UNIVERSITÀ
DEGLI STUDI
FIRENZE

Università degli Studi di Firenze
Dipartimento di Ingegneria dell'Informazione (DINFO)
Corso di Dottorato in Ingegneria dell'Informazione
Curriculum: Control, Optimization and Complex Systems

Maritime anomaly detection based on statistical methodologies: theory and applications

Candidate
Enrica d'Afflisio

Supervisors
Prof. Luigi Chisci

Prof. Giorgio Battistelli

Dr. Paolo Braca

PhD Coordinator
Prof. Fabio Schoen

ciclo XXXIV, 2018-2021

Università degli Studi di Firenze, Dipartimento di Ingegneria
dell'Informazione (DINFO).

A tutti quelli che restano pescetti in un mondo di squali

Acknowledgments

I would like to acknowledge the efforts, the input, the kindness and the patience of my supervisors, Prof. Luigi Chisci, Prof. Giorgio Battistelli, and Dr. Paolo Braca, and all my colleagues of the DKOE group who were always there for me when I was stuck. And I am always stuck.

In particular my thanks go to the loved ones. To Tommaso for always being my family. To Norma and Angeliki for keeping me from sinking. To my mum for having me in this crazy crazy world. To NBB for making this world even crazier and full of love.

To those who arrived and stayed. To those who always understood.

To ME.

Abstract

The proposed research aims at contributing to advances in the anomaly detection methodologies within the framework of maritime domain, in order to improve the ability to reveal, understand, anticipate and prevent illegitimate activities at sea. This work has been developed based on three fundamental tools: a prior information from a maritime traffic graph that can be derived from a route atlas or from historical data, the Ornstein-Uhlenbeck mean reverting stochastic process to model the vessel's dynamics in deep waters, and the complete or incomplete observation of the available data from heterogeneous sensor systems. Relying on the statistical hypothesis testing framework, the work treats the problem of detecting a vessel's anomalous deviations from the expected conditions in the presence of different levels of data unavailability. The problem is further complicated by the possible falsification of dynamic data self-reported by the vessel. A worst-case scenario in terms of detection capability is finally tackled by proposing an optimization methodology to make the trajectory of a malicious vessel as stealth as possible. The effectiveness of the proposed strategies has been assessed through experimental analyses concerning both synthetic and real-world maritime operational scenarios.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC Akaike's Information Criteria
AoI area of interest
AIS Automatic Identification System
BIC Bayesian Information Criteria
CMRE Centre for Maritime Research and Experimentation
CRLB Cramer-Rao Lower Bound
C-GLRT Coupled Generalized Likelihood Ratio Test (GLRT)
DBSCAN Density-Based Spatial Clustering of Applications with Noise
EEZ Exclusive Economic Zone
EM Expectation Maximization
EO Electro-Optical
ETA Estimated time of arrival
FIM Fisher Information Matrix
GLRT Generalized Likelihood Ratio Test
GMM Gaussian Mixture Model
GP Gaussian Processes
GPS Global Positioning System
GT Gross Tonnage
HFSW HF Surface Wave radar
H-GLRT Hierarchical GLRT
IMO International Maritime Organization
IOU Integrated Ornstein-Uhlenbeck

KDE Kernel Density Estimation
KL Kullback-Leibler
LRIT Long-Range Identification and Tracking
MLE Maximum Likelihood Estimate
MMSI Maritime Mobile Service Identity
MOU Mixed Ornstein-Uhlenbeck
MOS Model Order Selection
MS Maritime Surveillance
MSA Maritime Situational Awareness
MTG Maritime Traffic Graph
MTRD Maritime Traffic Route Discovery
NCV Nearly Constant Velocity
N-COST Non-Convex Optimized Stealth Trajectory
NN No-Name
OU Ornstein-Uhlenbeck
ROC Receiver Operating Characteristic
SAR Synthetic Aperture Radar
SaR Search and Rescue
S-AIS Satellite Automatic Identification System
SCA Suez Canal Authority
SDE stochastic differential equation
SOCP Second-Order Cone Program
STO NATO Science and Technology Organization
SVM Support Vector Machines
SVD Singular Value Decomposition
TREAD Traffic Route Extraction for Anomaly Detection
UTM Universal Transverse Mercator
VHF Very High Frequency
VTS Vessel Traffic Services
VMS Vessel Monitoring System

Chapter 1

Introduction

1.1 An overview of maritime surveillance

Along with its primal role in the development of ecosystems, its wildlife richness and being one of the last places on Earth mainly unknown to humankind, the World Ocean is an essential environment for the world economy, as some activities, such as transportation of goods, fishing, sailing and cruising occur at sea and impact the worldwide economy. For instance, 80 - 90% of global goods transportation and energy transportation are done by sea [111, 193, 222], and millions of people work on sea-related activities. However, the increasing shipping traffic leads to risks in coastal areas, harbors, and densely exploited maritime routes, as all vessels try to optimize their journey, creating conflicts in areas where a large amount of vessels are gathered [1].

Formally defined by the NATO Military Committee as "*an enabling capability which seeks to deliver the required Information Superiority in the maritime environment to achieve a common understanding of the maritime situation, in order to increase effectiveness in the planning and conduct of operations*," Maritime Situational Awareness (MSA) is the concrete understanding of anything associated with the maritime domain that could impact the security, safety, economy, or environment [9]. It aims at providing a seamless wide-area operational picture of ship traffic in coastal areas and the oceans in real time [40, 48, 90], and establishes a view of adversary activities, events, locations, and maneuvers. In the specific, Maritime Surveillance (MS) is essential for creating maritime awareness, represent-

ing an important domain for many national and international institutions, agencies, and bodies. Indeed, national authorities and international bodies require proper support to monitor increasing maritime activities and the related volume of data to guarantee safety, protect the environment, optimize traffic streams, and avoid illegal activities. Therefore, the MS efforts are aimed to enhance Search and Rescue (SaR) operations, provide effective response to accidents and disasters, monitor fisheries, prevent pollution and support law enforcement and national defense. In this regard, asset allocation, route planning and anomaly detection tasks are among all the possible applications within the MS context.

In order to accomplish these tasks, MS requires target detection, tracking, and identification, and deals with surface, subsurface, and airborne targets in considerable large areas. Besides, MS benefits from the availability of a large number of sensors, high data storage capacity, cheap devices and good database management systems able to handle huge volumes of maritime traffic data. With the aim of providing comprehensive maritime operational pictures, the huge amount of data collected by heterogeneous sensor systems need to be condensed in the right way, to avoid overwhelming the human decision-making process. This procedure allows to prevent the system-level failure due to potential communication interruption, hardware breakage or detection limits of individual systems, and, thus, to avoid the unique dependency on single system, as well as, to enhance the surveillance performance. Indispensable in this regard are the *multi-source data fusion* technologies [29], which use the data from multiple sensors to perform inferences that may not be possible from a single sensor alone. Data from different sensors are then combined using signal processing, pattern and image recognition, artificial intelligence, and information theory, in order to produce a single, consolidated surveillance picture. In research community, many works have been conducted to investigate the multi-source data fusion technologies for MS [27, 29, 31, 32, 35, 48, 75, 77, 92, 98, 137, 144].

