcal{V}}_1$ and $m \in \{4,5\} \in \mathcal{V}$.



(a) (3,4) is added to the sensing network.



(b) (3,5) is added to the sensing network.

Figure B.2: Information gain of an additional measurement.

Table B.1 shows the evaluated information gains in the case of Figure B.2. In this case, the measurement (3,4) has higher information gain for all the spacecraft than the measurement (3,5).

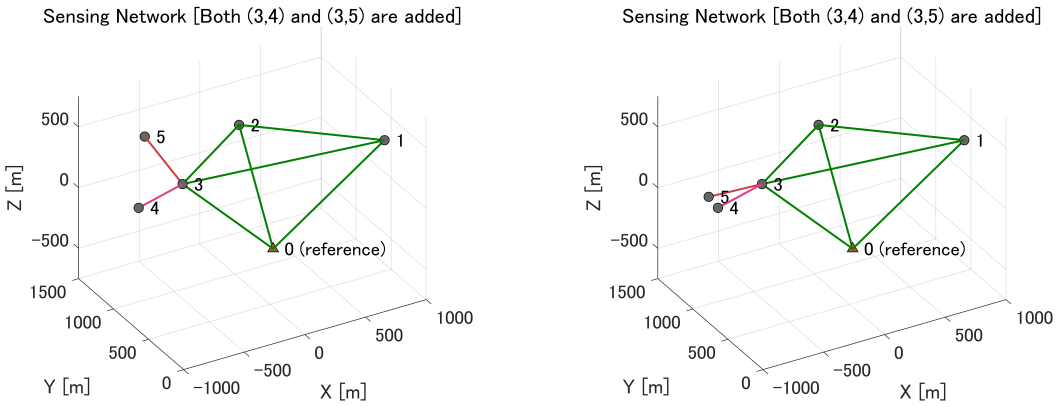
Table B.1: Information gain in the case of Figure B.2

Added edge	1st spacecraft	2nd spacecraft	3rd spacecraft
$(3,4)$	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,4)}\right) = 15.1$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,4)}\right) = 13.2$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,4)}\right) = 18.8$
$(3,5)$	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,5)}\right) = 8.8$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,5)}\right) = 12.9$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,5)}\right) = 17.1$

As the next example, we consider the information gain of the measurement (3.5) after already adding the measurement (3,4). The information gain is expressed as:

$$
\begin{aligned}
&\Psi^{(i)}\left(\boldsymbol{y}^{(lm)}\right) = \\
&\quad - \operatorname{trace}\left(\Upsilon_p^{(i)}\left[\left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_1)i} + \sum_{s\in\{r,\theta,\phi\}}\sum_{i,j\in\tilde{\mathcal{V}}_1, i\neq j}\boldsymbol{U}_k^{s(ij)} + \sum_{s\in\{r,\theta,\phi\}}\boldsymbol{U}_k^{s(lm)}\right)^{-1}\right]\right) \\
&\quad + \operatorname{trace}\left(\Upsilon_p^{(i)}\left[\left(\boldsymbol{Z}_{k|k-1}^{(\tilde{\mathcal{V}}_1)i} + \sum_{s\in\{r,\theta,\phi\}}\sum_{i,j\in\tilde{\mathcal{V}}_1, i\neq j}\boldsymbol{U}_k^{s(ij)} + \sum_{s\in\{r,\theta,\phi\}}\boldsymbol{U}_k^{s(3,4)}\right)^{-1}\right]\right)
\end{aligned}
\tag{B.8}
$$

where $\tilde{\mathcal{V}}_1 = \{1,2,3\}$, $l \in \{3\} \subseteq \tilde{\mathcal{V}}_1$ and $m \in \{4,5\} \in \mathcal{V}$. In this case, we consider two locations of the 5th spacecraft as shown in Figure B.3.



