DECENTRALIZED SCALAR FIELD MAPPING USING GAUSSIAN PROCESSES

By

HOSSEIN PAPI

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2026

To My Madar, Parvin, My Pedar, Ahmad, and my Khahar, Zeinab

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Rushikesh Kamalapurkar, for his guidance, support, and mentorship throughout my master's studies. His advice and feedback helped shape both my development as a researcher and as a person. Beyond his technical guidance, I am especially grateful for the way he approaches research and mentorship. His curiosity, discipline, and genuine appreciation for scientific inquiry have made the research process more meaningful and enjoyable for me and have helped me develop a deeper respect for the purpose and practice of science.

I would also like to thank my cochair, Dr. Dixon, for his time, helpful advice, and support during this work.

I am grateful to my lab mates in the SCC Lab for their discussions, encouragement, and assistance throughout this project. In particular, I would like to thank Muzaffar Qureshi for his support, collaboration, and helpful conversations during the development of this research.

TABLE OF CONTENTS

page

LIST OF FIGURES

Figures page

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

DECENTRALIZED SCALAR FIELD MAPPING USING GAUSSIAN PROCESSES

By

Hossein Papi

August 2026

Chair: Rushikesh Kamalapurkar
Major: Mechanical Engineering

Scalar-field mapping is an important problem in robotic sensing, where mobile agents estimate an unknown, spatially distributed quantity from sparse and noisy measurements in applications such as environmental monitoring, wildfire assessment, autonomous exploration, and others. In these settings, centralized mapping methods can be difficult to use. This thesis develops a decentralized Gaussian-process mapping framework that allows multiple sensing agents to improve their local field estimates through minimal neighbor-to-neighbor information exchange.

The proposed method uses overlapping local sensing domains as the interface for decentralized cooperation. Each agent maintains a local Gaussian-process model using its own measurements and selected information received from neighboring agents. When neighboring agents estimate the same overlap region using different local datasets, their posterior means and variances may disagree. This disagreement is used as a signal for deciding which information should be exchanged. Instead of transmitting full local datasets, agents exchange minimal inducing-point packets containing a location, posterior mean, and posterior variance. Selected packets are then retained as fictitious measurements in the local model of receiving agent.

The method is evaluated in simulation on a decentralized scalar-field mapping problem with four agents. The results show that minimal packet exchange can improve prediction accuracy while preserving decentralized computation and avoiding full data sharing. Overall, this thesis demonstrates that overlap-based posterior disagreement can provide a practical mechanism for communication-efficient decentralized scalar-field mapping.

CHAPTER 1
INTRODUCTION

1.1 Motivation

Scalar-field mapping involves estimating a spatially distributed physical quantity across a domain using a finite, often sparse, set of localized measurements. In multi-agent robotics, this task is typically executed by dispersing mobile sensors to sample the environment and cooperatively reconstruct the underlying field. This framework is essential for missions where human entry is unsafe or impractical, such as radar-based spatial mapping, smoke concentration, radiation-field mapping, bathymetry, and chemical concentration fields. Related spatial prediction problems also appear in urban and coastal environment, where climate impacts, land-cover change, and shoreline vulnerability must be assessed using finite, sparse, and often incomplete spatial observations [35, 28, 34, 11, 26, 8, 2]. In multi-agent robotics, this task is typically executed by dispersing mobile sensors to sample the environment and cooperatively reconstruct the underlying field.

While centralized mapping algorithms offer a straightforward approach by pooling all measurements at a central base station, they scale poorly. As the operational environment expands, the communication bandwidth required to transmit raw data back to a central node quickly becomes a bottleneck. Furthermore, centralized systems introduce a single point of failure, making them poorly suited for hazardous deployments. While decentralized alternatives mitigate these scaling issues, they introduce severe coordination challenges. Previous work has highlighted that independent, single-agent mapping suffers from poor spatial coverage, whereas existing multi-agent frameworks often demand excessive data sharing to maintain consistency across the network [24].

1.2 Gaussian Processes for Scalar-Field Estimation

Scalar-field estimation focuses on reconstructing an unknown continuous function that maps a spatial coordinate to a scalar physical quantity. Let $g : \mathcal{X} \to \mathbb{R}$ represent this unknown field over a bounded domain $\mathcal{X} \subset \mathbb{R}^d$, where $g(x)$ denotes a spatially varying property such as local temperature or another spatially varying physical quantity. In robotic applications, this field is sampled selectively via sparse, noise-corrupted measurements gathered by mobile platforms. A

single observation taken at a location x_j is modeled as

$$y_j = g(x_j) + \varepsilon_j, \tag{1-1}$$

where $\varepsilon_j \sim \mathcal{N}(0, \sigma_n^2)$ isolates the stochastic measurement noise. The fundamental objective is to reliably infer the true state of field g at unobserved query locations using only this highly restricted, finite data array.

A Gaussian process[38]is a nonparametric Bayesian model, and is a distribution defined over functions, $f(x)$, where$f(x)$ is a function mapping the input space Ω_x to,

$$g(x) : \Omega_x \rightarrow \mathbb{R}, x \in \Omega_x. \tag{1-2}$$

A non-parametric GP is defined by its mean function $m(x)$ and covariance kernel $k(x, x')$:

$$g(x) \sim \mathcal{GP}\left(m(x), k(x, x')\right). \tag{1-3}$$

The choice of the covariance kernel $k(x, x')$ embeds fundamental physical assumptions regarding the field, including spatial correlation lengths, structural smoothness, and signal variance [27]. Because environmental features rarely follow clean analytical profiles, this non-parametric flexibility is useful. Rather than selecting a fixed, low-dimensional basis model, a GP acts as a flexible prior that updates into a posterior distribution upon conditioning on real-world measurements. Critically, this posterior representation gives both a predicted mean and a predictive variance at each query location, making GPs useful not only for interpolation but also for quantifying uncertainty.

This ability to represent uncertainty makes GPs useful for active sensing. Autonomous platforms cannot simply sample blindly; they must actively compute where to deploy their sensors to yield the highest informational return. The local GP posterior variance acts as an explicit exploratory metric. By mapping this variance across the domain, agents can pinpoint data-starved regions, drive targeted motion planning, and strike a mathematically grounded balance between

exploring unmapped zones and exploiting known areas of high intensity [32, 3, 1]. Consequently, GPs have become a cornerstone in adaptive multi-UAV sampling, environmental tracking, and safety-critical hazard mapping [24, 6, 21, 5]. For instance, a drone swarm can use local GP uncertainty bounds to actively avoid blind spots while simultaneously tracking the peak intensity of a hazardous plume without requiring exhaustive grid searches [24].

Furthermore, the strength of the GP framework lies in its data efficiency. Real-world robotic deployments are limited by flight times, hardware costs, and dangerous terrains, making dense, exhaustive sampling impractical. GPs address this data sparsity by leveraging the spatial correlation defined by the kernel; a single measurement updates the model's confidence about the entire local neighborhood. When hyperparameters are properly tuned, the posterior smoothly interpolates across unobserved gaps while letting the variance naturally widen as queries move further from known data tracks. This makes GPs a natural fit for limited-pass sensing paradigms, where an agent might only have a single opportunity to traverse a corridor or wildfire front.

Despite these clear benefits, exact GP inference becomes difficult as the number of measurements increases. Training a standard GP model on N total measurements requires the inversion or Cholesky decomposition of an $N \times N$ covariance matrix, scaling cubically at $O(N^3)$, while predictions scale at $O(N^2)$ [27, 18]. While these overheads are trivial for small, static datasets, they become computationally challenging for online, real-time mapping over large areas. This computational bottleneck is severely compounded in multi-agent networks; requiring resource-constrained edge computers to process or transmit a large global measurement history is often impractical.

Sparse GP formulations resolve this bottleneck by representing the full dataset using a smaller set of pseudo-inputs, commonly referred to as inducing points. The objective here is to capture the underlying covariance structure while reducing the size of the matrices that require inversion. These approximate methods offer a structured, mathematically sound trade-off between absolute predictive accuracy and computational runtime [23, 31, 36, 30, 14]. As synthesized by Quinonero-Candela and Rasmussen, these diverse sparse techniques can be understood as exact inference over modified,

approximate priors [23]. This inducing-point view is useful for scalable multi-robot applications: it shows that a compact set of artificial measurements can summarize much of the information in a larger raw dataset.

Spatially structured sensing data is uniquely suited for these local and sparse approximations. Because spatial correlation decays with distance, localized GP variants exploit the fact that distant observations have a negligible statistical impact on local predictions compared to immediate neighbors [12, 7]. Modern nearest-neighbor architectures and covariance-based block selection strategies use this correlation to drop irrelevant data points entirely, reducing computational cost [19]. These local approximations are attractive for environmental robotics, allowing localized chunks of a map to be updated rapidly without triggering a global re-computation of the entire field.

In the context of this thesis, Gaussian processes serve as the main probabilistic model for decentralized field estimation. The posterior mean gives each robot its local operational map, while the posterior variance offers a continuous diagnostic tool to monitor local map quality. Sparse inducing-point representations provide a mechanism for compressing local posterior information into lightweight, edge-specific data packets. This thesis builds directly on these principles. Rather than attempting to replace or bypass the GP framework, this work reshapes it for a decentralized landscape where agents must reduce local map disagreements over overlapping domains through the exchange of compact, low-dimensional summaries.

1.3 Decentralized Multi-Agent Mapping

A centralized approach to scalar-field mapping assumes that all measurements collected by all agents are transmitted to a single processing unit, where a global model is trained and predictions are generated for the entire domain. In principle, this architecture is attractive because the central unit can use all available data simultaneously and avoid inconsistencies between local models. For Gaussian-process-based mapping, a centralized model can exploit correlations across the entire dataset and produce a unified posterior distribution over the field. However, centralized mapping becomes difficult in large-scale or time-critical deployments because it requires reliable communication from every agent to the central node, sufficient bandwidth to transmit all

measurements, and enough computation to train or update the global model. Since exact GP training scales cubically with the number of observations, the computational burden grows rapidly as agents collect more data [27, 18].

These limitations motivate the move from single-agent or centralized mapping to decentralized multi-agent mapping. Multi-agent systems are useful because multiple robots can collect measurements from different parts of the environment simultaneously, improving spatial coverage and reducing the time required to observe a large domain. Prior work on scalar-field mapping has emphasized that single-agent approaches face limited spatial coverage, while multi-agent systems can improve the speed and accuracy of field estimation by allowing agents to collect data in parallel and use information from neighboring agents [24]. Multi-agent teams have been used in environmental sensing, coverage control, active search, adaptive sampling, mobility-on-demand modeling, and field learning problems [6, 21, 5, 13, 33, 15].

In decentralized mapping, each agent maintains its own local dataset and computes a local estimate of the scalar field. Let $D_i = (x_j^i, y_j^i)_{j=1}^{n_i}$ denote the dataset collected by agent i. Instead of constructing a global dataset $D = \cup_i D_i$ at a central node, agent i trains or updates a local GP using D_i and possibly limited information received from nearby agents. This architecture reduces the need for global data transmission and allows agents to continue operating when communication with a central node is unavailable. Such a structure is especially important in GPS-denied, underwater, underground, or disaster-response where communication range may be limited or time-varying [25, 16, 18].