1.2 Data sources for maritime surveillance

As previously mentioned, in support to the MS framework comes the rapid development of information and communications technologies, with the increasing availability of surveillance sensors, data stored in databases, networked system solutions and robust signal processing techniques. MS data

relate to vessels and their attributes (e.g., position, identification, history, etc.) and to contextual geographically-linked information. Currently, MS mostly relies on data collected by heterogeneous systems, all with their advantages and limitations. All around the world, the on-duty Vessel Traffic Services (VTS) [20] deployed at maritime traffic control and monitoring centres collect the traffic data from the MS networks that typically comprise multi-infrastructure sources. Specifically, vessel tracking data and attributes can be grouped into self-reporting (voluntary broadcast by collaborative systems) or observation-based (collected by active or passive sensors) depending on the way such data are acquired. Information registries and databases contain additional information about the vessel that is essential to build a solid awareness of maritime activities:

- **Collaborative data sources** comprise the systems working in a cooperative fashion as, e.g., systems for collision avoidance as the Automatic Identification System (AIS) [111, 114], for security and safety as the Long-Range Identification and Tracking (LRIT) [112], or for fisheries monitoring as the Vessel Monitoring System (VMS) [76].

- **Non-collaborative data sources** include coastal radars [20, 40, 99, 145,152,175], active and passive sonar, video and infrared cameras [95]. Furthermore, space-based sensors are of notable importance in routinely providing data for MS. Specifically, space-based Synthetic Aperture Radar (SAR) and Electro-Optical (EO) are effective for ship detection and are being operationally utilized by nations [219]. In addition, space-based AIS, or Satellite Automatic Identification System (S-AIS), provides capability for non-coastal regions with a detection performance and persistence that depend on satellite revisit rates, ship traffic density, and other interference sources.

A conceptual representation of some of collaborative and non-collaborative data sources is provided in Fig. 1.1. Basically, in ports and harbour, the coastal radars assist in detecting and tracking vessels' navigation in busy waterways; in remote ocean area and open oceans, the LRIT, S-AIS and SAR consist of the primary approaches for ship surveillance.

Each of the aforementioned stand-alone systems is usually not able to ensure reliable surveillance for handling complex scenarios (due to the specific effects of each technology as transitional loss of availability, errors, limits of coverage, etc.). For example, high resolution coastal radar technology is

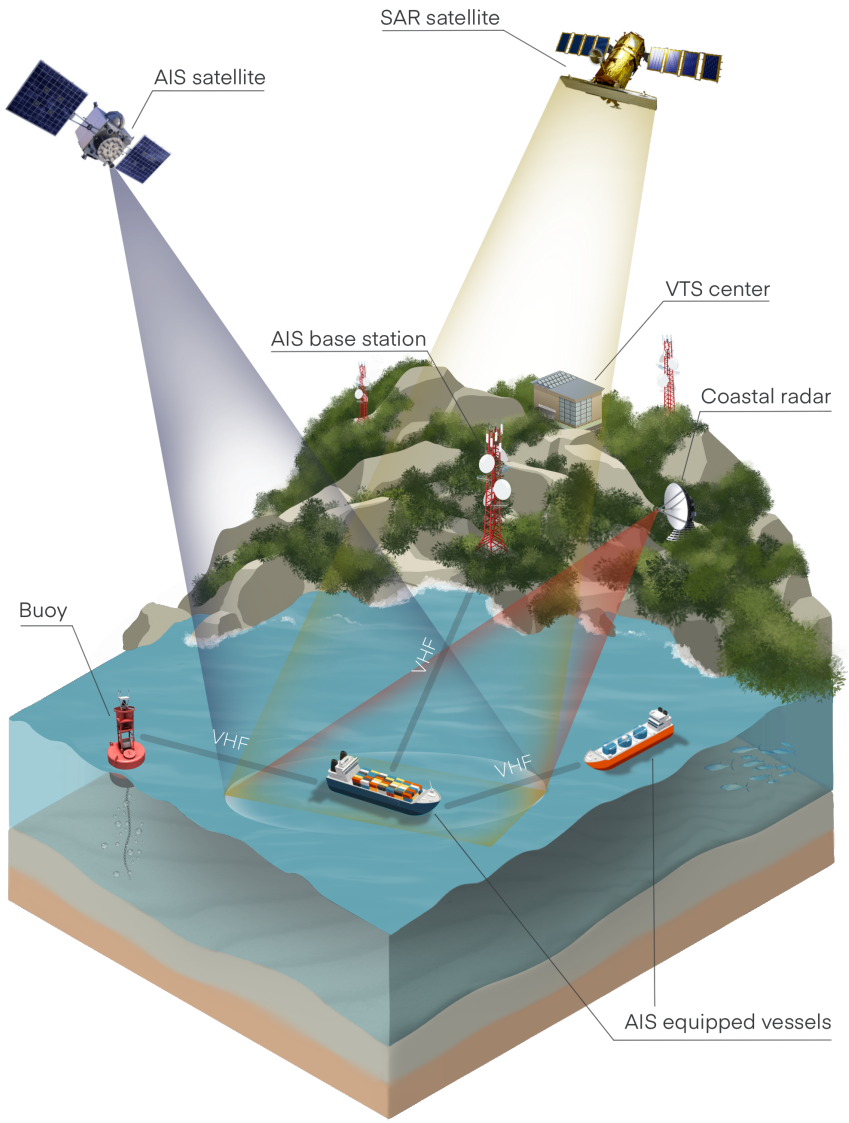


Figure 1.1: A conceptual representation of collaborative and non-collaborative data sources. An AIS equipped vessel transmits messages to other AIS equipped vessels and to shore-based stations, using dedicated radio channels. Alternatively, the messages can be gathered by the VTS center through satellite transmission (S-AIS). Besides, SAR satellites and coastal radars take actions in a non-collaborative way.

effective with high accuracy and persistent monitoring, but usually presents difficulties (occlusions, shadows, fragmentation, lack of identification) which make it necessary to supplement them with cooperative technologies such as AIS.

In the following, we provide a descriptive summary of the primary information systems in maritime traffic service networks.

Automatic Identification System

AIS technology is currently an important and widely used solution in smart transportation, with an estimated number of over 300,000 installations. Precisely, AIS, refers to a technology introduced to enhance the safety of vessel traffic by automatically exchanging up-to-date information as well as tracking and monitoring ships. AIS technology enables the exchange of both the dynamic (like position, velocity, course, rate of turn etc.) and static (vessel identifier, dimension etc.) information between ships and between ship and shore-based AIS station in near real time through Very High Frequency (VHF) radio transmission, as it will be better described in Appendix A.

AIS is mostly land based, also known as terrestrial AIS, of which the AIS signals are received by shore-based monitoring stations; AIS receivers can be also placed on low orbiting satellites to collect AIS signals, called S-AIS [47, 109] and thus vessels can be indeed tracked in the remote areas and open oceans beyond the reach of terrestrial AIS.

Since the 2002 International Maritime Organization (IMO) Safety on Life at Sea (SOLAS) convention [111], AIS has been a mandatory installation for international voyaging ships weighing at least 300 Gross Tonnage (GT) and all passenger ships, regardless of size. Nowadays, over half a million vessels use AIS to transmit their location, which is collected by a network of receivers deployed in over 140 countries and transferred for display and vessel tracking through platforms such as MarineTraffic[1], that has 40 million users annually alone [5]. As visible in Fig. 1.2, the amount of information reported by onboard AIS transceivers is impressive.