(a) Pattern 1 **(b)** Pattern 2

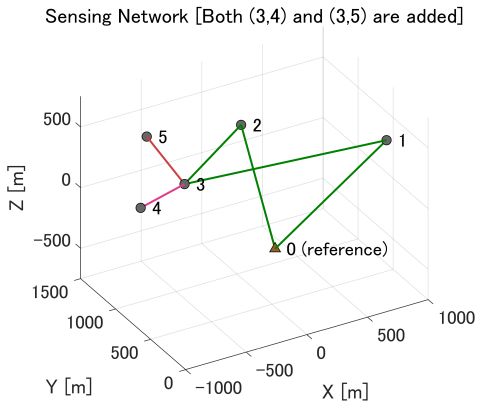
Figure B.3: The edge (3,5) is added after the edge (3,4) is added against the complete graph.

Table B.2 shows the information gain of the measurement (3,5) for both locations of the 5th spacecraft. The configuration of Pattern 1 provides slightly higher information gain than the configuration of Pattern 2.

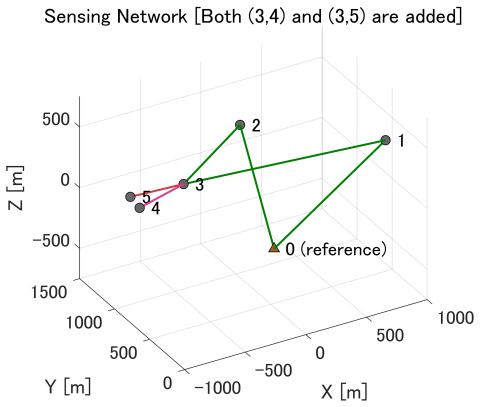
Table B.2: Information gain in the case of Figure B.3

Pattern	1st spacecraft	2nd spacecraft	3rd spacecraft
Pattern 1	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,5)}\right) = 4.84$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,5)}\right) = 2.83$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,5)}\right) = 5.62$
Pattern 2	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,5)}\right) = 4.85$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,5)}\right) = 2.78$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,5)}\right) = 5.56$

Finally, we consider adding the measurement (3,5) on the incomplete graph, in which the measurements (0,3) and (1,2) are removed from the complete graph as shown in Figure B.4. The information gain of the measurement (3,5) is evaluated after adding the measurement (3,4) on the above-mentioned incomplete graph. Eq. (B.8) is used for the calculation of the information gain.



(a) Pattern 1



(b) Pattern 2

Figure B.4: The edge (3,5) is added after the edge (3,4) is added against the complete graph.

Table B.3 shows the evaluation results. In contrast to the case of Figure B.3, the configuration of Pattern 2 provides remarkably higher information gain than the configuration of Pattern 1.

Table B.3: Information gain in the case of Figure B.4

Pattern	1st spacecraft	2nd spacecraft	3rd spacecraft
Pattern 1	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,5)}\right) = 3.43$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,5)}\right) = 1.35$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,5)}\right) = 7.44$
Pattern 2	$\Psi^{(1)}\left(\boldsymbol{y}^{(3,5)}\right) = 6.30$	$\Psi^{(2)}\left(\boldsymbol{y}^{(3,5)}\right) = 3.31$	$\Psi^{(3)}\left(\boldsymbol{y}^{(3,5)}\right) = 10.0$

B.1.3 Discussion on Measurement Information Gain

As the above-mentioned numerical evaluation shows, the information gain of the inter-spacecraft measurements heavily relies on the already obtained measurements. Therefore, it would be meaningless to consider the information of a single measurement. In addition, it is considered to evaluate the information gain with the onboard resources for an active selection of transferred measurement data to reduce inter-spacecraft communication load. To assess the information gain on a particular spacecraft, it needs to receive the related information, including measurement data and the estimation result of the communication target. However, if the spacecraft receives those data to evaluate the information gain, the original purpose of reducing the communication load is no longer achieved. Therefore, the onboard evaluation of the information gain is fundamentally difficult, especially for multi-

agent systems like spacecraft swarms. Based on the above analysis, it would be reasonable to decide the policy of measurement selection heuristically and develop a high-performance estimation algorithm working well under that policy.