A major line of prior work has studied decentralized or distributed GP prediction by decomposing the full GP problem across agents or data blocks. For example, decentralized data fusion and active sensing methods distribute the computational load of approximate GP prediction among mobile sensors while providing predictive performance comparable to a centralized partially independent training conditional approximation [6]. This type of approach shows that decentralized GP prediction can be theoretically connected to centralized sparse approximations, allowing the computational cost to be shared among sensors. However, the method still requires a carefully

chosen support set and may introduce communication overhead as the support set grows.

Other approaches use product-of-experts, Bayesian committee machines, nested aggregation, or nearest-neighbor aggregation to combine local GP predictions. Distributed GP methods partition data across local experts and aggregate their predictions, reducing computation per agent but introducing challenges in uncertainty calibration and consistency [9, 16, 19]. Product-of-experts methods can become overconfident when local experts are combined without careful variance correction, while generalized or robust Bayesian committee methods attempt to correct this effect by weighting expert predictions [37, 20, 19]. Nested GP aggregation and covariance-based nearest-neighbor methods further reduce computation and communication by selecting only statistically relevant agents for prediction [16, 19]. These methods are useful comparison points for this thesis because they address the same broad goal of scalable multi-agent GP prediction, but they generally focus on aggregating predictions or hyperparameters rather than exchanging overlap-targeted inducing-point packets.

Federated GP learning provides another related direction. In federated learning, agents collaborate to learn a shared model without exchanging raw local datasets. Recent federated and decentralized GP training methods use distributed optimization, such as ADMM, to estimate GP hyperparameters while preserving local data privacy [18, 29, 19]. These methods are important because GP performance depends strongly on kernel hyperparameters, and inaccurate hyperparameter estimates can degrade prediction quality. Federated GP methods therefore address a different but complementary problem: how agents can agree on a global or shared model structure without pooling all raw measurements. However, these methods may require multiple communication rounds, consensus iterations, or global coordination of model parameters, which can be difficult in time-limited mapping missions.

Sparse and online GP approaches have also been developed for multi-agent coverage and adaptive sampling. Distributed online sparse GP regression has been used to estimate unknown environmental density functions while simultaneously controlling agents for coverage [13]. Multi-agent adaptive estimation and coverage control methods model an unknown sensory function

as a Gaussian random field and design control laws that balance coverage and estimation [5]. These approaches highlight the close coupling between estimation and motion planning in multi-agent mapping. In contrast, the focus of this thesis is not primarily on optimizing coverage trajectories, but on improving decentralized GP map quality through compact communication between neighboring agents.

The key challenge in decentralized multi-agent mapping is that local computation and limited communication create fragmented information. Each agent observes only a subset of the domain, and its GP posterior is shaped by its local measurements, local hyperparameters, and any information received from neighbors. If agents do not communicate, their local maps may be accurate near their own measurements but poor elsewhere. If they communicate all raw data, the method becomes expensive and may approach centralized data fusion. Therefore, a practical decentralized method should share enough information to improve predictive performance, but not so much that it loses the scalability and robustness advantages of decentralization.

This thesis addresses this challenge through compact, edge-specific information exchange. Instead of requiring each agent to transmit its full dataset or participate in a global aggregation procedure, neighboring agents whose local domains overlap exchange compact inducing-point packets. These packets summarize the information collected over covered region of each agent, rather than being limited to the overlapping area alone. This design is motivated by the observation that the most useful information for a neighbor is not necessarily the sender's entire dataset, but rather a compact representation of the sender's posterior behavior in the region where both agents make predictions. This efficient communication strategy allows the method to preserve decentralized computation while still improving consistency and accuracy in shared regions of the field. In Chapter 3, this approach can be compared against several classes of existing methods: local-only GP baselines, centralized or approximate centralized GP fusion, decentralized data-fusion methods such as D2FAS [6], expert-aggregation methods such as PoE, BCM, and NPAE [4, 9], federated GP training methods [17, 18], and distributed sparse GP methods for coverage control [10, 22].

1.4 Overlapping Domains and Posterior Disagreement

In decentralized scalar-field mapping, agents often operate over local subdomains rather than the entire environment. Let $\mathcal{X}_i \subset \mathcal{X}$ denote the local domain assigned to agent i. In many multi-agent deployments, neighboring local domains are not disjoint. Instead, they overlap:

$$\mathcal{O}^{ij} := \mathcal{X}^i \cap \mathcal{X}^j \neq \emptyset.$$

These overlap regions are important because they represent portions of the field where more than one agent may have data. Overlap can arise naturally from coverage requirements, communication geometry, safety margins, sensor footprints, or partitioning strategies that intentionally include shared boundaries between neighboring agents. In a centralized GP, all data are processed together, so the posterior over an overlap region is unique. In a decentralized setting, however, each agent constructs its own GP posterior using only its local dataset and locally available information. Consequently, two neighboring agents may produce different predictions for the same physical location.

This mismatch can be described as posterior disagreement. Suppose agent i and agent j both model the scalar field over an overlap point $x \in \mathcal{O}_{ij}$. Local GP of Agent i posterior may produce

$$(g(x) \mid D_i) \sim \mathcal{N}(\mu_i(x), \sigma_i^2(x)),$$

while local GP posterior of agent j may produce

$$(g(x) \mid D_j) \sim \mathcal{N}(\mu_j(x), \sigma_j^2(x)).$$

Even though both agents are estimating the same underlying scalar field, their posterior means and variances may differ because their local datasets contain different measurements. Agent i may have observed points near one side of the overlap region, while agent j may have observed points near the other side. Their local kernels, hyperparameters, noise estimates, or sparse approximations may

also differ. Therefore, the overlap region becomes a place where decentralized information fragmentation is directly visible.

Posterior disagreement is not necessarily a failure of the GP framework. Instead, it is a natural consequence of local Bayesian inference using different data. A GP posterior is conditioned on the data available to the model; if neighboring agents condition on different datasets, then their posterior distributions need not agree. This is one of the central distinctions between centralized and decentralized mapping. In the centralized case, all measurements are pooled and there is a single posterior conditioned on $D = \cup_i D_i$. In the decentralized case, each posterior is conditioned on a partial view of the field. Thus, disagreement over O_{ij} can be interpreted as a symptom of missing cross-agent information.

Several existing approaches address this issue indirectly. Decentralized data fusion methods attempt to approximate the predictions of a centralized GP model by distributing computation across mobile sensors [6]. Expert aggregation methods combine local GP experts to form a joint prediction at a query location [9, 16, 19]. Federated GP training methods encourage agents to agree on shared hyperparameters or global model structure without sharing raw data [18, 29]. Distributed sparse GP methods use inducing variables or pseudo-representations to reduce the cost of sharing information [23, 31, 36, 29]. These methods provide important tools for scalable GP inference, but they do not always focus explicitly on the overlap region between neighboring local maps as the primary object of communication. Section 3.2 provide a comprehensive comparison detailing how our method addresses the limitations of these approaches.

The overlap region provides a useful opportunity because it is both local and informative. It is local because only neighboring agents need to coordinate over O_{ij}, avoiding global communication. It is informative because disagreement in this region indicates that the agents have complementary information about the same part of the field. If posterior of agent i is confident in a portion of the overlap where posterior of agent j is uncertain, then information from agent i may be valuable to agent j. Conversely, if both agents are uncertain or if their predictions already agree, communication may be less useful. This suggests that posterior disagreement can be used not only

as a diagnostic quantity but also as a guide for targeted information exchange.

This thesis uses overlapping local domains as the interface for decentralized cooperation. Instead of transmitting full datasets or requiring global aggregation, neighboring agents whose local domains overlap exchange compact, edge-specific inducing-point packets. Each packet is constructed from the local Gaussian-process model of sender, which is based on measurements collected over the covered region of sender, rather than being restricted to the overlapping area alone. The receiver then evaluates candidate packets and retains selected packet information as fictitious measurements in its local GP model. In this way, overlap between neighboring domains provides a mechanism for identifying useful communication links, while the exchanged packets allow local models to improve without requiring full data sharing.

This perspective differs from standard global sparse GP approximations because the inducing points are not selected only to approximate a single global GP. Instead, they are selected in an edge-specific way for communication between a sender and a receiver. It also differs from expert aggregation methods because the receiver does not merely combine final predictions from multiple GP experts at query time; instead, it incorporates selected information into its own local model so that the predictions can benefit from previous exchanges. This retained-information mechanism is important in limited-pass sensing scenarios because each communication event can have a cumulative effect on the local posterior.

The overlap-based view also provides a clear basis for comparison in Chapter 3. A local-only GP baseline represents the case where no cross-agent information is used, so posterior disagreement in overlap regions is not corrected. A centralized GP represents the idealized case where all data are available, reducing disagreement but requiring full communication. Decentralized data fusion methods approximate centralized prediction through support sets and distributed computation. Expert aggregation methods combine local predictions but may suffer from overconfidence, conservative uncertainty, or dependence on aggregation assumptions. Federated GP methods coordinate model parameters but may not directly exchange overlap-specific field information. The proposed method can be evaluated against these alternatives by measuring not only global

prediction accuracy, such as RMSE, but also probabilistic calibration, such as NLPD, and local consistency over overlapping regions.

Thus, overlapping domains are not merely a geometric detail of decentralized mapping. They define the interface where neighboring local models meet, disagree, and potentially improve each other. By treating posterior disagreement as useful information rather than only as an inconsistency, this thesis develops a decentralized GP mapping approach that uses compact, targeted communication to improve local estimates while preserving the scalability of multi-agent operation.

1.5 Research Problem

This thesis studies decentralized scalar-field mapping over a compact domain $\mathcal{X} \subset \mathbb{R}^n$, where a network of mobile sensing agents estimates an unknown field $f : \mathcal{X} \to \mathbb{R}$ from noisy local measurements. Each agent i is assigned to a local subdomain $\mathcal{X}^i$, maintains a local dataset $\mathcal{D}_t^i$, and builds a local Gaussian-process model using the information available to it. Since the agents do not have access to the full network dataset, the main challenge is to improve local and network-level prediction accuracy without relying on centralized computation or full data exchange.

For two agents i and j, the overlap region is defined as

$$O^{ij} := \mathcal{X}^i \cap \mathcal{X}^j .$$

In this shared region, both agents are estimating the same physical field, but their local GP posteriors may be different because they are conditioned on different datasets. Agent i's posterior mean and variance over the overlap are shaped by $\mathcal{D}_t^i$, while agent j's posterior is shaped by $\mathcal{D}_t^j$. As a result, the two agents may disagree about the value of the field or the uncertainty associated with that value. This disagreement is not simply an error; it is useful information. It indicates that one agent may have local information that can improve the estimate of another agent.

The research problem is therefore to design a decentralized information-exchange mechanism that uses overlap-based disagreement between neighboring local GP models to improve the predictive accuracy of the network, without requiring full data sharing or centralized computation. In this thesis, each agent maintains an augmented information set consisting of its own local

measurements and selected fictitious measurements received from neighboring agents. Each in-neighbor generates a library of candidate packets using its own local data and local GP model, together with the overlap geometry between the sender and receiver. A packet has the form

$$\mathbf{u} = (u, m, s),$$

where u is a location, m is the sender's posterior mean at that location, and s is the sender's posterior variance at that location. The receiver then applies a local selection rule to decide which candidate packet should be assimilated into its own GP model.