As the AIS network of receivers grows and its technology develops from a simple navigational and situational awareness tool for safety at sea to the backbone of a global ship tracking network, AIS data is able to provide ever-more accurate and valuable information to track the position of

[1] http://www.marinetraffic.com

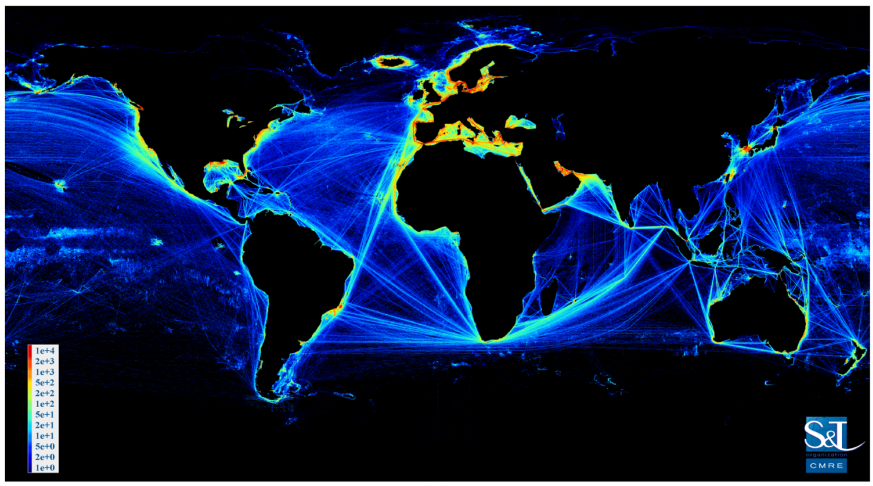


Figure 1.2: Density of traffic in the world computed using AIS data collected from multiple AIS networks from April to September 2012 at the NATO Science and Technology Organization Centre for Maritime Research and Experimentation [154].

ships anywhere in open waters or sea passageways in near real time. AIS is currently a major technology and solution in traffic monitoring and vessel assistance. Shipowners and maritime authorities rely on AIS to supplement traditional radars for collision avoidance, SaR operations, accident investigation and location tracking, in addition to complementary systems for visual observation, audio exchange, and LRIT.

Large volumes of streams of AIS messages were recently processed to visualize ship routes and traffic density maps [78], identify global patterns of transshipment behavior [159], analyze the effects of COVID-19 on global maritime mobility [156], and to train deep learning networks for long-term prediction of vessel trajectories [45]. This information is also fundamental to enable advanced anomaly detection systems for the maritime domain to make decisions based on near real-time data.

Long-Range Identification and Tracking System

LRIT system was established by IMO in 2006, in order to complement other existing maritime information systems to enhance MSA [112]. LRIT is a satellite-based, real-time reporting system that collects and disseminates vessel position information [213]. According to the LRIT regulation, all passen-

ger ships, ships used for the purpose of cargo-carrying, speedier craft of 300 GT and above, and the mobile offshore drilling units must report their position at regular intervals (at least every six hours). The LRIT system consists of the transmitting device and satellite communication equipment, communication service providers, the overall LRIT application service providers, LRIT data centres, the data distribution plan and the International data exchange [112]. S-AIS and LRIT shares some similarity as both systems allow for ship detection and identification from space, but they entirely differ in the working mechanism.

Coastal radar

Coastal radar is a shore-based sensor radar system that is designed to meet the requirements of VTS to detect and track vessels in the area covered by the radar site. Coastal radar gives regular information about the position and velocity of a vessel, serving as the basic infrastructure of VTS for maritime traffic surveillance in ports, harbour and busy waterways. The shore-based radar systems mainly focus on collision monitoring and grounding prevention, and generally aim at achieving the safety and security aspect in the covered area [20]. In research community, the HF Surface Wave radar (HFSW) radars (a type of over-the-horizon radar that works with ground-wave) attract increasing interests recently due to their wider coverage beyond the conventional microwave radar detection ranges. HFSW radars provide additional data sources and enable wide-area surveillance in MSA applications [40, 99, 145, 175].

Satellite-borne Synthetic Aperture Radar

SAR has been utilized for maritime surveillance worldwide [122]. The high-resolution imagery data captured by SAR platforms are able to be responsible for ship detection over wide water area and in all weather conditions [43, 225]. Satellite-borne SAR is developed to overcome the coverage limitations of coastal systems and is suited for oceanographic observations. The popular SAR platforms include RADARSAT [206], ENVISAT-ASAR [104], Sentinel-1 [103,192], Cosmo-SkyMed SAR system [147,148] and TerraSAR-X [167, 178, 216] and TanDEM-X [160] etc.

In addition EO satellites are also used for vessel tracking and monitoring [150,219]. The optical satellite images can extend the SAR based systems

through more frequent revisit times and have a higher spatial resolution for the detection of smaller vessels [150].

1.3 The anomaly detection problem: definition and practical scenarios

Although there is no univocal definition, in general, an *anomaly* always represents a deviation from what is standard, normal, or expected. In today's data-driven world, anomalies are often referred to as non-conforming patterns, outliers, discordant observations, exceptions, aberrations, surprises, peculiarities or contaminants in different application domains, representing anything not fitting the general trend. Anomalies might be induced in the data for a variety of reasons, such as malicious activity like credit card fraud, cyber-intrusion, terrorist activity or breakdown of a system [50], and they can be identified through various mechanisms, generally known as *anomaly detection* algorithms. Anomaly detection has been studied in the statistics community as early as the 19th century [71], and has found extensive use in a wide variety of applications, such as fraud detection for credit cards, insurance or health care, intrusion detection for cyber-security, fault detection in safety critical systems, and military surveillance for enemy activities.

A straightforward anomaly detection approach is to define a normal behavior and declare any observation in the data which does not belong to this normal condition as an anomaly. But several factors make this apparently simple approach very challenging:

- Defining every possible normal behavior is very difficult. In addition, the boundary between normal and anomalous behavior is often not precise.
- The exact notion of an anomaly is different for different application domains.
- When anomalies are the result of malicious actions, the malicious adversaries often adapt themselves to make the anomalous observations appear like normal, thereby, making the task of defining normal behavior more difficult.
- In many domains, normal behavior keeps evolving and a current notion of normal behavior might not be sufficiently representative in the

future.

– Often the data contains noise which tends to be similar to the actual anomalies and, hence, is difficult to distinguish and remove.

Due to the above challenges, the anomaly detection problem, in its most general form, is not easy to solve. Indeed, most of the existing anomaly detection techniques solve a specific formulation of the problem. The formulation is induced by various factors such as nature of the data, availability of labeled data, type of anomalies to be detected, etc. Often, these factors are determined by the application domain in which the anomalies need to be detected. Researchers have adopted concepts from diverse disciplines such as *statistics*, *data mining*, *machine learning*, *information theory*, *spectral theory*, and have applied them to specific problem formulations [50].