B.2 Information Utility Metrics

To enable quantitative evaluation of the information gain provided by sensor measurements, it is required to define the measure of information utility. The intuition is that the information content is inversely related to the size of the uncertain region of the estimate of the random variable $\boldsymbol{x}$ [58]. We consider an estimate following the Gaussian distribution $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \Sigma)$. The determinant of the covariance matrix is proportional to the volume of the rectangular region enclosing the covariance ellipsoid. In the case of $\boldsymbol{x} \in \mathbb{R}^3$, the covariance matrix is defined as:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \tag{B.9}$$

Let the eigenvalues of Σ be λ_1, λ_2, and λ_3. The determinant of Σ is calculated with the following equation:

$$\det(\Sigma) = \prod_{i=1}^{3} \lambda_i \tag{B.10}$$

The standard form of the covariance ellipsoid is expressed as:

$$\frac{x^2}{\lambda_1^2} + \frac{y^2}{\lambda_2^2} + \frac{z^2}{\lambda_3^2} = 1 \tag{B.11}$$

Hence, the information utility function for this approximation is chosen as:

$$\Psi\left(p(\boldsymbol{x})\right) = -\det(\Sigma) \tag{B.12}$$

Although the volume of the high probability region seems to be a helpful measure, there are cases in which this measure underestimates the residual uncertainty. If the smallest principal axis shrinks to zero, the volume of the uncertainty ellipsoid is zero, while the uncertainties along the remaining principal axes might remain large. An alternative measure using only the covariance of a distribution $p(\boldsymbol{x})$ would be the trace, which is proportional

to the circumference of the rectangular region enclosing the covariance ellipsoid. That is because the summation of the eigenvalues equals to the trace as:

$$\mathrm{trace}(\Sigma) = \sum_{i=1}^{3} \lambda_i \tag{B.13}$$

Then, the information utility function is defined as:

$$\Psi\left(p(\boldsymbol{x})\right) = -\mathrm{trace}(\Sigma) \tag{B.14}$$

B.3 Information Gain of Range and Bearing Angle Measurements

We summarize the information metrics of range and bearing angle measurements. Here we consider the information gain of the measurements between the i-th and the j-th spacecraft for the state estimation of the i-th spacecraft. Firstly, the information contribution matrix of the range measurement is defined as:

$$\boldsymbol{U}_k^{r(ij)} = \frac{1}{\sigma_r^2}\left(\boldsymbol{H}_{r,k}^{(i)}\right)^T \boldsymbol{H}_{r,k}^{(i)} \tag{B.15}$$

$$\boldsymbol{H}_{r,k}^{(i)} = \frac{\partial h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right)}{\partial \boldsymbol{x}^{(i)}} = \begin{bmatrix} \frac{\partial h_r}{\partial p_x^{(i)}} & \frac{\partial h_r}{\partial p_y^{(i)}} & \frac{\partial h_r}{\partial p_z^{(i)}} & 0 & 0 & 0 \end{bmatrix}. \tag{B.16}$$

Then, the trace of the information contribution matrix is computed as:

$$\begin{aligned} \mathrm{trace}\left(\boldsymbol{U}_k^{r(ij)}\right) &= \frac{1}{\sigma_r^2} \sum_{\alpha\in\{x,y,z\}} \left(\frac{\partial h_r}{\partial p_\alpha^{(i)}}\right)^2 \\ &= \frac{1}{\sigma_r^2} \cdot \frac{1}{\left(h_r\left(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}\right)\right)^2} \\ &\qquad \times \left(\left(p_x^{(j)} - p_x^{(i)}\right)^2 + \left(p_y^{(j)} - p_y^{(i)}\right)^2 + \left(p_z^{(j)} - p_z^{(i)}\right)^2\right) \\ &= \frac{1}{\sigma_r^2} \end{aligned} \tag{B.17}$$

where the Jacobian terms are replaced with Eq.(2.16).