The key question studied in this thesis is whether the disagreement between local GP models on overlapping subdomains can be used as a decentralized metric for deciding which compact packet should be transmitted and retained. More specifically, the goal is to determine whether a receiver-side selection rule, based on the uncertainty and consistency of candidate inducing-point packets, can improve local and network-level prediction performance compared with a local-only GP baseline. This question is important because it connects three competing requirements: the agents should improve their estimates by using information from neighbors, the communication should remain compact, and the computation should remain local.

The research problem can be summarized as follows: given a network of mobile sensing agents with local datasets, overlapping subdomains, and limited communication, design sender-side packet-generation rules and receiver-side packet-selection rules that allow agents to improve their local GP estimates through compact fictitious-measurement assimilation. The desired approach should preserve the decentralized structure of the network, avoid transmission of full local datasets, and use overlap-based posterior disagreement as a practical signal for deciding which neighbor information is most useful.

1.6 Organization

The remainder of this thesis is organized as follows. Chapter 2 presents the main technical contribution of the thesis in manuscript form. This chapter develops the decentralized Gaussian-process mapping framework used in this work. It first formulates the scalar-field mapping

problem for a network of sensing agents with local subdomains, local measurements, and overlapping regions. Then, it introduces the Gaussian-process models used by each agent to estimate the unknown field from noisy measurements. The chapter also presents the sparse inducing-point representation used to construct compact communication packets, the greedy selection procedure used to generate candidate packets, and the receiver-side selection rule used to decide which packet should be assimilated as a fictitious measurement. Through this structure, Chapter 2 shows how local GP models can be improved through compact neighbor-to-neighbor communication without requiring full data sharing or centralized computation.

Chapter 2 also includes the simulation study used to evaluate the proposed method. The simulations compare the proposed decentralized packet-based approach with a local-only Gaussian-process baseline. The comparison is based on prediction accuracy and uncertainty quality, using metrics such as root-mean-square error and negative log predictive density. These results are used to examine whether selected fictitious measurements from neighboring agents can improve local and network-level scalar-field estimates.

Chapter 3 concludes the thesis. It first summarizes the overall problem, the proposed decentralized GP mapping approach, and the main results obtained from the simulation study. It then presents the major findings of the thesis and discusses the main contributions of the proposed method in relation to existing approaches. Chapter 3 also identifies the limitations of the current framework, including assumptions related to communication, overlap geometry, packet selection, and simulation-based validation. Finally, the chapter outlines possible directions for future work, such as improving packet-selection rules, extending the approach to more complex communication settings, considering time-varying fields, and evaluating the method in more realistic robotic sensing scenarios. The thesis ends with a final conclusion that summarizes the significance of compact information exchange for decentralized scalar-field mapping.

CHAPTER 2
DECENTRALIZED SCALAR FIELD MAPPING USING GAUSSIAN PROCESSES

2.1 Introduction

This chapter presents the decentralized Gaussian-process mapping framework used in this thesis. The method considers a network of agents that collect local measurements, maintain local GP models, and exchange compact information packets with neighboring agents. The main objective is to improve local and network-level prediction without constructing a centralized GP or sharing full local datasets. The chapter first formulates the decentralized scalar-field mapping problem, then develops the local GP model, the edge-specific inducing-point packet-generation rule, and the receiver-side packet-selection rule. The chapter concludes with a simulation study comparing the proposed packet-based method with a local-only GP baseline.

2.2 Problem Formulation

Let $\mathcal{X} \subset \mathbb{R}^n$ be a compact domain, and let $f : \mathcal{X} \to \mathbb{R}$ denote the unknown scalar field to be estimated. A network of N mobile sensing agents, indexed by $\mathcal{V} = \{1, \ldots, N\}$, communicates over a connected directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of communication links. For each agent $i \in \mathcal{V}$, the out-neighbor and in-neighbor sets are defined as

$$\mathcal{N}_i^{\text{out}} := \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}, \qquad \mathcal{N}_i^{\text{in}} := \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}\}.$$

Each agent i is assigned a local sensing subdomain $\mathcal{X}^i \subseteq \mathcal{X}$, and the collection of local subdomains covers the full domain,

$$\mathcal{X} \subseteq \bigcup_{i \in \mathcal{V}} \mathcal{X}^i.$$

For any pair of agents $i, j \in \mathcal{V}$, the overlap region is defined as

$$\mathcal{O}^{ij} := \mathcal{X}^i \cap \mathcal{X}^j.$$

∗ This chapter is based on the manuscript "Decentralized Scalar Field Mapping using Gaussian Process" by Hossein Papi[1], Muzaffar Qureshi[1], Kyle Volle[2], and Rushikesh Kamalapurkar[1]. [1]Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida, USA; [2]Torch Technologies, Shalimar, Florida, USA.

These overlap regions are important because they identify portions of the domain where neighboring agents may estimate the same physical field using different local datasets.

The agents operate on a synchronized measurement clock indexed by $t \in \mathbb{N}$. Let $\mathcal{D}_t^i \subseteq \mathcal{X}^i \times \mathbb{R}$ denote the set of measurements collected by agent i up to time t. Each measurement pair $(x_\tau^i, y_\tau^i) \in \mathcal{D}_t^i$ satisfies

$$y_\tau^i = f(x_\tau^i) + \epsilon_\tau^i,$$

where $x_\tau^i \in \mathcal{X}^i$ and $\epsilon_\tau^i \sim \mathcal{N}(0, (\sigma^i)^2)$. The measurement noises is assumed to be independent across agents and i.i.d. over time for each fixed agent.

In addition to its own local measurements, agent i may receive candidate data packets $\mathcal{C}_t^{j\to i}$ from its in-neighbors. Each in-neighbor $j \in \mathcal{N}_i^{\text{in}}$ generates candidate packets using its local dataset $\mathcal{D}_t^j$, local GP model, and the overlap geometry $\mathcal{O}^{ji}$ associated with receiver i, given by

$$\mathcal{C}_t^{j\to i} = \Phi_t^{j\to i}\left(\mathcal{D}_t^j, \mathcal{O}^{ji}\right),$$

where $\Phi_t^{j\to i}$ denotes the sender-side packet-generation rule. Each packet has the form

$$\mathbf{u} = (u, m, s) \in \mathcal{U} := \mathcal{X} \times \mathbb{R} \times \mathbb{R}_{\geq 0}, \tag{2-1}$$

where u is a location, m is the posterior of sender mean at u, and s is the posterior of sender variance at u. The pooled candidate packet library available to agent i is

$$\bar{\mathcal{C}}_t^i := \bigcup_{j\in\mathcal{N}_i^{\text{in}}} \mathcal{C}_t^{j\to i}.$$

Agent i applies a local selection rule

$$\psi_t^i : \mathbb{T}_t^i \to \mathcal{U}$$

which selects one packet from the pooled candidate library for assimilation into its local GP model.

Let

$$\mathcal{M}_t^i := \left\{ \left(u_{\ell,t}^i, m_{\ell,t}^i, r_{\ell,t}^i \right) \right\}_{\ell=1}^{M_t^i} \subseteq \mathcal{X} \times \mathbb{R} \times \mathbb{R}_{>0} \tag{2-2}$$

denote the set of packets selected for assimilation by agent i up to time t. These selected packets are treated as fictitious measurements. Defining the augmented information set

$$\mathcal{A}_t^i := \left(\mathcal{D}_t^i, \mathcal{M}_t^i \right),$$

the local GP maintained by agent i is conditioned on both its local measurements and selected fictitious measurements:

$$f \mid \mathcal{A}_t^i \sim \mathcal{GP}\left(\mu_t^i(\cdot), \Sigma_t^i(\cdot,\cdot) \right), \tag{2-3}$$

where μ_t^i and Σ_t^i denote the local posterior mean and covariance functions, respectively.

The objective of this chapter is to test whether discrepancies between neighboring local GP models over overlapping subdomains can be used as a decentralized metric for improving network-level predictive performance. To this end, the chapter develops packet-generation and packet-selection rules that use overlap-based consistency to guide which compact information should be transmitted and assimilated. The following sections first describe the local GP models, then introduce sparse inducing-point approximations for reduced-order packet generation, and finally present the receiver-side selection rule used to choose which packets are retained as fictitious measurements.

2.3 Methodology

This section illustrates the local GP models maintained by each agent from its augmented information set. Later Section 2.4 describes the optimal selection of inducing points to construct targeted communication packets for neighboring agents.

2.3.1 Gaussian Process Regression

We first make precise the local GP model that each agent $i \in \mathcal{V}$ generates from its augmented information set $\mathcal{A}_t^i = (\mathcal{D}_t^i, \mathcal{M}_t^i)$. This GP model is the object that the transmission rule $\Phi_t^{j\to i}$ and the selection rule ψ_t^i are ultimately intended to improve. The raw dataset $\mathcal{D}_t^i$ contains locally

collected measurements, while $\mathcal{M}_t^i$ contains packets selected for assimilation from received packet libraries $\bar{C}_t^i$. Each agent i employs a squared exponential covariance kernel of the form

$$k^i(x, x') = \alpha^i \exp\left(-\frac{\|x - x'\|^2}{2(\ell^i)^2}\right), \qquad x, x' \in \mathcal{X}, \tag{2-4}$$

where $\alpha^i > 0$ and $\ell^i > 0$ denote the signal variance and length-scale hyperparameters of agent i, respectively.

Let $S_t^i := \left|\mathcal{D}_t^i\right|$, $M_t^i := \left|\mathcal{M}_t^i\right|$, and write the local measurements of agent i at clock step t as $\mathcal{D}_t^i = \left\{\left(x_{j,t}^i, y_{j,t}^i\right)\right\}_{j=1}^{S_t^i}$, where $x_{j,t}^i \in \mathcal{X}^i$ and $y_{j,t}^i \in \mathbb{R}$. Likewise, write the packets selected for assimilation as (2-2) where $u_{\ell,t}^i \in \mathcal{X}$, $m_{\ell,t}^i \in \mathbb{R}$, and $r_{\ell,t}^i > 0$ are the location, mean and variances estimates associated with the selected packet, respectively. For compactness of notation, define $X_{\text{raw},t}^i := \left[x_{1,t}^i, \ldots, x_{S_t^i,t}^i\right]^\top$, $Y_{\text{raw},t}^i := \left[y_{1,t}^i, \ldots, y_{S_t^i,t}^i\right]^\top$, $U_t^i := \left[u_{1,t}^i, \ldots, u_{L_t^i,t}^i\right]^\top$, $M_t^{i,\text{fic}} := \left[m_{1,t}^i, \ldots, m_{L_t^i,t}^i\right]^\top$, and the augmented input and output arrays $X_t^i := \begin{bmatrix} X_{\text{raw},t}^i \\ U_t^i \end{bmatrix}$, $Y_t^i := \begin{bmatrix} Y_{\text{raw},t}^i \\ M_t^{i,\text{fic}} \end{bmatrix}$.