1.3.1 The maritime anomaly detection problem

Anomaly detection in the maritime domain was identified by the operators/-analysts of the operational community as an important aspect for MSA requiring research and development [146]. More specifically, automatic anomalous maritime vessel behavior detection consists of finding anomalous movement behavior of vessels moving in maritime areas such as in harbours or in open sea using computational methods [197]. Anomalous vessel behavior refers to vessel movement behavior that differs from the typical or normal vessel movement behavior. For instance, as it will be better described in Subsection 1.3.4, anomalous vessel behavior can refer to a sudden change in vessel kinematic behavior (such as unusual speed or location), deviation from standard sea lanes, unexpected port arrivals, close approach, and zone entry [127]. Therefore, anomalous vessel behavior detection is one of the main tasks for MSA, allowing the identification of suspicious situations or threats in the maritime domain and taking appropriate action. Indeed, surveillance operators have to constantly search and predict emerging conflict situations, e.g., risk for collision, anomalous vessels or suspicious activities from a large number of vessels within vast sea area, and a prompt detection of such situations provides critical time to take appropriate action with, possibly before potential problems occur.

Nowadays, in addition to the available nominative and positioning information broadcast by collaborative systems (e.g., AIS, VMS, LRIT), other information related to environmental, contextual and geographical data are

also at hand and can be aggregated as complementary and correlative sources to the exploitable data [181]. Therefore, merged and analyzed, together with a large variety of contextual information of different natures and types, vessel positioning data can be used to monitor maritime mobility, uncover maritime activities, illegal trafficking or risks for the environment, living resources and the navigation. However, exploring and monitoring the data manually is a demanding task, not only due to the complexity and heterogeneous nature of the data itself but also due to other factors like uncertainty, fatigue, cognitive overload, or other time constraints. Consequently, increasing automation through a large number of advanced methods and techniques for maritime anomaly detection has enabled the system and the operator to spot complex situations by correlating various events from all surveillance sensors and classify them into important incidents.

According to the literature [189,197], the main approaches for anomalous vessel behavior detection can be divided into three categories: *data-driven*, *rule-based*, and *hybrid*, as outlined in Fig. 1.3.

Rule-based approaches for maritime anomaly detection refer to the generation of alerts based on a set of rules [191], such as maximum speed allowed in a port, presence in areas restricted to navigation or inconsistencies between the ship's claimed and actual activity. These approaches are composed of two phases: *i)* creating and specifying predefined patterns that one wants to detect using rules; *ii)* the new data are matched to the predefined rules to detect predefined anomalies of interest, which are those data that deviate from the predefined rules. However, rule-based models are generally deterministic and do not take advantage of the variability found in large datasets, instead relying on heuristics to assign the parameters of the model. In contrast to rule-based models, the data-driven approaches do not foresee a prescribed set of rules, and they can learn arbitrarily complex patterns from large amounts of data. The hybrid approach combines the data-driven and rule-based approaches to perform anomalous maritime vessel behavior detection, however, recently, only limited effort has been devoted to the use of this type of approaches [120, 195, 196].

Since data-driven approaches appear easier to apply on a large scale to gain efficient classification performance and detect different types of anomalies [189], the majority of studies in this research area fall under the data-driven category, which will be considered in depth in the following subsections.

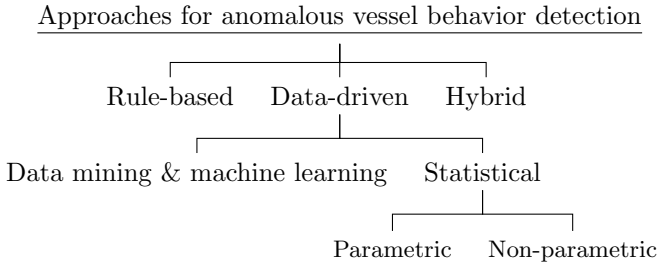


Figure 1.3: Approaches for anomalous vessel behavior detection divided into three categories: data-driven, rule-based, and hybrid. Data-driven approaches, in turn, can be divided into distinct categories: statistical (parametric and non-parametric), machine learning and data mining (vector-based and grid-based).

1.3.2 Data-driven methods and techniques for maritime anomaly detection

Data-driven approaches typically involve two common phases:

1. *Normalcy extraction*: a model representing the normal behavior is learned from the available historical data;

2. *Anomaly detection*: new vessel movement data are matched to the learned normalcy model of vessel behavior, considering any mismatch as anomalous behavior.

Essential for effective anomaly detection is thus building an accurate model of normalcy. Most of the data-driven approaches build models representing the normal vessel behavior through the analysis of historical AIS data [80, 87, 168, 171, 187]; indeed, accumulating tracks from AIS data over a long time period can establish a pattern of typical movements, namely *pattern of life* [19]. More specifically, patterns of life are understood as observable human activities that can be described as patterns related to a specific action taking place at a specified time and place. Essentially, vessel-based maritime activity can be described in space and time, while classified to a number of known activities at sea (e.g., fishing, dredging, etc.). The spatial element describes recognised areas where maritime activity takes place; thus, including ports, fishing grounds, offshore energy infrastructure, dredging areas and others. The transit paths to and from these areas also describe the spatial element, (e.g., commercial shipping and ferry routes etc.), while the

temporal element often holds additional information for categorising these activities (e.g., fishing period, time of year).

Once the picture of the maritime traffic is derived through the analysis of the increasingly available stream of AIS data, this historical knowledge can be applied to: *i*) classify the routes, i.e., identify compatible routes, assigning a probability that the vessel is actually following a specific route; *ii*) predict future vessel positions, i.e., predict the route along which a vessel is going to move, in agreement with the partially observed track and given the vessel static information (e.g., vessel type); *iii*) support decisions and action planning at different information levels, enhancing the detection of anomalous behaviors, i.e., behaviors deviating from the learned traffic normality; *iv*) optimize sensor resource management, i.e., the problem of allocating the available sensor resources in order to obtain the optimal awareness of the situation.

According to the available proposals and studies [189, 197, 227], the data-driven methodologies for maritime anomaly detection can be divided into: *data mining and machine learning* approach and *statistical* approach, as outlined in Fig. 1.3.

Data mining and machine learning approach

Data mining and machine learning techniques [97, 121, 230] are able to identify patterns emerging within huge amounts of data, fused from various uncertain sources and generated from monitoring thousands of vessels a day, so as to act proactively and minimize the impact of possible threats. The general aim of such an approach include frequent pattern discovery, trajectory pattern clustering in a multidimensional feature space, trajectory classification, forecasting, and anomalies/outliers detection.

Clustering-based methods: Cluster analysis is one of the basic tools for exploring the underlying structure of a given data set and its primary objective of cluster analysis is to partition a given data set of multidimensional vectors (patterns) into so-called homogeneous clusters such that patterns within a cluster are more similar to each other than patterns belonging to different clusters. The huge volume of spatial trajectories enables opportunities for analyzing the mobility patterns of vessels, which can be represented by an individual trajectory containing a certain pattern or a group of trajectories sharing similar patterns. Indeed, clustering is a form of unsupervised learning and involves automatic grouping of data into multiple clusters ac-

cording to distinguishing features. Assuming that the data set reflects what is considered more or less normal, the resulting set of fairly large clusters is regarded as a model of normalcy. New data presented to the model are regarded as more or less anomalous based on the distance to the nearest cluster.
Therefore, vessel tracks can be clustered into a number of routes, where the features of the track are attributes such as longitude and latitude, speed and course, and various clustering-based methods can be applied for maritime anomaly detection:

- *Distance-based methods* optimize a global criteria based on the distance between patterns. *k-Means* is a distance-based method that defines k disjoint clusters on the basis of the feature value of the objects to be grouped. *k-Medoids* is very similar to the k-Means algorithm, but it represents each cluster by the most centric object in the cluster, rather than by the implicit mean that may not belong to the cluster. Its main drawback is the selection of the number of clusters, which can be optimized by using an Expectation Maximization (EM) algorithm.