Next, the information contribution matrix of the range measurement is defined as:

$$\boldsymbol{U}_k^{b(ij)} = \left(\boldsymbol{H}_{b,k}^{(i)}\right)^T \boldsymbol{R}_b^{-1} \boldsymbol{H}_{b,k}^{(i)} = \boldsymbol{U}_k^{\theta(ij)} + \boldsymbol{U}_k^{\phi(ij)} \tag{B.18}$$

$$\boldsymbol{U}_k^{\theta(ij)} = \frac{1}{\sigma_b^2}\left(\boldsymbol{H}_{\theta,k}^{(i)}\right)^T \boldsymbol{H}_{\theta,k}^{(i)} \tag{B.19}$$

$$\boldsymbol{U}_k^{\phi(ij)} = \frac{1}{\sigma_b^2}\left(\boldsymbol{H}_{\phi,k}^{(i)}\right)^T \boldsymbol{H}_{\phi,k}^{(i)} \tag{B.20}$$

where the Jacobian matrices are defined as:

$$\boldsymbol{H}_{\theta,k}^{(i)} = \frac{\partial h_\theta\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)}{\partial \boldsymbol{x}^{(i)}} = \begin{bmatrix} \frac{\partial \theta^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial \theta^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial \theta^{(ij)}}{\partial p_z^{(i)}} & 0 & 0 & 0 \end{bmatrix} \tag{B.21}$$

$$\boldsymbol{H}_{\phi,k}^{(i)} = \frac{\partial h_\phi\left(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)}\right)}{\partial \boldsymbol{x}^{(i)}} = \begin{bmatrix} \frac{\partial \phi^{(ij)}}{\partial p_x^{(i)}} & \frac{\partial \phi^{(ij)}}{\partial p_y^{(i)}} & \frac{\partial \phi^{(ij)}}{\partial p_z^{(i)}} & 0 & 0 & 0 \end{bmatrix} \tag{B.22}$$

The trace of the information contribution matrix of the azimuth angle measurement is formulated as:

$$\begin{aligned} \text{trace}\left(\boldsymbol{U}_k^{\theta(ij)}\right) &= \frac{1}{\sigma_b^2}\sum_{\alpha\in\{x,y,z\}}\left(\frac{\partial h_\theta}{\partial p_\alpha^{(i)}}\right)^2 \\ &= \frac{1}{\left(1+\xi^2\right)^2}\left(1+\xi^2\right)\cdot\frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2}+0^2 \\ &= \frac{1}{1+\frac{\left(p_y^{(j)}-p_y^{(i)}\right)^2}{\left(p_x^{(j)}-p_x^{(i)}\right)^2}}\cdot\frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2} \\ &= \frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2}, \end{aligned} \tag{B.23}$$

and the trace of the information contribution matrix of the elevation angle measurement is formulated as:

$$\begin{aligned} \text{trace}\left(\boldsymbol{U}_k^{\phi(ij)}\right) &= \frac{1}{\sigma_b^2}\sum_{\alpha\in\{x,y,z\}}\left(\frac{\partial h_\phi}{\partial p_\alpha^{(i)}}\right)^2 \\ &= \left(\frac{\zeta}{1+\zeta^2}\right)^2\cdot\frac{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2}{\left(\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2\right)^2} \\ &\quad+\frac{1}{\left(1+\zeta^2\right)^2}\cdot\frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2} \\ &= \frac{1}{1+\frac{\left(p_z^{(j)}-p_z^{(i)}\right)^2}{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2}} \\ &\quad\times\frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2} \\ &= \frac{1}{\left(p_x^{(j)}-p_x^{(i)}\right)^2+\left(p_y^{(j)}-p_y^{(i)}\right)^2+\left(p_z^{(j)}-p_z^{(i)}\right)^2} \end{aligned} \tag{B.24}$$

where the Jacobian terms are replaced with Eqs.(2.24) to (2.30).