Finally, define the associated diagonal noise matrix

$$R_t^i := \operatorname{diag}\left((\sigma^i)^2, \ldots, (\sigma^i)^2, r_{1,t}^i, \ldots, r_{L_t^i,t}^i\right). \tag{2-5}$$

Conditioned on $\mathcal{A}_t^i = (\mathcal{D}_t^i, \mathcal{M}_t^i)$, the local GP posterior of agent i at a test point $x^* \in \mathcal{X}^i$ is given by $f(x^*) \mid \mathcal{A}_t^i \sim \mathcal{N}\left(\mu_t^i(x^*), \left(\sigma_t^i(x^*)\right)^2\right)$, where $\mu_t^i(x^*)$ and $\sigma_t^i(x^*)$ denote the posterior mean and posterior standard deviation of agent i, respectively.

Given the augmented information set $\mathcal{A}_t^i$, define the kernel vector $\mathbf{k}_t^i(x^*) \in \mathbb{R}^{S_t^i + L_t^i}$ and kernel matrix $K_t^i \in \mathbb{R}^{(S_t^i + L_t^i) \times (S_t^i + L_t^i)}$ by $\mathbf{k}_t^i(x^*) := \left[k^i(x^*, X_t^i(1,:)), \ldots, k^i(x^*, X_t^i(S_t^i + L_t^i, :))\right]^\top$, $\left[K_t^i\right]_{j,\ell} := k^i(X_t^i(j,:), X_t^i(\ell,:))$, $j, \ell = 1, \ldots, S_t^i + L_t^i$.

Using $\mathbf{k}_t^i(x^*)$, K_t^i, and the noise matrix R_t^i in (2-5), the posterior mean and variance of agent i are computed as

$$\mu_t^i(x^*) = \mathbf{k}_t^i(x^*)^\top \left(K_t^i + R_t^i\right)^{-1} Y_t^i, \tag{2-6}$$

and

$$\left(\sigma_t^i(x^*)\right)^2 = k^i(x^*, x^*) - \mathbf{k}_t^i(x^*)^\top \left(K_t^i + R_t^i\right)^{-1} \mathbf{k}_t^i(x^*), \tag{2-7}$$

where the first S_t^i diagonal entries of R_t^i correspond to local sensor-noise variance, and the remaining M_t^i diagonal entries correspond to the retained fictitious measurements variances associated with previously selected packets.

Note that the computational cost of the GP prediction grows cubically with the total size of the augmented information set. In particular, since agent i conditions on D_t^i raw measurements and M_t^i retained fictitious measurements, each evaluation of the posterior mean requires $O\big((D_t^i + M_t^i)^3\big)$ time, and each evalaution of the posterior variance requires $O\big((D_t^i + M_t^i)^2\big)$ time. To reduce this burden while preserving the agent-specific local GP structure, the next subsection introduces an inducing-point approximation for each agent.

2.3.2 Inducing Points in Gaussian Process

To obtain a scalable local approximation of a local GP, fix an arbitrary agent $i \in \mathcal{V}$ and clock step $t \in \mathbb{N}$. Recall from Section 2.3.1 that the local posterior of agent i is conditioned on the augmented information set $\mathcal{A}_t^i = (\mathcal{D}_t^i, \mathcal{M}_t^i)$, whose total size is $D_t^i + M_t^i$. To generate a reduced order representation of its GP, agent i employs a set of p_t^i inducing points, with $p_t^i \ll D_t^i + M_t^i$ [27].

Specifically, let

$$\bar{X}_t^i := \left[\bar{x}_{1,t}^i, \ldots, \bar{x}_{p_t^i,t}^i\right]^\top \subset \mathcal{X}^i, \tag{2-8}$$

denote the inducing-point set of agent i at clock step t. Using the kernel $k^i(\cdot,\cdot)$, define
$\left[K_{pp,t}^i\right]_{r,s} := k^i(\bar{x}_{r,t}^i, \bar{x}_{s,t}^i),\ r, s = 1, \ldots, p_t^i$, and
$\left[K_{Xp,t}^i\right]_{j,r} := k^i\left(X_t^i(j,:), \bar{x}_{r,t}^i\right), j = 1, \ldots, M_t^i + L_t^i,\ r = 1, \ldots, p_t^i$, where $X_t^i(j,:)$ denotes the j-th input location in the augmented input array X_t^i, while $\bar{x}_{r,t}^i$ denotes the r-th inducing point selected for the sparse approximation.

The corresponding low-rank approximation of the training covariance is

$$K_t^i \approx \tilde{K}_t^i := K_{Xp,t}^i \left(K_{pp,t}^i\right)^{-1} \left(K_{Xp,t}^i\right)^\top. \tag{2-9}$$

To equate the exact variance, we use the diagonal correction

$$\Sigma_t^{i,(p)} := K_{Xp,t}^i \left(K_{pp,t}^i\right)^{-1} \left(K_{Xp,t}^i\right)^\top + \Lambda_t^{i,(p)} + R_t^i, \tag{2-10}$$

where $\Lambda_t^{i,(p)} := \mathrm{Diag}\left\{K_t^i - K_{Xp,t}^i \left(K_{pp,t}^i\right)^{-1} \left(K_{Xp,t}^i\right)^\top\right\}$. For a test point $x^* \in \mathcal{X}^i$, define $k_{p,t}^i(x^*) := \left[k^i(\bar{x}_{1,t}^i, x^*), \ldots, k^i(\bar{x}_{p_t^i,t}^i, x^*)\right]^\top$, $\Omega_t^i := \Lambda_t^{i,(p)} + R_t^i$, $\Gamma_t^i := \left(\Omega_t^i\right)^{-1} K_{Xp,t}^i$, and $Q_t^i := K_{pp,t}^i + \left(K_{Xp,t}^i\right)^\top \left(\Omega_t^i\right)^{-1} K_{Xp,t}^i$. The resulting sparse GP predictive mean and variance are

$$\mu_t^{i,(p)}(x^*) = \left(k_{p,t}^i(x^*)\right)^\top \left(Q_t^i\right)^{-1} \left(\Gamma_t^i\right)^\top Y_t^i, \tag{2-11}$$

$$\left(\sigma_t^{i,(p)}(x^*)\right)^2 = k^i(x^*, x^*) - \left(k_{p,t}^i(x^*)\right)^\top \left[\left(K_{pp,t}^i\right)^{-1} - \left(Q_t^i\right)^{-1}\right] k_{p,t}^i(x^*). \tag{2-12}$$

The inducing-point construction reduces the dominant local training cost from $O\left((D_t^i + M_t^i)^3\right)$ to $O\left((D_t^i + M_t^i)(p_t^i)^2 + (p_t^i)^3\right)$, and the storage cost from $O\left((D_t^i + M_t^i)^2\right)$ to $O\left((D_t^i + M_t^i)p_t^i + (p_t^i)^2\right)$.

The computational burden in decentralized GP inference has two components: the *local cost* incurred by each agent for training, storage, and prediction as its local augmented dataset evolves, and the *communication cost* required to transmit candidate packets between agents. Because received information increases the computational load of the receiving agent and incurs a high communication cost, direct exchange of exact GP representations is not scalable.

In this thesis, we adopt the idea of inducing points to facilitate the design of the transmission rule $\Phi_t^{j\to i}$. Commensurate with the central hypothesis of the chapter, the transmission rules are designed to reduce the mismatch between the local GPs on the overlapping subdomains. To that end, each sender selects a small set of inducing points for a specific receiver so that the transmitted packets are the most informative over their overlap region O^{ji}. In the next subsection, an edge-dependent inducing-point selection rule is developed for communication-efficient inter-agent overlap consistency.

2.4 Domain Targeted Inducing-Point Selection

In Section 2.2, the transmission rule $\Phi_t^{j\to i}$ was introduced abstractly as a map from agent j's retained information and the overlap geometry to a candidate set sent to agent i. Here, we realize that idea by constructing edge-specific inducing points that are selected not to summarize the entirety of agent j's subdomain, but to reduce predictive uncertainty over the portion of O^{ji} that is most relevant to agent i. Fix a communicating pair $(j,i) \in \mathcal{E}$ and a clock step $t \in \mathbb{N}$. Let $N_t^j := D_t^j + M_t^j$ denote the total size of the augmented information set of sender agent j. Let $p_t^{j\to i} \in \mathbb{N}$ denote the total number of edge-specific inducing points to be selected by sender agent j for receiver agent i at clock step t, and let $m \in \{0, 1, \dots, p_t^{j\to i} - 1\}$ denote the current greedy selection stage. The localized inducing-point method of [7] selects inducing points for a single predictive location by minimizing a weighted integrated prediction-error criterion. Here, that construction is extended from a single target location to a finite target set, so that agent j selects edge-specific inducing points to improve prediction collectively over the portion of the overlap region that is most relevant to agent i.

For notational convenience, write the squared exponential kernel of agent j as $k^j(x,x') = \nu^j \kappa^j(x,x')$, $\nu^j := (\alpha^j)^2$, $\kappa^j(x,x') := \exp\!\left(-\frac{\|x-x'\|^2}{\theta^j}\right)$, $\theta^j := 2(\ell^j)^2$, where α^j and ℓ^j are the kernel hyperparameters already introduced in Section 2.3.1.

Let $\mathcal{T}_t^{j\to i} = \left\{z_{1,t}^{j\to i}, \dots, z_{q_t^{j\to i},t}^{j\to i}\right\} \subseteq O^{ji}$ denote the discretized target evaluation set in the overlap region, where $q_t^{j\to i} \in \mathbb{N}$ is the number of target locations used for communication from agent j to agent i at clock step t. Let $\bar{X}_{m,t}^{j\to i} = \left[\bar{x}_{1,t}^{j\to i}, \dots, \bar{x}_{m,t}^{j\to i}\right]^\top \subseteq O^{ji}$ be the current ordered edge-specific inducing-point set after m selections, and let $\bar{x}_{m+1,t}^{j\to i} = \left[\bar{x}_{m+1,1,t}^{j\to i}, \dots, \bar{x}_{m+1,n,t}^{j\to i}\right]^\top \in O^{ji}$ denote a candidate next inducing point. The corresponding augmented inducing-point set is $\bar{X}_{m+1,t}^{j\to i} = \left[\bar{x}_{1,t}^{j\to i}, \dots, \bar{x}_{m,t}^{j\to i}, \bar{x}_{m+1,t}^{j\to i}\right]^\top$.

To evaluate the utility of $\bar{x}_{m+1,t}^{j\to i}$ over the full target set $\mathcal{T}_t^{j\to i}$, define the weighting function

$$\omega_{\mathcal{T}_t^{j\to i}}(\tilde{x}) = \sum_{b=1}^{q_t^{j\to i}} \alpha_{b,t}^{j\to i}\, \kappa^j\!\left(\tilde{x}, z_{b,t}^{j\to i}\right), \qquad \sum_{b=1}^{q_t^{j\to i}} \alpha_{b,t}^{j\to i} = 1, \tag{2-13}$$

where $\alpha_{b,t}^{j\to i} \geq 0$ specifies the relative importance of the target location $z_{b,t}^{j\to i}$.