- *Density-based methods* identify distinctive groups/clusters in the data, based on the idea that a cluster in a data space is a contiguous region of high point density, separated from other such clusters by contiguous regions of low point density. Recently, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm have become very popular for their convenient properties: these methods do not require to specify the number of clusters and have the ability to derive arbitrarily shaped clusters and identify outliers [73].

The strength of clustering-based algorithms stem from the fact that they are inherently unsupervised and have an intuitive criteria for detecting outliers/anomalies. Their limitations include the quadratic computational complexity and a possible incorrectness when handling high dimensional data.

Classification-based methods: The key problem of intelligent maritime surveillance is the modelling of typical sailing patterns of training vessel trajectories and classification of vessel behavior to be monitored. In machine learning, *classification* is considered as an instance of supervised learning in data analysis and pattern recognition that requires the construction of a *classifier*, that is, a function that assigns a class label to instances described by a set of attributes. Using supervised learning approaches, trajectories or

segments of a trajectory can be classified into some categories, which can be motions, human activities, or transportation modes. In general, trajectory classification is comprised of three major steps: *i*) division of a trajectory into segments using segmentation methods; *ii*) extraction of features from each segment (or point); *iii*) building of a model to classify each segment (or point). For anomaly detection purposes, the data are generally classified into two categories, namely normal or anomalous.
Common learning models used for classification are:

- A *Decision Tree* is a machine learning classifier that uses a tree graph model, similar to flow chart structure, to introduce conditional statements in which each node represents a conditional test on an attribute to output a class represented by one of the leaves. Then, the internal nodes are a test property, each branch represents test result, and final nodes, or leaves, represent the class to which any object belongs. A training process is used to build the tree that minimizes the classification error. *Random forest* is an ensemble learning method that relies on a combination of many decision trees for classification purpose. It uses bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree.

- *Neural networks* represent one of the most effective data-driven approaches for the supervised learning task. They are configured as a set of interconnected nodes designed to imitate the functioning of the human brain, where each node has a weighted connection to several other nodes in neighbouring layers. A Neural Network basically infers a function that maps an input to an output based on example input-output pairs derived from existing data. Compared with traditional Neural Networks, *Recurrent Neural Networks* turn out to be more capable of processing time series data, consequently, they are well suited to the tasks of maritime anomaly detection since the vessel trajectory data consists of tracking points with time stamps, a type of typical time series data [231].

- *Associative learning* is a type of data mining approach to discover taxonomic relationships between objects or concepts of interest [183]. This method, based on biological principles, is able to determine anomalous behaviors and predict vessels' navigation states that include both

location and velocity prediction.

- *Support Vector Machines (SVM)* are machine learning models that can be used for classification and anomaly detection. Basically, the SVM algorithm's goal is to create the best line or decision boundary that can decompose the multidimensional feature space into sets supporting categorization of new data points. In their simplest forms, SVM suffer from a number of problems that have limited their use in vessel anomaly detection, including a lack of partial assignment, a restriction to binary classes, high computational complexity and difficulties in summarizing and communicating the learned models.
- *Logistic Regression* is a classification technique borrowed by machine learning from the field of statistics. Specifically, Logistic Regression is a parametric a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The intention behind using Logistic Regression is to find the best fitting model to describe the relationship between the dependent and the independent variable. Different from Linear Regression, the outcome of Logistic Regression on one sample is the probability that it is positive or negative, where the probability depends on a linear measure of the sample.
- *Fuzzy Logic*: was first introduced in [229] and relies on the theory of fuzzy sets. This approach exploits the notion of degree in the verification of a condition, enabling conditions to be in intermediate states between the states of conventional evaluations, thus allowing variables to be "partially" true, or "not definitely yes" etc. Fuzzy logic has been used for static anomaly detection as it is considered to be an ideal tool when dealing with imprecise or contradictive data, which can be modelled adequately with fuzzy sets, and combined with human logic.

Statistical approach

Statistical approaches to maritime anomaly detection fit a statistical model representing normal vessel behavior to the given historical vessel movement data, and then apply a statistical inference test to determine whether a new vessel observation belongs to this model or not. Observations that have a low probability of being generated from the learned model, based on the applied test statistic, are declared as anomalous behaviors.

Basically, traffic statistics provide quantitative modelling and representation of the traffic features which can be used for determining some important traffic parameter values (or distribution) and thresholds to help distinguish the normal and abnormal navigation behavior.

In [189], authors identify two main statistical methods: parametric and non-parametric methods.

- *Parametric methods*: a specific statistical model relying on the historical data is assumed. Gaussian Mixture Model (GMM) is a very popular unsupervised parametric method which approximates the (unknown) statistical distribution of the normal vessel traffic by using the EM algorithm to estimate the distribution parameters. A GMM can be regarded as an ensemble model of K multivariate Gaussian distributions (mixture components).
  Bayesian Networks represent another parametric method, which assumes an underlying probabilistic model and it allows capturing uncertainty about the model in a principled way by determining probabilities of the outcomes. More specifically, These networks are directed acyclic graphs, with evidence propagation governed by Bayes' theorem [176], and allow efficient and effective representation of the joint probability distribution over a set of random variables: each node corresponds to a random variable and each edge represents the conditional probability for the corresponding random variables. Naïve Bayes classifiers are among the simplest Bayesian network models with strong independence assumptions between the features. These classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem, and, coupled with Kernel Density Estimation (KDE), they can achieve higher accuracy levels.

- *Non-parametric methods*: no assumption is made about the underlying distribution of the data. In this case, KDE is a popular method which derives a non-parametric model of traffic normalcy. One tricky step is the selection of the free parameters of KDE: the kernel, which specifies the shape of the distribution, and the kernel bandwidth, which controls the size of the kernel.
  Gaussian Processes (GP) are Bayesian non-parametric statistical models, which are mainly applied to regression and classification tasks, and are able to cope with unreliable, noisy, or partially missing data. An

advantage with GP is that the model is non-parametric so it is not necessary to build in features of anomalous behavior. A limitation of this approach is that although GP provide a flexible and robust approach to anomaly detection, they are not typically suitable for large datasets due to high computational complexity in training [123].

1.3.3 Representation of the maritime traffic knowledge

Fundamental to clustering is the feature model, i.e., the representation, of the data in which we want to detect anomalies, what characteristics of the data we chose and how we model these characteristics. According to the available proposals and studies [189, 227], the maritime traffic knowledge representation can be categorized as *vector-based* and *grid-based* methods.

Vector-based methods

Normalcy condition can be represented by means of vector-based methods, which intend to establish the maritime traffic network through extracting the network nodes, namely *waypoints*, and the sea routes, namely *waterways*, and thus to model the traffic states along a navigation voyage. The waypoints are regions where ships regularly stop or change their velocity, such as like entry/exist/turning point. The waterway pattern essentially reflects the vessels' navigating routines derived from the maritime traffic operations, traffic planning and regulations, vessels' manoeuvrability and hydrographical features etc., which represents the practical sea routes through voluminous real-world navigation instances [227]. The temporal information like route travel time distributions and daily patterns is able to be extracted. The historical route patterns can be associated to the navigating vessels in real time to support traffic analysis and traffic forecasting [169].