Bibliography

[1] Fred Y Hadaegh, Soon-Jo Chung, and Harish M Manohara. On development of 100-gram-class spacecraft for swarm applications. *IEEE Systems Journal*, 10(2):673–684, 2014.

[2] A Budianu, Arjan Meijerink, and Marinus Jan Bentum. Swarm-to-earth communication in olfar. *Acta astronautica*, 107:14–19, 2015.

[3] MJ Bentum, MK Verma, RT Rajan, AJ Boonstra, CJM Verhoeven, EKA Gill, AJ van der Veen, H Falcke, M Klein Wolt, B Monna, et al. A roadmap towards a space-based radio telescope for ultra-low frequency radio astronomy. *Advances in Space Research*, 65(2):856–867, 2020.

[4] Martijn de Kok, Jose Velazco, and Mark Bentum. Cubesat array for detection of rf emissions from exoplanets using inter-satellite optical communicators. In *2020 IEEE Aerospace Conference*, pages 1–12. IEEE, 2020.

[5] Rebecca C Foust, E Sorina Lupu, Yashwanth K Nakka, Soon-Jo Chung, and Fred Y Hadaegh. Autonomous in-orbit satellite assembly from a modular heterogeneous swarm. *Acta Astronautica*, 169:191–205, 2020.

[6] Shaozhao Lu, Yao Zhang, Xingang Li, and Quan Hu. Decentralized closed-loop optimization for 6-dof self-assembly satellites. *Acta Astronautica*, 189:593–605, 2021.

[7] Nicolas Lee and Sigrid Close. Meteoroid impact detection for exploration of asteroids: Small satellites for asteroid characterization. *Journal of Spacecraft and Rockets*, 55(1):202–213, 2018.

[8] Julie C Castillo-Rogez, Karen Meech, Soon-Jo Chung, and Damon Landau. Approach to exploring interstellar objects and long-period comets. 2019.

[9] Hugo Sanchez, Dawn McIntosh, Howard Cannon, Craig Pires, Josh Sullivan, Simone D ' Amico, and Brendan O ' Connor. Starling1: Swarm technology demonstration. 2018.

[10] Luis C Carrillo-Arce, Esha D Nerurkar, José L Gordillo, and Stergios I Roumeliotis. Decentralized multi-robot cooperative localization using covariance intersection. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1412–1417. IEEE, 2013.

[11] Amir Rahmani, Saptarshi Bandyopadhyay, Federico Rossi, Jean-Pierre de la Croix, Joshua Vander Hook, and Michael Wolf. Space vehicle swarm exploration missions: A study of key enabling technologies and gaps. 2019.

[12] Saptarshi Bandyopadhyay, Rebecca Foust, Giri P Subramanian, Soon-Jo Chung, and Fred Y Hadaegh. Review of formation flying and constellation missions using nanosatellites. *Journal of Spacecraft and Rockets*, 53(3):567–578, 2016.

[13] Nathan Stacey and Simone D ' Amico. Autonomous swarming for simultaneous navigation and asteroid characterization. In *AAS/AIAA Astrodynamics Specialist Conference*, volume 1, 2018.

[14] Martin Lo and Shane Ross. The lunar l1 gateway: portal to the stars and beyond. 2001.

[15] Giuseppe Di Mauro, Margaret Lawn, and Riccardo Bevilacqua. Survey on guidance navigation and control requirements for spacecraft formation-flying missions. *Journal of Guidance, Control, and Dynamics*, 41(3):581–602, 2018.

[16] Lorenzo Pettazzi, Dario Izzo, and S Theil. Swarm navigation and reconfiguration using electrostatic forces. In *7th International Conference on Dynamics and Control of Systems and Structures in Space, Greenwich, London, England*, pages 257–267. Citeseer, 2006.

[17] Corinne Lippe and Simone D ' Amico. Optimal spacecraft swarm reconfiguration through chief orbit refinement. *Acta Astronautica*, 183:162–175, 2021.