Next, define the kernel vector $k_{m+1,t}^{j\to i}(x) := \left[k^j(x, \bar{x}_{1,t}^{j\to i}), \ldots, k^j(x, \bar{x}_{m+1,t}^{j\to i})\right]^\top \in \mathbb{R}^{m+1}$, the Gram matrix $\left[K_{m+1,t}^{j\to i}\right]_{r,s} := k^j(\bar{x}_{r,t}^{j\to i}, \bar{x}_{s,t}^{j\to i}), \qquad r, s = 1, \ldots, m+1$, and the cross-covariance matrix between the sender active inputs $X_t^j = \{x_{u,t}^j\}_{u=1}^{M_t^j}$ and the augmented edge-specific inducing set $\left[K_{X,m+1,t}^{j\to i}\right]_{u,r} := k^j(X_t^j(u,:), \bar{x}_{r,t}^{j\to i}), \quad u = 1, \ldots, N_t^j, r = 1, \ldots, m+1$.

Using the exact augmented training covariance matrix K_t^j from Section 2.3.1, define $\Lambda_{m+1,t}^{j\to i} := \mathrm{Diag}\left\{K_t^j - K_{X,m+1,t}^{j\to i}\left(K_{m+1,t}^{j\to i}\right)^{-1}\left(K_{X,m+1,t}^{j\to i}\right)^\top\right\}$, $\Omega_{m+1,t}^{j\to i} := \Lambda_{m+1,t}^{j\to i} + R_t^j$, and $Q_{m+1,t}^{j\to i} := K_{m+1,t}^{j\to i} + \left(K_{X,m+1,t}^{j\to i}\right)^\top \left(\Omega_{m+1,t}^{j\to i}\right)^{-1} K_{X,m+1,t}^{j\to i}$.

The sparse predictive variance of sender agent j, computed from the augmented information set $\mathcal{A}_t^j$ and the augmented edge-specific inducing set $\bar{X}_{m+1,t}^{j\to i}$, is then

$$\left(\sigma_{m+1,t}^{j\to i}(\tilde{x})\right)^2 = k^j(\tilde{x},\tilde{x}) - \left(k_{m+1,t}^{j\to i}(\tilde{x})\right)^\top \left[\left(K_{m+1,t}^{j\to i}\right)^{-1} - \left(Q_{m+1,t}^{j\to i}\right)^{-1}\right] k_{m+1,t}^{j\to i}(\tilde{x}). \tag{2-14}$$

We now define the batch-targeted inducing-point criterion

$$\mathrm{BTIP}_t^{j\to i,(m+1)}\left(\bar{x}_{m+1,t}^{j\to i}; \mathcal{T}_t^{j\to i}\right) = \int_{O^{ji}} \omega_{\mathcal{T}_t^{j\to i}}(\tilde{x}) \frac{\left(\sigma_{m+1,t}^{j\to i}(\tilde{x})\right)^2}{\nu^j} d\tilde{x}. \tag{2-15}$$

Since $k^j(\tilde{x},\tilde{x}) = \nu^j$, the normalization by ν^j removes the process scale from the variance term. Although (2-15) is induced by the predictive variance, we interpret it as a weighted spatially integrated surrogate of prediction error over the target region. Hence, minimizing this objective selects inducing points that reduce uncertainty near the prescribed target set $\mathcal{T}_t^{j\to i}$.

Substituting (2-13) into (2-15) gives

$$\mathrm{BTIP}_t^{j\to i,(m+1)}\left(\bar{x}_{m+1,t}^{j\to i}; \mathcal{T}_t^{j\to i}\right) = \sum_{b=1}^{q_t^{j\to i}} \alpha_{b,t}^{j\to i} \int_{O^{ji}} \kappa^j\left(\tilde{x}, z_{b,t}^{j\to i}\right) \frac{\left(\sigma_{m+1,t}^{j\to i}(\tilde{x})\right)^2}{\nu^j} d\tilde{x}. \tag{2-16}$$

For each target point $z \in O^{ji}$, define the matrix $W_{m+1,t}^{*,j\to i}(z) \in \mathbb{R}^{(m+1)\times(m+1)}$ whose (r,s)-entry

is

$$\left[W^{*,j\to i}_{m+1,t}(z)\right]_{r,s} := \tfrac{1}{\nu^j}\int_{O^{ji}} \kappa^j(\tilde{x},z)\,k^j(\tilde{x},\bar{x}^{j\to i}_{r,t})\,k^j(\tilde{x},\bar{x}^{j\to i}_{s,t})\,d\tilde{x}, \tag{2-17}$$

$r,s = 1,\ldots,m+1$. The corresponding batch-weighted matrix is
$W^{*,j\to i}_{m+1,t}(\mathcal{T}^{j\to i}_t) = \sum_{b=1}^{q^{j\to i}_t} \alpha^{j\to i}_{b,t}\, W^{*,j\to i}_{m+1,t}(z^{j\to i}_{b,t})$.

For the overlap region O^{ji}, the criterion in (2-16) admits the following closed-form representation. Although the original BTIP formulation is presented over rectangular domains, the packet-selection procedure in the proposed method is performed only over the overlap region O^{ji}. To preserve the analytical closed-form expression, we evaluate the criterion over the smallest axis-aligned rectangle that contains O^{ji}. Consequently, the rectangular bounds appear only in the computation of the integral constants, while the optimization itself is restricted to candidate inducing points satisfying $\bar{x} \in O^{ji}$.

$$\mathrm{BTIP}^{j\to i,(m+1)}_t\left(\bar{x}^{j\to i}_{m+1,t};\mathcal{T}^{j\to i}_t\right) = C_{\mathcal{T}^{j\to i}_t} - \mathrm{tr}\left[\left((K^{j\to i}_{m+1,t})^{-1} - (Q^{j\to i}_{m+1,t})^{-1}\right)W^{*,j\to i}_{m+1,t}(\mathcal{T}^{j\to i}_t)\right], \tag{2-18}$$

where

$$C_{\mathcal{T}^{j\to i}_t} = \sum_{b=1}^{q^{j\to i}_t} \alpha^{j\to i}_{b,t}\frac{\sqrt{\pi\theta^j}}{2}\cdot\prod_{\ell=1}^{n}\left[\mathrm{erf}\left(\frac{z^{j\to i}_{b,\ell,t}-a^{j\to i}_\ell}{\sqrt{\theta^j}}\right) - \mathrm{erf}\left(\frac{z^{j\to i}_{b,\ell,t}-b^{j\to i}_\ell}{\sqrt{\theta^j}}\right)\right]. \tag{2-19}$$

where $a^{j\to i}_\ell$ and $b^{j\to i}_\ell$ denote the lower and upper bounds, respectively, of the smallest axis-aligned rectangle enclosing the overlap region O^{ji} along dimension ℓ. These bounds are introduced for evaluating the closed-form integral in (2-16); the inducing-point optimization remains constrained to the overlap region O^{ji}. The next inducing point is selected greedily as
$\bar{x}^{j\to i,\star}_{m+1,t} = \arg\min_{\bar{x}\in O^{ji}} \mathrm{BTIP}^{j\to i,(m+1)}_t\left(\bar{x};\mathcal{T}^{j\to i}_t\right)$, and the edge-specific inducing-point set is updated according to $\bar{X}^{j\to i}_{m+1,t} = \left[\bar{X}^{j\to i}_{m,t}, \bar{x}^{j\to i,\star}_{m+1,t}\right]$.

Thus, inducing points are selected sequentially to improve prediction over the target set $\mathcal{T}^{j\to i}_t$.

For gradient-based optimization, the derivative of the objective with respect to coordinate

$\ell \in \{1, \dots, n\}$ of the candidate point $\bar{x}_{m+1,t}^{j\to i}$ is

$$\frac{\partial}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}} \mathrm{BTIP}_t^{j\to i,(m+1)}\left(\bar{x}_{m+1,t}^{j\to i}; \mathcal{T}_t^{j\to i}\right) = - \operatorname{tr}\left[\left(\frac{\partial (K_{m+1,t}^{j\to i})^{-1}}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}} - \frac{\partial (Q_{m+1,t}^{j\to i})^{-1}}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}}\right) W_{m+1,t}^{*,j\to i}(\mathcal{T}_t^{j\to i})\right] - \operatorname{tr}\left[\left((K_{m+1,t}^{j\to i})^{-1} - (Q_{m+1,t}^{j\to i})^{-1}\right) \frac{\partial W_{m+1,t}^{*,j\to i}(\mathcal{T}_t^{j\to i})}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}}\right], \quad (2\text{-}20)$$

where

$$\frac{\partial W_{m+1,t}^{*,j\to i}(\mathcal{T}_t^{j\to i})}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}} = \sum_{b=1}^{q_t^{j\to i}} \alpha_{b,t}^{j\to i} \frac{\partial W_{m+1,t}^{*,j\to i}(z_{b,t}^{j\to i})}{\partial \bar{x}_{m+1,\ell,t}^{j\to i}}. \quad (2\text{-}21)$$

A practical initialization is given by the weighted centroid

$$z_{c,t}^{j\to i} = \sum_{b=1}^{q_t^{j\to i}} \alpha_{b,t}^{j\to i}\, z_{b,t}^{j\to i}, \qquad \bar{x}_{1,t}^{j\to i} = z_{c,t}^{j\to i}. \quad (2\text{-}22)$$

After the edge-specific inducing-point set $\bar{X}_t^{j\to i,\star} := \left[\bar{x}_{1,t}^{j\to i,\star}, \dots, \bar{x}_{p_t^{j\to i},t}^{j\to i,\star}\right]^\top$ has been selected, sender agent j constructs the communicated packet library for receiver agent i by evaluating its local posterior at those locations. Specifically, the transmitted packet is

$$C_t^{j\to i} = \Phi_t^{j\to i}\left(\mathcal{D}_t^j, \mathcal{O}^{ji}\right) = \left\{\left(\bar{x}_{r,t}^{j\to i,\star}, \mu_t^j\left(\bar{x}_{r,t}^{j\to i,\star}\right), \Sigma_t^j\left(\bar{x}_{r,t}^{j\to i,\star}, \bar{x}_{r,t}^{j\to i,\star}\right)\right)\right\}_{r=1}^{p_t^{j\to i}} \subseteq \mathcal{U}. \quad (2\text{-}23)$$

Thus, sender agent j does not transmit the entire local dataset or its full posterior representation. Instead, it transmits only the edge-specific inducing locations selected by the batch-targeted rule together with the sender posterior mean and variance predictions at those locations. These data packets provide a novel compact communication interface tailored to the overlap region $\mathcal{O}^{ji}$, and they are the objects from which the receiver-side candidate library is formed in the next section.