A vector-based approach relying on point clustering is the Traffic Route Extraction for Anomaly Detection (TREAD) algorithm developed in [169]. TREAD generates a set of historical patterns of life represented by waypoints and route features. Waypoints are defined as stationary objects like ports and offshore platforms and entry and exit points [90], and are clustered using DBSCAN methodology [73].

Successful advancements on vector-based knowledge discovery [61, 80] have recently led to the development of an unsupervised graph-based methodology to identify the spatiotemporal dynamics of ship routes, and efficiently

extract a compact representation of global maritime patterns from large volumes of historical AIS data, in the form of a Maritime Traffic Graph (MTG). This method builds on recent advances in long-term vessel motion modeling [154,155] whereby the dynamics of ships can be effectively described by a piecewise Ornstein-Uhlenbeck (OU) mean-reverting stochastic process. This approach, extensively validated against real-world datasets [154, 157], relies on model parameters. The OU model statistically represents the dynamics of maritime traffic allowing to synthesize historical ship trajectories into a sequence of waypoints connected together by a network of navigational legs represented by the MTG.

Starting from raw AIS streams, the MTG model can be automatically extracted based on the following processing steps:

1. **Detection of navigational waypoints**: Based on the OU dynamic model for accurate long-term ship prediction, change detection algorithms [61,158] can be applied to identify specific geospatial waypoints where the mean long-term velocity parameter of the underlying OU process abruptly changes. The detected waypoints represent: *i*) ports, where a ship's speed is null either before (leaving the port) or after the change (entering the port); *ii*) navigational waypoints, where the direction of the ship changes (while the speed is possibly constant); *iii*) entry, exit, and entry/exit points, i.e., virtual boundary regions of the area of interest that summarize the entering and exiting traffic.

2. **Clustering of navigational waypoints**: Based on the assumption that most maritime traffic is inherently regular, change points are expected to be concentrated around specific geospatial regions. To find these significant waypoint areas, standard density-based clustering techniques can be used in order to group together multiple change points into a lower number of distinct waypoint clusters.

3. **Merging and pruning procedures**: Pruning and merging techniques can be used to progressively improve and simplify the overall MTG by reducing the number of graph entities (i.e., nodes and edges), and thus implicitly encode knowledge about ships' patterns using a lower-dimensional representation. The number of graph edges can be reduced by eliminating those links characterized by low weights, these being least likely to represent recurrent patterns; and by merging closely-spaced edges (connecting waypoints in a cluster with waypoints

in a different cluster) into one, as these are more efficiently represented by a single route.

Thus, the vector-based approaches allow for high compactness network representation of the waypoints and the traffic routes connecting the waypoints and thus vessel motions and traffic patterns in large areas and even at a global scale can be characterized as a graph-based network.

However, most of vector-based methods had indeed only been implemented in bounding areas. And, besides, the vector-based approaches are only considered operative in areas with high traffic but they are difficult to be effective for unregulated areas; in other words, the vectorial representation has its limitations where the behavior of vessels is difficult to categorize.

Grid-based methods

Grid-based techniques usually divide the monitored maritime traffic zone using a spatial grid whose cells contain the motion properties (e.g., location, heading and speed) of the crossing vessels.

This approach is fast enough to detect simple point anomalies but it suffers from some limitations [227]. First of all, the grid-based methods have been considered effective only for small area surveillance and the computational burden was regarded as its limitation when increasing the scale [37, 188]. Another limitation pertains to that a priori determination of the grid size is required [127, 169]. The selection of the optimal cell size is critical and often not straightforward and the adoption of a grid structure looses the spatio-temporal correlation of vessel tracks, making difficult or impossible the detection of structured anomalies (e.g., start/stop events); for example, it is not reasonable for open sea (with less traffic) and harbour-like areas (with dense traffic) to adopt uniform grid size. Other factors like vessel size and speed may also affect the decision of grid size. The vector-based methods have also been proposed to overcome these above-mentioned difficulties.

1.3.4 Maritime anomalous behaviors

The main maritime anomalies, described in the scientific literature [11, 115, 189], are schematically reported below:

- **Kinematic anomaly**: a ship exhibits significant changes to course (e.g., 90° in a short timeframe), change of speed (high-speed or low-

speed), deviation from known shipping lines in the high sea, instantaneous stops and turns without reporting any mechanical issues, etc.

- **Historical (shipping route) anomaly**: a ship is found travelling outside the normal route it is expected to follow in a given area and time as shown in Fig. 1.4(a). Or there is a mismatch in its type of activity reported for the given area or time period (e.g., fishing in a forbidden area).

- **Rendezvous anomaly**: ships that come alongside each other within a given distance either at a full stop or low speed. Potential rendezvous aim at transferring either crew, passengers and cargo. An example of rendezvous is given in Fig. 1.4(b).

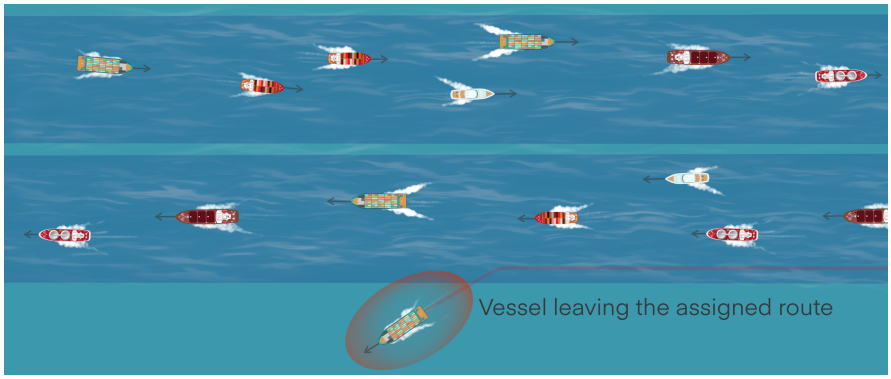
- **Positional anomaly**: ships that are loitering out of usual areas such as anchorage areas or waiting areas, as depicted in Fig. 1.4(c).

- **Estimated time of arrival (ETA) anomaly**: ETA is an attribute of civilian vessel information to indicate when it is planning to arrive at the indicated destination. ETA anomaly is due to the inconsistency between the time required for a vessel to reach its destination (based on characteristic, e.g., cruising speed and positional information) and the actual time of the journey.

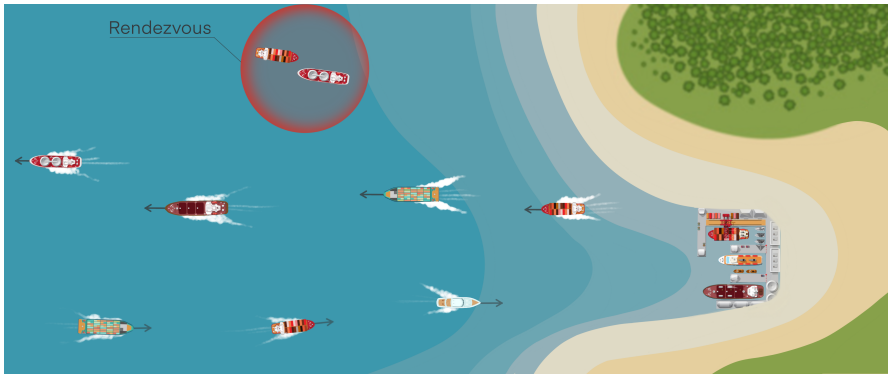
- **Destination anomaly**: undeclared destination, destination inconsistent with present course, speed and ETA, or unexpected port arrival. If a ship of a particular type arrives at a port that has no facilities to handle it, then this could be considered an anomaly.