[18] Daniel Morgan, Soon-Jo Chung, and Fred Y Hadaegh. Model predictive control of swarms of spacecraft using sequential convex programming. *Journal of Guidance, Control, and Dynamics*, 37(6):1725–1740, 2014.

[19] Reza Olfati-Saber and Jeff S Shamma. Consensus filters for sensor networks and distributed sensor fusion. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 6698–6703. IEEE, 2005.

[20] Reza Olfati-Saber. Distributed kalman filter with embedded consensus filters. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 8179–8184. IEEE, 2005.

[21] Reza Olfati-Saber. Distributed kalman filtering for sensor networks. In *2007 46th IEEE Conference on Decision and Control*, pages 5492–5498. IEEE, 2007.

[22] Ondrej Hlinka, Ondrej Slučiak, Franz Hlawatsch, and Markus Rupp. Distributed data fusion using iterative covariance intersection. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1861–1865. IEEE, 2014.

[23] Saptarshi Bandyopadhyay and Soon-Jo Chung. Distributed bayesian filtering using logarithmic opinion pool for dynamic sensor networks. *Automatica*, 97:7–17, 2018.

[24] Hao Li and Fawzi Nashashibi. Cooperative multi-vehicle localization using split covariance intersection filter. *IEEE Intelligent transportation systems magazine*, 5(2):33–44, 2013.

[25] Solmaz S Kia, Stephen Rounds, and Sonia Martinez. Cooperative localization for mobile agents: A recursive decentralized algorithm based on kalman-filter decoupling. *IEEE Control Systems Magazine*, 36(2):86–101, 2016.

[26] Anusna Chakraborty, Clark N Taylor, Rajnikant Sharma, and Kevin M Brink. Cooperative localization for fixed wing unmanned aerial vehicles. In *2016 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 106–117. IEEE, 2016.

[27] Jan Dentler, Somasundar Kannan, Souad Bezzaoucha, Miguel Angel Olivares-Mendez, and Holger Voos. Model predictive cooperative localization control of multiple uavs using potential function sensor constraints. *Autonomous Robots*, 43(1):153–178, 2019.

[28] Xiwang Dong, Yangfan Li, Chuang Lu, Guoqiang Hu, Qingdong Li, and Zhang Ren. Time-varying formation tracking for uav swarm systems with switching directed topologies. *IEEE transactions on neural networks and learning systems*, 30(12):3674–3685, 2018.

[29] Wonsuk Lee, Hyochoong Bang, and Henzeh Leeghim. Cooperative localization between small uavs using a combination of heterogeneous sensors. *Aerospace science and technology*, 27(1):105–111, 2013.

[30] Yaohong Qu, Jizhi Wu, Bing Xiao, and Dongli Yuan. A fault-tolerant cooperative positioning approach for multiple uavs. *IEEE Access*, 5:15630–15640, 2017.

[31] Jiuqing Wan, Liping Zhong, and Fan Zhang. Cooperative localization of multi-uavs via dynamic nonparametric belief propagation under gps signal loss condition. *International Journal of Distributed Sensor Networks*, 10(2):562380, 2014.

[32] Yongcai Wang, Tianyuan Sun, Guoyao Rao, and Deying Li. Formation tracking in sparse airborne networks. *IEEE Journal on Selected Areas in Communications*, 36(9):2000–2014, 2018.

[33] Jialong Zhang, Jianguo Yan, Maolong Lv, Xiangjie Kong, and Pu Zhang. Uav formation flight cooperative tracking controller design. In *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 856–861. IEEE, 2018.

[34] Pablo O Arambel, Constantino Rago, and Raman K Mehra. Covariance intersection algorithm for distributed spacecraft state estimation. In *Proceedings of the 2001 American Control Conference.(Cat. No. 01CH37148)*, volume 6, pages 4398–4403. IEEE, 2001.

[35] William Bezouska and David Barnhart. Decentralized cooperative localization with relative pose estimation for a spacecraft swarm. In *2019 IEEE Aerospace Conference*, pages 1–13. IEEE, 2019.