2.5 Receiver-Side Candidate Inducing-Point Selection

At every measurement step t, the receiver agent $i \in \mathcal{V}$ has access to its local measurements $\mathcal{D}_t^i$ and the edge-specific packets $C_t^{j\to i} \subseteq \mathcal{U}$ received from all of its in-neighbors $j \in \mathcal{N}_i^{\text{in}}$. Given any candidate packet $\mathbf{u} = (u, m, s) \in \bar{C}_t^i$ as defined in (2-1), receiver agent i interprets (m, s) as a one-point fictitious measurement at location u. The corresponding updated posterior mean is

$$\mu_{i,t}^+(x;\mathbf{u}) = \mu_t^i(x) \ + \Sigma_t^i(x,u)\left(\Sigma_t^i(u,u)+s\right)^{-1} \times \left(m - \mu_t^i(u)\right), \tag{2-24}$$

and the corresponding updated posterior covariance is

$$\Sigma_{i,t}^+(x,x';\mathbf{u}) \ = \Sigma_t^i(x,x') \ - \Sigma_t^i(x,u)\left(\Sigma_t^i(u,u)+s\right)^{-1} \times \Sigma_t^i(u,x'). \tag{2-25}$$

The receiver agent i selects one packet from the pooled candidate library $\bar{C}_t^i$ at each t. For any candidate packet $\mathbf{u} \in \bar{C}_t^i$, define the local receiver cost

$$J_t^i(\mathbf{u}) := \| \sum_{j\in\mathcal{N}_i^{\text{in}}} \sum_{\mathbf{v}=(v,m_v,s_v)\in C_t^{j\to i}} \left(\alpha \left|\mu_{i,t}^+(v;\mathbf{u}) - m_v\right|^2 + \ \beta \left|\Sigma_{i,t}^+(v,v;\mathbf{u}) - s_v\right|^2\right)\| \tag{2-26}$$

where $\alpha \geq 0$ and $\beta \geq 0$ are weights for mean and variance consistency, respectively. $J_t^i(\mathbf{u})$ measures how well the receiver posterior obtained by assimilating packet $\mathbf{u}$ agrees, at all received packet locations, with the GP posterior of its in-neighbors. The receiver-side packet selection rule for agent i is therefore given as

$$\mathbf{u}_t^{i,\star} = \psi_t^i(\bar{C}_t^i) = \arg\min_{\mathbf{u}\in\bar{C}_t^i} J_t^i(\mathbf{u}). \tag{2-27}$$

After the minimizing packet

$$\mathbf{u}_t^{i,\star} = (u_t^{i,\star}, m_t^{i,\star}, s_t^{i,\star}) \tag{2-28}$$

is selected, receiver agent i treats it as a fictitious measurement with pseudo-noise variance $r_t^{i,\star} := s_t^{i,\star}$. The packets selected for assimilation are then updated according to $\mathcal{M}_{t+1}^i = \mathcal{M}_t^i \cup \left\{\left(u_t^{i,\star}, m_t^{i,\star}, r_t^{i,\star}\right)\right\}$. Thus, the selected packet is not used only for a temporary

posterior correction at time t, but is retained as a pseudo-measurement and incorporated into the augmented information set used to form the local GP posterior at subsequent measurement steps. Collecting the selected inducing points for all agents as

$$\mathbf{u}_t \coloneqq \left(\mathbf{u}_t^1, \ldots, \mathbf{u}_t^N\right) \in \bar{C}_t^1 \times \cdots \times \bar{C}_t^N. \tag{2-29}$$

and defining the network-level objective as

$$J_t(\mathbf{u}_t) \coloneqq \sum_{i=1}^{N} J_t^i(\mathbf{u}_t^i). \tag{2-30}$$

Hence, the network-level packet-selection problem is

$$\mathbf{u}_t^\star \in \arg \min_{\mathbf{u}_t \in \bar{C}_t^1 \times \cdots \times \bar{C}_t^N} J_t(\mathbf{u}_t). \tag{2-31}$$

Therefore, if each local problem has a unique minimizer, then the global problem also has a unique minimizer. In summary, the receiver-side packet-selection problem is a finite exact optimization that is globally defined but rigorously separable across receivers. Each receiver selects the packet that minimizes its local overlap-consistency cost, and the collection of these local minimizers simultaneously solves the network-level problem.

2.6 Simulation Results

The developed method was evaluated in a decentralized mapping simulation with $N = 4$ over the square domain $[-6, 6] \times [-6, 6]$. The scalar field was generated from a smooth nonlinear test function over a 41×41 evaluation grid. The four agent centers were placed at $(-3.0, -2.6)$, $(2.5, -2.8)$, $(3.0, 2.8)$, and $(-2.7, 3.0)$, and each agent was assigned a circular local domain of radius 4.4. At each clock step, every agent collected one local measurement, with a Gaussian noise standard deviation 0.08, over a total of 20 steps.

For local GP modeling, the agent-specific length-scales are chosen as $[0.85,\ 1.10,\ 1.35,\ 0.95]$, the signal scales as $[0.90,\ 1.15,\ 1.05,\ 0.80]$, and the noise standard

deviations as $[0.07,\ 0.09,\ 0.08,\ 0.10]$. At every step, each sender generated an edge-specific packet over each overlap region using 4 transmitted inducing points per directed edge. Each receiver then solved the exact finite-library packet-selection problem described in Section 2.5 and retained the selected packet as a fictitious measurement for subsequent updates. The retained pseudo-measurement variance was formed using the sender's predictive variance. In the receiver-side cost, the mean- and variance-consistency terms were weighted by $\alpha = 1.0$ and $\beta = 0.25$, respectively.

Figure 2-1 shows the true scalar field together with the four local agent domains. Figure 2-2 shows the final local GP posterior mean for each agent at the end of the simulation, along with that agent's measurement locations. Figure 2-3 shows the retained inducing-point packets for each agent, where the number next to each selected packet indicates the time step at which it was incorporated.

For an evaluation set $\{x_n\}_{n=1}^{N}$, with ground-truth field value $f(x_n)$, GP predictive mean $\mu(x_n)$, and predictive variance $\sigma^2(x_n)$, these metrics are defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\big(\mu(x_n) - f(x_n)\big)^2}, \tag{2-32}$$

$$\mathrm{NLPD} = \tfrac{1}{N}\textstyle\sum_{n=1}^{N}\left[\tfrac{1}{2}\log\big(2\pi(\sigma^2(x_n)+\sigma_\epsilon^2)\big) + \frac{\big(f(x_n)-\mu(x_n)\big)^2}{2(\sigma^2(x_n)+\sigma_\epsilon^2)}\right], \tag{2-33}$$

where σ_ϵ^2 denotes the observation-noise variance. RMSE measures the posterior mean accuracy, while NLPD evaluates probabilistic calibration by jointly penalizing prediction error and uncertainty. Figure 2-4 compares the proposed inducing-point assimilation strategy against a self-only baseline at the network level. In the self-only baseline, each agent trains its GP using only locally collected measurements. In contrast, the proposed method augments each local dataset with selected inducing-point packets received from neighboring agents. This comparison highlights the impact of inter-agent communication on predictive performance within each agent's local domain and in the overlap regions shared with other agents.

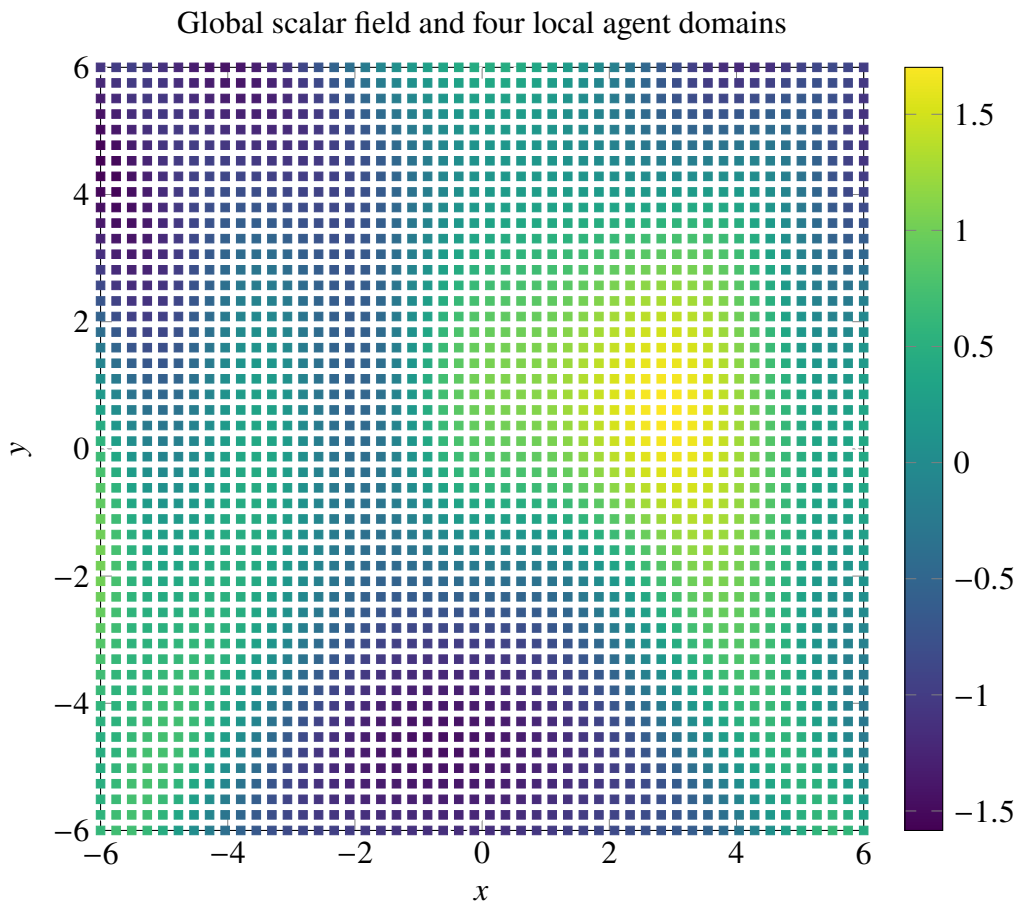

Figure 2-1. True scalar field, local agent domains, and measurement locations.

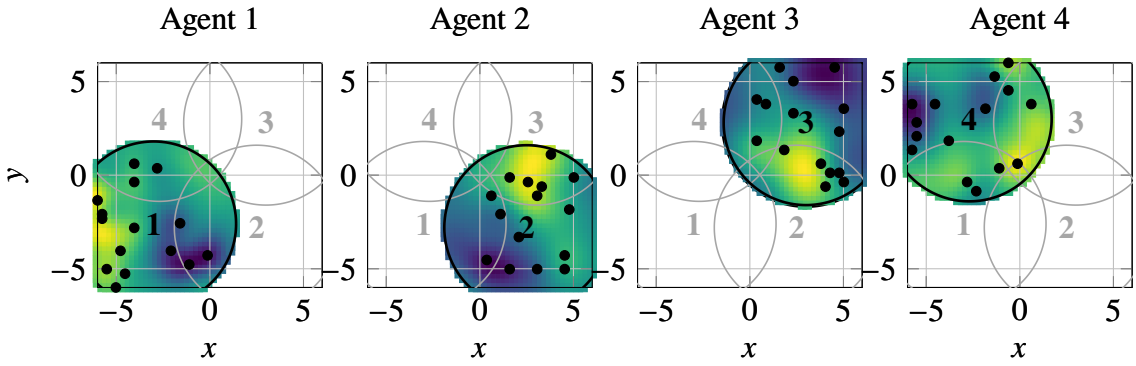

Figure 2-2. Local GP prediction mean is shown at the culmination of the episode. The individual agent measurement locations are shown as black dots.