Vessels exhibiting one or more of the listed anomalous conducts could be connected to illegal activities at sea. More specifically, the *stealth illegal activities* [115], in which the culprits aim to remain out of sight and undetected by law-enforcement bodies throughout the whole duration of the activity,[2] include drug and arm smuggling, human trafficking, poaching, prohibited imports/exports, unauthorized cargo transshipment, unreported or unregulated fishing (ignoring marine protected areas), unauthorized maritime arrivals and illegal immigration, marine pollution and waste dumping,

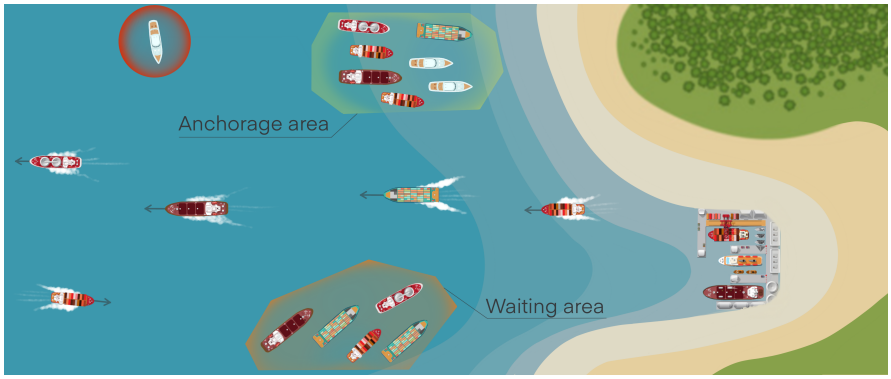
[2]Differently, the *adversarial illegal activities* [115] involve open and direct confrontation with other vessels, which is typically not hidden. Piracy and maritime terrorism are examples of such activities [161].



(a) Shipping route anomaly.



(b) Rendezvous anomaly.



(c) Positional anomaly.

Figure 1.4: Examples of maritime anomalies. Panel (a) shows a vessel leaving the assigned route. Panel (b) shows a representation of a rendezvous between two vessels. Panel (c) shows an example of positional anomaly.

and other unlawful activities. In some cases, the geographic characteristics of certain areas can either hinder or support such illegal operations. For instance, the desolation and other physical attributes of certain coastal areas make them attractive sites for staging smuggling operations. Besides, the occurrence of natural phenomena such as rain, fog, haze, blizzards, darkness and certain sea conditions, making visibility poor or nil, turns out to be beneficial for the malicious perpetrators.

Unexpected AIS activity

The huge amount of AIS data acts as a fundamental support to automatically identify the aforementioned anomalous set patterns depending on the illicit activity in which the vessels are engaged, however, AIS datasets could be affected by limitations and pose some significant challenges. These actually constitute effective anomalies associated to an unexpected AIS activity. More specifically, we need to take into consideration that a part of the information contained in AIS messages are entered manually by the crew, both at the initialization of the system for permanent data and at every new journey for travel-related data (such as the destination for instance), and that the AIS messages are transmitted in a non-secure channel.[3] Consequently, three main issues arise [25, 108, 113]:

1. *Unintentional human errors*: Each human-filled field is subject to errors. Both static data (ship identification number, type of the vessel, name of the vessel, physical characteristics) and dynamic data (position, navigation status, ETA or destination) are subject to errors [108].

2. *Internal AIS data falsification or interruption*: The crew can deliberately modify the AIS messages, such as the Maritime Mobile Service Identity (MMSI) number, or the destination, with the purpose of dissimulating.
   The crew can also interrupt the transmission simulating a malfunction; switching the AIS transponder OFF could indicate a will to hide some illegal activities, such as smuggling on coast or with other ships, or fishing in unauthorized areas.

[3] Recent advances in maritime radio technology brought to the development of the VHF data exchange system [8], which builds on the capabilities of AIS to overcome today's network limitations by providing higher data rates, and by enabling two-way and encrypted communications.

3. *External AIS data spoofing*: The AIS information are manipulated from an external actor in order to mislead both the crew on board and the outside world on the behavior of a vessel. The spoofer creates a ghost vessel that would cross the trajectory of another one, forcing the real one to change its heading, and could hypothetically be guided to hazardous locations [25].

1.3.5 Real-world maritime anomalous scenarios

We report below two specific cases of anomalies in maritime domain related respectively to the illegal fishing activities around the Galápagos Marine Reserve happened in 2017, and the grounding of the *Ever Given* in the Suez Canal in 2021.

Illegal fishing activities around the Galápagos Marine Reserve and waters off Ecuador

Galápagos, protected as a UNESCO World Heritage site since 1978, is home to about 30-odd species of sharks, and some of those are critically endangered. Sharks in the Galápagos Islands are being decimated by fleets of fishing vessels, many of them Chinese, and that is bringing these vital creatures to the brink of extinction [38]. The longliners, in the specific, are ostensibly fishing for tuna, but they also encroach on the paths of shark species that migrate into and out of the waters of the Galápagos.

In August of 2017, a Chinese fishing fleet comprising about 300 boats (e.g., fishing, cargo and fuel boats) was detected in international waters and near the Exclusive Economic Zone (EEZ) of Galápagos Islands and waters off Ecuador's coast in the Southeastern Tropical Pacific. The Chinese factory-mother ship (Fu Yuan Yu Leng 999) was caught in the Galapágos Marine Reserve carrying about 300 tons of fish, including several endangered species [17, 38]. The upper subplot of Fig. 1.5 shows the AIS track of the mother ship navigating in the waters of the Pacific Ocean from July 7th to December 9th 2017. As represented in the lower subplot of Fig. 1.5, between the 8th and the 11th of August 2017, the vessel shuts the engines down and starts drifting, with an apparent deviation from its route to rendezvous with four tuna longliners at about 1700 miles away from Galápagos EEZ. Each fishing vessel spends about 12 hours moving along with the vessel at a distance of about 30 meters, which indicates the boats were likely tied