[36] Philip Ferguson and Jonathan How. Decentralized estimation algorithms for formation flying spacecraft. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, page 5442, 2003.

[37] Fred Y Hadaegh, B Kang, and Daniel P Scharf. Rule-based estimation and control of formation flying spacecraft. 2001.

[38] Kai Matsuka, Aaron O Feldman, Elena S Lupu, Soon-Jo Chung, and Fred Y Hadaegh. Decentralized formation pose estimation for spacecraft swarms. *Advances in Space Research*, 67(11):3527–3545, 2021.

[39] Terence McLoughlin and Mark Campbell. Distributed estimate fusion filter for large spacecraft formations. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, page 6666, 2008.

[40] Hermann Kaptui Sipowa, Jay W McMahon, and Taralicin Deka. Distributed unscented-information kalman filter (uikf) for cooperative localization in spacecraft formation flying. In *AIAA Scitech 2020 Forum*, page 1917, 2020.

[41] X Rong Li and Vesselin P Jilkov. Survey of maneuvering target tracking: Iii. measurement models. In *Signal and Data Processing of Small Targets 2001*, volume 4473, pages 423–446. SPIE, 2001.

[42] Henk Wymeersch, Jaime Lien, and Moe Z Win. Cooperative localization in wireless networks. *Proceedings of the IEEE*, 97(2):427–450, 2009.

[43] Bradley T Burchett. Unscented kalman filters for range-only cooperative localization of swarms of munitions in three-dimensional flight. *Aerospace Science and Technology*, 85:259–269, 2019.

[44] Kyle T Alfriend, Srinivas R Vadali, Pini Gurfil, Jonathan P How, and Louis Breger. *Spacecraft formation flying: Dynamics, control and navigation*, volume 2. Elsevier, 2009.

[45] WH Clohessy and RS Wiltshire. Terminal guidance system for satellite rendezvous. *Journal of the Aerospace Sciences*, 27(9):653–658, 1960.

[46] Samuel A Schweighart and Raymond J Sedwick. High-fidelity linearized j model for satellite formation flight. *Journal of Guidance, Control, and Dynamics*, 25(6):1073–1080, 2002.

[47] Francesco Bullo, Jorge Cortés, and Sonia Martinez. *Distributed control of robotic networks*. Princeton University Press, 2009.

[48] Zhe Chen et al. Bayesian filtering: From kalman filters to particle filters, and beyond. *Statistics*, 182(1):1–69, 2003.

[49] Kazufumi Ito and Kaiqi Xiong. Gaussian filters for nonlinear filtering problems. *IEEE transactions on automatic control*, 45(5):910–927, 2000.

[50] Sebastian Thrun. Probabilistic robotics. *Communications of the ACM*, 45(3):52–57, 2002.

[51] Simon J Julier and Jeffrey K Uhlmann. A non-divergent estimation algorithm in the presence of unknown correlations. In *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, volume 4, pages 2369–2373. IEEE, 1997.

[52] FWJ Van den Berg, HCJ Hoefsloot, HFM Boelens, and AK Smilde. Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria. *Chemical engineering science*, 55(4):827–837, 2000.

[53] Kaare Brandt Petersen, Michael Syskind Pedersen, et al. The matrix cookbook. *Technical University of Denmark*, 7(15):510, 2008.

[54] Yuanxin Wu, Dewen Hu, Meiping Wu, and Xiaoping Hu. A numerical-integration perspective on gaussian filters. *IEEE Transactions on Signal Processing*, 54(8):2910–2921, 2006.

[55] Simo Särkkä. *Bayesian filtering and smoothing*. Number 3. Cambridge university press, 2013.

[56] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[57] Javier E Contreras-Reyes and Reinaldo B Arellano-Valle. Kullback–leibler divergence measure for multivariate skew-normal distributions. *Entropy*, 14(9):1606–1626, 2012.

[58] Maurice Chu, Horst Haussecker, and Feng Zhao. Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks. *The International Journal of High Performance Computing Applications*, 16(3):293–313, 2002.