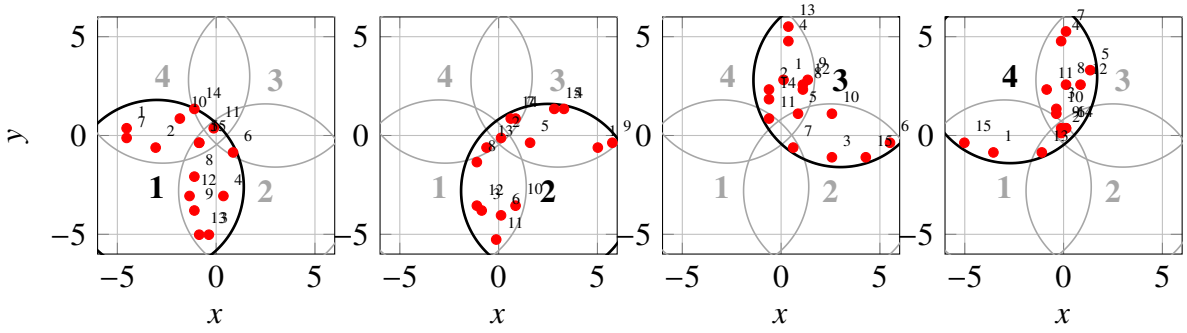

Figure 2-3. Retained selected inducing-points for each agent. The number next to each packet denotes the time step.

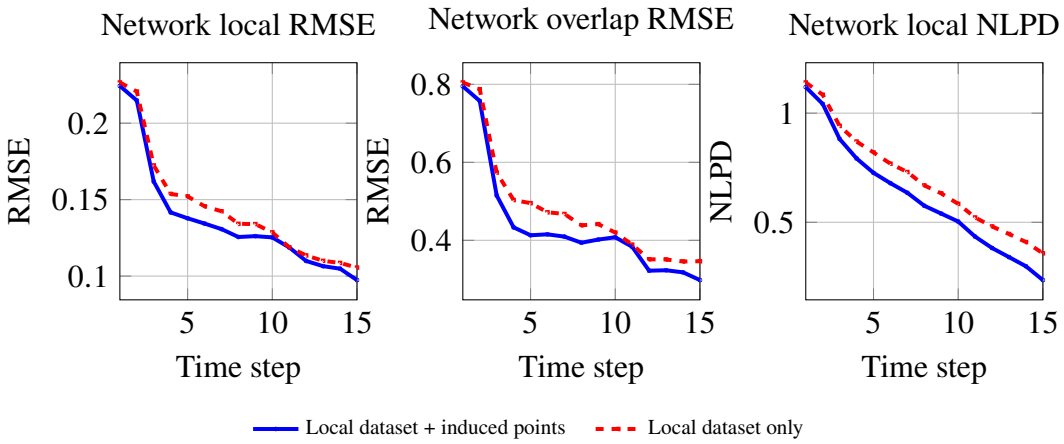

Figure 2-4. Network-level comparison between the shared-information method and the self-only baseline averaged across agents

.

2.7 Discussion

The results show that the developed method improves consistency over overlap regions while preserving accurate local GP predictions. Since each sender transmits only a compact overlap-targeted summary, the method remains decentralized and avoids sharing full local packets. A key feature is that communication has a cumulative effect. The selected packets are retained as fictitious measurements, so the local posteriors evolve using both newly collected measurements and previously accepted data packets. In comparison with GP models trained via local data sets, the developed algorithm performs better in the simulation studies, with results showing lower RMSE, along with reduced NLPD in the overlap regions while keeping the estimate errors preserved in the local map. This shows that agents improve their predictions and uncertainty estimates and reach better consensus than when operating in isolation.

The developed method currently assumes accurate transfer of information, but in practical scenarios, information sharing between agents can also introduce noise in the packets. This can degrade estimates, leading to reduced accuracy and increased error in the estimate compared to the local GP. Additionally, the framework relies on sufficient overlap between agents; with minimal overlap, the impact of information sharing diminishes and performance approaches that of isolated local models.

CHAPTER 3
CONCLUSIONS AND FUTURE WORK

3.1 Major Findings

The results of this thesis show that overlap between local sensing domains can be used as more than a geometric relationship between neighboring agents. In decentralized scalar-field mapping, overlapping regions provide a useful place to compare local Gaussian-process models that are built from different datasets. When two neighboring agents estimate the same portion of the field but obtain different posterior means or posterior variances, this disagreement reveals that the agents have different information about the environment. One major finding of this thesis is that this disagreement can be used as a practical decentralized signal for deciding which information should be exchanged between agents.

The results also show that compact packet-based communication can improve decentralized GP mapping without requiring transmission of full local datasets. Instead of transmitting all local measurements or constructing a centralized GP model, each sender communicates only a small set of inducing-point packets. Each packet contains a location, the posterior mean of sender, and the posterior variance of sender at that location. This representation is small enough to support decentralized communication, but still informative enough for a receiving agent to update its own local model. The simulation results indicate that selected packets can improve the posterior estimate of receiver, especially when the selected information comes from a neighboring agent with useful knowledge about a related region of the field.

Another important finding is that retaining selected packets as fictitious measurements gives communication a lasting effect. If a received packet were used only for a temporary prediction correction, then its influence would disappear after the current update. In the proposed framework, the selected packet is added to the augmented information set of receiver and is used in later GP updates. This means that information received from neighbors can continue to improve future predictions. This retained-information structure is especially important in limited-pass sensing scenarios, where agents may not have many opportunities to revisit the same locations or repeatedly exchange large amounts of data.

The simulation study also shows that the proposed approach can improve performance compared with a local-only GP baseline. In the local-only case, each agent trains its GP using only its own measurements, so the resulting estimate depends only on the local data collected inside that agent's subdomain. In contrast, the proposed method allows each agent to benefit from selected posterior information received from neighboring agents. The comparison using root-mean-square error and negative log predictive density shows that the packet-based method can improve both prediction accuracy and uncertainty quality. This is important because accurate mean prediction alone is not sufficient for GP-based mapping; the uncertainty estimate must also remain meaningful.

The results further indicate that local estimates can improve without requiring the network to form a single global GP model. The proposed method preserves local computation and local GP models while still allowing the network to improve through neighbor-to-neighbor communication. This is useful because many robotic sensing applications cannot rely on constant global communication or centralized processing. The results therefore support the main hypothesis of the thesis: overlap-based posterior discrepancy can guide compact information exchange and improve decentralized scalar-field mapping while maintaining the distributed structure of the multi-agent system.

3.2 Comparison with Existing Decentralized Approaches

The proposed method is a decentralized mapping approach in which each agent maintains its own local Gaussian-process model and exchanges selected packet information with neighboring agents. The method keeps inference local to each agent, while communication is used to improve local prediction quality rather than to construct a single shared posterior.

Compared with local-only GP mapping, the proposed method uses communication to reduce the information gap between neighboring agents. In the local-only case, each agent trains its GP only on its own measurements. This avoids communication cost, but it also ignores useful information collected by nearby agents. As a result, local estimates can remain inconsistent in regions where neighboring sensing domains overlap. The proposed approach keeps the decentralized structure of the local-only method, but allows each agent to improve its local model by

assimilating selected information from neighbors.

Compared with decentralized data-fusion methods such as D2FAS [6], the proposed method does not try to approximate a centralized GP through a common support set or global fusion structure. D2FAS distributes GP prediction and active sensing across mobile sensors, but it is still built around a distributed approximation of centralized data fusion. In this thesis, each agent maintains its own local GP model, and communication is limited to compact edge-specific packets generated by neighboring agents.

The proposed framework also differs from expert-aggregation and federated GP methods. Expert-aggregation approaches, such as generalized product-of-experts and distributed GP methods [4, 9, 20, 37], combine local expert predictions at query time to form a final predictive distribution. In contrast, this thesis does not combine final predictions from multiple experts. A receiver selects a candidate packet and retains it inside its own GP model, so the received information can influence future posterior estimates. Federated GP methods [17, 18] mainly focus on learning shared hyperparameters or common model structures across agents. The focus here is different: agents exchange compact information about the field itself, represented by inducing locations, sender posterior means, and sender posterior variances.

Compared with distributed sparse GP and distributed coverage-control methods [10, 22], the proposed method separates sender-side packet generation from receiver-side packet selection. The sender compresses its local posterior information into candidate inducing-point packets for a neighboring receiver, while the receiver decides which packet is most useful for its own local model. Once selected, the packet is stored as a fictitious measurement and continues to influence later GP updates. Overall, the proposed approach provides a fully decentralized compact-communication framework that avoids centralized inference and full raw-data transmission while improving local GP estimates through limited communication with neighboring agents.

3.3 Limitations

Although the simulation results support the proposed framework, several limitations should be considered. First, the current framework assumes reliable communication of the selected packet

information. In the proposed method, each packet contains a location, a posterior mean, and a posterior variance generated by the local GP model of sender. The receiver then uses this information as a fictitious measurement. If the packet is corrupted by communication noise, delay, or mismatch between the sender and receiver coordinate frames, then the receiver may assimilate information that is not fully consistent with the true field. This could increase local prediction error instead of reducing it. Therefore, the current approach is most suitable for settings where the transmitted packet information can be trusted or where communication errors are small.

The method also depends on the existence of useful overlap between neighboring local domains. The overlap region is the main place where neighboring agents can compare their local posterior behavior and identify useful information exchange. If two agents have very small overlap, or if their overlap contains little informative variation in the field, then the benefit of communication may be limited. In the extreme case where the local domains do not overlap, the proposed overlap-based consistency metric cannot be applied directly. This means that the method is naturally suited to decompositions where neighboring subdomains share enough common region to support meaningful comparison.

The current receiver-side selection rule also selects a limited amount of information at each measurement step. This makes the optimization simple and communication-efficient, but it may ignore other useful packets that could improve the local posterior if they were assimilated together. In some cases, selecting only one packet may be too conservative, especially when several neighbors provide complementary information. Extending the method to multi-packet selection could improve performance, but it would also increase the computational cost and require a more careful design of the selection criterion.

Finally, the present formulation assumes fixed GP hyperparameters and a relatively simple scalar-field model. In practice, the correct kernel parameters may not be known in advance and may vary across agents or environments. If the length-scale, signal variance, or noise variance are poorly chosen, the local posterior mean and uncertainty estimates may be inaccurate, which can affect both packet generation and packet selection. Similarly, real scalar fields may be time-varying,

discontinuous, anisotropic, or affected by external dynamics. These cases are beyond the current scope of the thesis, but they are important for future development of the framework.

3.4 Future Work

Several extensions would make the proposed decentralized GP mapping framework more suitable for practical robotic deployments. One important direction is to move from simulation-based validation to experimental validation with real robotic platforms. The current results show that compact packet exchange can improve decentralized scalar-field mapping in a controlled setting, but real systems introduce additional challenges such as localization error, communication delay, packet loss, sensor bias, actuator limits, and imperfect synchronization between agents. Testing the method on physical robots or high-fidelity robotic simulators would help evaluate how robust the proposed packet-generation and packet-selection rules are under practical operating conditions.

Another direction is to improve the receiver-side selection rule. In the current framework, each receiver selects one candidate packet at each measurement step and retains it as a fictitious measurement. This keeps the decision problem simple and communication-efficient, but it may not fully use the information available from multiple neighbors. Future work could therefore consider an information-theoretic selection rule that evaluates each candidate packet, or candidate subset, using the Kullback–Leibler divergence between the receiver's current GP posterior and the hypothetical posterior obtained after assimilating that subset. Interpreted as an expected weight of evidence, this divergence measures how strongly the selected packets distinguish the updated posterior from the receiver's current belief, providing a principled measure of their information value [38]. Because a large posterior change is not necessarily an accurate or beneficial change, the KL information gain should be combined with reliability constraints and penalties for packet redundancy, correlation, and overconfident uncertainty estimates.

A related extension is multi-packet selection, where a receiver chooses a small subset of packets instead of a single packet. This would allow the receiver to combine complementary information from different agents, especially when several neighbors observe different parts of the

field. However, this extension would require careful control of computational cost, packet redundancy, and possible overconfidence caused by assimilating correlated information.

The framework can also be extended by adding adaptive packet-management strategies. Since retained packets become part of the receiver's augmented information set, the local GP dataset grows over time. This is useful because communication has a cumulative effect, but it may become expensive during long missions. This could include pruning, merging, or reweighting methods for fictitious measurements so that agents keep the most informative packets while removing packets that become redundant or outdated. Such methods would help preserve the scalability of the approach as the mission length and number of agents increase.

Another important extension is adaptive hyperparameter learning. In this thesis, the GP kernel hyperparameters are treated as fixed during the simulation. In practical scalar-field mapping problems, the correct length-scale, signal variance, and noise variance may not be known ahead of time and may differ across agents. Future work can develop decentralized or federated hyperparameter-learning methods so that agents can update their local GP models as more measurements become available. This would make the framework more flexible when the field has different smoothness properties in different regions of the domain.

Future work should also consider more realistic communication and sensing models. The present method assumes that candidate packets are transmitted accurately and that the receiver can directly assimilate the selected packet as a fictitious measurement. In practice, communication may be noisy, delayed, intermittent, or affected by coordinate-frame mismatch between agents. A disturbance-tolerant version of the method could include confidence weights, packet validation tests, robust cost functions, or filtering methods that prevent unreliable packets from degrading the receiver's local GP model. This would be especially important for field deployments in cluttered, hazardous, or communication-limited environments.

Finally, the proposed framework can be extended to time-varying scalar fields and more complex multi-agent planning problems. Many real environmental fields, such as smoke, temperature, chemical concentration, and pollution, change over time. Extending the method to

spatiotemporal GPs would allow agents to model both spatial and temporal uncertainty. In addition, future work can connect the packet-selection framework with active sensing and motion planning, so that agents not only decide which information to exchange, but also where to move next. This would create a more complete decentralized mapping system in which sensing, communication, and planning are designed together.

LIST OF REFERENCES

[1] Ruzena Bajcsy, Yiannis Aloimonos, and John K. Tsotsos, *Revisiting active perception*, Autonomous Robots **42** (2018), no. 2, 177–196.

[2] Nils Bore, Ignacio Torroba, and John Folkesson, *Sparse gaussian process slam, storage and filtering for auv multibeam bathymetry*, 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), 2018, pp. 1–6.

[3] Eric Brochu, Vlad M. Cora, and Nando de Freitas, *A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning*, 2010.

[4] Yanshuai Cao and David J. Fleet, *Generalized product of experts for automatic and principled fusion of gaussian process predictions*, 2014.

[5] Andrea Carron, Marco Todescato, Ruggero Carli, Luca Schenato, and Gianluigi Pillonetto, *Multi-agents adaptive estimation and coverage control using gaussian regression*, Proceedings of the European Control Conference, ECC, 2015, pp. 2490–2495.

[6] Jie Chen, Kian Hsiang Low, Colin Keng-Yan Tan, Ali Oran, Patrick Jaillet, John M. Dolan, and Gaurav S. Sukhatme, *Decentralized data fusion and active sensing with mobile sensors for modeling and predicting spatiotemporal traffic phenomena*, Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence, UAI, AUAI Press, 2012, pp. 163–173.

[7] D. Austin Cole, Ryan Christianson, and Robert B. Gramacy, *Locally induced gaussian processes for large-scale simulation experiments*, Statistics and Computing **31** (2021), 33.

[8] R. Andres Cortez, Herbert G. Tanner, and Ron Lumia, *Distributed robotic radiation mapping*, Experimental Robotics (Oussama Khatib, Vijay Kumar, and George J. Pappas, eds.), Springer Tracts in Advanced Robotics, vol. 54, Springer, Berlin, Heidelberg, 2009, pp. 147–156.

[9] Marc Peter Deisenroth and Jun Wei Ng, *Distributed gaussian processes*, Proceedings of the 32nd International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 37, PMLR, 2015, pp. 1481–1490.

[10] James Di, Ehsan Zobeidi, Alec Koppel, and Nikolay Atanasov, *Distributed gaussian process mapping for robot teams with time-varying communication*, Proceedings of the American Control Conference, ACC, 2022, pp. 4840–4846.

[11] Christopher Diehl, Eduard Feicho, Alexander Schwambach, Thomas Dammeier, Eric Mares, and Torsten Bertram, *Radar-based dynamic occupancy grid mapping and object detection*, 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020, pp. 1–6.

[12] Robert B. Gramacy and Daniel W. Apley, *Local gaussian process approximation for large computer experiments*, Journal of Computational and Graphical Statistics **24** (2015), no. 2, 561–578.

[13] Shuo He, M. Tang, J. Fu, and J. Liang, *Distributed online sparse gaussian process regression for multi-agent coverage control*, Proceedings of the 42nd Chinese Control Conference, 2023, pp. 5464–5469.

[14] S. Sathiya Keerthi and Wei Chu, *A matching pursuit approach to sparse gaussian process regression*, Advances in Neural Information Processing Systems, vol. 18, 2005, pp. 643–650.

[15] Stephanie Kemna, John G. III Rogers, Carlos Nieto-Granda, Stuart Young, and Gaurav S. Sukhatme, *Multi-robot coordination through dynamic voronoi partitioning for informative adaptive sampling in communication-constrained environments*, Proceedings of the IEEE International Conference on Robotics and Automation, 2017, pp. 2124–2130.

[16] George P. Kontoudis and Daniel J. Stilwell, *Decentralized nested gaussian processes for multi-robot systems*, 2021 IEEE International Conference on Robotics and Automation, ICRA, 2021, pp. 8881–8887.

[17] _____, *Fully decentralized, scalable gaussian processes for multi-agent federated learning*, 2022.

[18] _____, *Scalable, federated gaussian process training for decentralized multi-agent systems*, IEEE Access **12** (2024), 77800–77815.

[19] _____, *Multi-agent federated learning using covariance-based nearest neighbor gaussian processes*, IEEE Transactions on Machine Learning in Communications and Networking **4** (2026), 115–138.

[20] Haitao Liu, Jianfei Cai, Yi Wang, and Yew-Soon Ong, *Generalized robust bayesian committee machine for large-scale gaussian process regression*, Proceedings of the 35th International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 80, PMLR, 2018, pp. 3131–3140.

[21] Ruochen Ouyang, Kian Hsiang Low, Jie Chen, and Patrick Jaillet, *Multi-robot active sensing of non-stationary gaussian process-based environmental phenomena*, Proceedings of the International Conference on Autonomous Agents and Multiagent Systems, 2014.

[22] Federico Pratissoli, Mattia Mantovani, Amanda Prorok, and Lorenzo Sabattini, *Distributed coverage control for time-varying spatial processes*, IEEE Transactions on Robotics **41** (2025), 1602–1617.

[23] Joaquin Quiñonero-Candela and Carl Edward Rasmussen, *A unifying view of sparse approximate gaussian process regression*, Journal of Machine Learning Research **6** (2005), 1939–1959.

[24] Muzaffar Qureshi, Tochukwu Elijah Ogri, Zachary I. Bell, and Rushikesh Kamalapurkar, *Scalar field mapping with adaptive high-intensity region avoidance*, Proceedings of the IEEE Conference on Control Technology and Applications, 2024, pp. 388–393.

[25] Muzaffar Qureshi, Tochukwu Elijah Ogri, H. Ramos, Zachary I. Bell, and Rushikesh Kamalapurkar, *Gaussian process-based scalar field estimation in gps-denied environments*, 2025.

[26] Prabhash Ragbir, Ajith Kaduwela, Xiaodong Lan, Adam Watts, and Zhaodan Kong, *A control-theoretic spatio-temporal model for wildfire smoke propagation using uav-based air pollutant measurements*, Drones **8** (2024), no. 5, 169.

[27] Carl Edward Rasmussen and Christopher K. I. Williams, *Gaussian processes for machine learning*, MIT Press, Cambridge, MA, 2006.

[28] Afsheen Sadaf, Julie Bruck, and Mojtaba Tahmasebi, *Coastal resilience through nature-based solutions – a gis suitability analysis model for living shorelines at aberdeen proving ground, maryland*, https://ssrn.com/abstract=4866580, 2025, SSRN preprint, not peer reviewed.

[29] Sanket A. Salunkhe and George P. Kontoudis, *Federated gaussian process learning via pseudo-representations for large-scale multi-robot systems*, 2026.

[30] Alexander J. Smola and Peter L. Bartlett, *Sparse greedy gaussian process regression*, Advances in Neural Information Processing Systems, vol. 13, 2001, pp. 619–625.

[31] Edward Snelson and Zoubin Ghahramani, *Sparse gaussian processes using pseudo-inputs*, Advances in Neural Information Processing Systems, vol. 18, 2006, pp. 1257–1264.

[32] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias Seeger, *Gaussian process optimization in the bandit setting: No regret and experimental design*, Proceedings of the 27th International Conference on Machine Learning, 2010.

[33] Varun Suryan and Pratap Tokekar, *Learning a spatial field in minimum time with a team of robots*, IEEE Transactions on Robotics **36** (2020), no. 5, 1562–1576.

[34] Mojtaba Tahmasebi, Julie Bruck, and Afsheen Sadaf, *Modeling future land cover and predicting carbon storage and sequestration in coastal areas: Integrating invest model and molusce plugin of qgis*, AGU Fall Meeting Abstracts, vol. 2024, December 2024, pp. GC21K–0019.

[35] Mojtaba Tahmasebi, Julie Bruck, Michael Volk, Emre Tepe, Abhinav Alakshendra, Afsheen Sadaf, and Jack A. Puleo, *Coastal carbon at risk: forecasting the impacts of sea-level rise on future land cover*, Frontiers in Ecology and Evolution **13** (2025), 1608422.

[36] Michalis K. Titsias, *Variational learning of inducing variables in sparse gaussian processes*, Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, vol. 5, PMLR, 2009, pp. 567–574.

[37] Volker Tresp, *A bayesian committee machine*, Neural Computation **12** (2000), no. 11, 2719–2741.

[38] Hongchuan Wei, Wenjie Lu, and Silvia Ferrari, *An information value function for nonparametric gaussian processes*, 2014, Presented at the NIPS Workshop on Bayesian Nonparametric Models for Reliable Planning and Decision-Making Under Uncertainty.

BIOGRAPHICAL SKETCH

Hossein Papi received his bachelor's degree in mechanical engineering in Isfahan, Iran before coming to the United States to pursue advanced graduate study and research in mechanical engineering. He received his Master of Science degree in mechanical engineering from the Department of Mechanical and Aerospace Engineering at the University of Florida. He completed his graduate studies under the supervision of Dr. Kamalapurkar in the Systems, Cognition, and Control Laboratory. His research focused on decentralized scalar field mapping using Gaussian processes. His academic interests include control systems, robotics, autonomous systems, and machine learning-based methods for dynamic systems.