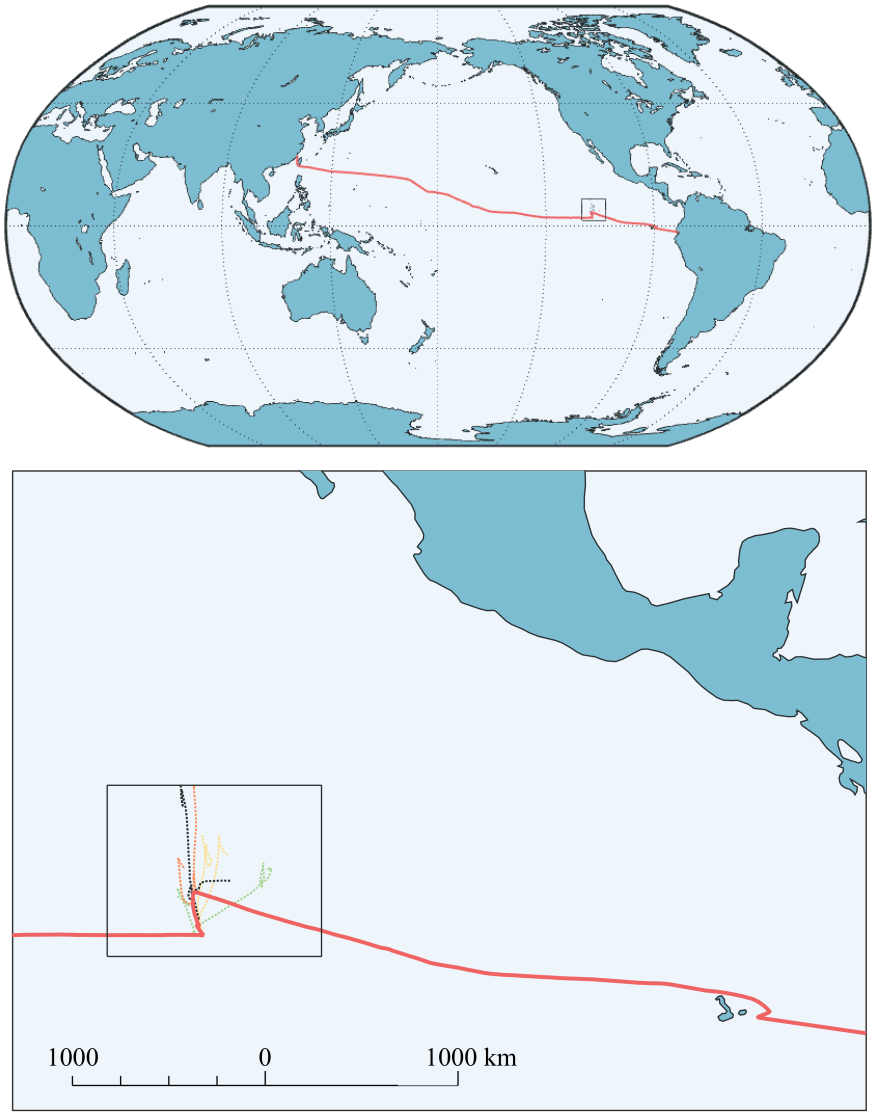


Figure 1.5: The red track of the mother ship Fu Yuan Yu Leng 999 reveals rendezvous with four fishing longliners (orange, yellow, green and black trajectories) at about 1700 miles away from Galápagos between the 8th and the 11th of August 2017.

up [17]. Subsequently, the vessel and the illegal fish catch were confiscated by Ecuador's Armada and the Galápagos National Park [16, 26].

Suez Canal blocked after massive container ship *Ever Given* gets stuck sideways

On March 23rd 2021, at 05:40 UTC, one of the world's biggest container vessels operated by Evergreen Marine Corp., the 400 meters (1300 feet) *Ever Given*, was navigating northbound through the Suez Canal on its way to Rotterdam during a dust storm with wind speeds reaching 40 knots, when it became stuck in a 300 meters wide channel at coordinates 30.01761°N, 32.58018°E (see Figs. 1.6 and 1.7). According to the Suez Canal Authority (SCA), the ship lost its ability to steer amid high winds and a dust storm, causing it to run aground diagonally and become wedged with its bow in one bank, and its stern nearly touching the other. The cause of the grounding is still unknown. While initial investigations exclude any mechanical or engine failure and strong winds during a dust storm are widely seen as a major factor, human and technical errors cannot be ruled out [2]. As is well-known, there was adverse weather and the ship lost control: it deviated in its course and collided with the bank, becoming wedged into the side of the canal (see Figs. 1.6 and 1.7).

The *Ever Given* grounding had immediate consequences, resulting in the complete blockage of the Suez Canal, extensive "traffic jams" of more than 360 ships (see Fig. 1.7), and disruptions in the maritime global trade for six days. A similar disruption, but less intensive, was observed at global level during the first half of 2020 because of the COVID-19 pandemic, see details in [156]. As one of the world's busiest trade routes, representing nearly 12% globally, this blockage of the Suez Canal had a significant negative impact on trade between Europe and Asia. It has been estimated that $9.6 billion worth of trade was held up each day as a result of the stranded ship, causing supply chain disruptions all over the world. On March 29th, at 13:04 UTC, the ship was finally freed and started moving towards the Great Bitter Lake for technical inspection, while traffic completely resumed in the canal at 17:00 UTC. A crucial assistance in the freeing of the *Ever Given* from the Suez Canal was played by an unusual full moon, which boosted a spring or "king" tide, helping the tugs refloat the ship [4].

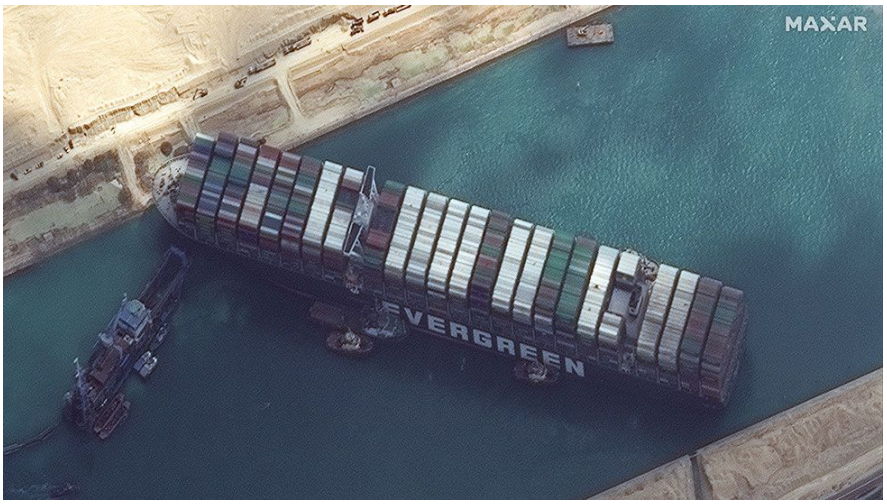


Figure 1.6: A high-resolution satellite view (captured on March 25th 2021) of the *Ever Given* container ship, which grounded in the Suez Canal on March 23rd 2021. Credit: Satellite image © 2021 Maxar Technologies.

Other relevant maritime incidents

According to the European Maritime Safety Agency (EMSA), over the period 2014–2019 the total number of reported marine incidents was 19418, involving a total number of 21392 ships and causing 496 fatalities, mainly occurring during collisions [12]. Recent examples of major real-world accidents that shed light on the key role of maritime surveillance data for safety and security at sea include fatal crashes involving commercial and navy ships such as the USS Fitzgerald [7], and the USS John S. McCain [6], the collision between the container ship Delphis Gdansk and the cargo ship BBC Neptune in the Great Belt Strait [3], the collision between the container vessel ANL Wyongand and the gas carrier King Arthur investigated by the UK Marine Accident Investigation Branch [13].

1.4 The objective

The cases and incidents presented in Subsection 1.3.5 illuminate the critical role of maritime surveillance on a global scale, and automatic systems for maritime anomaly detection, providing early warning and alert triggering to

Figure 1.7: Two views from Europe's Sentinel 1 radar observation mission show nearly 200 ships gathered on March 25th 2021 waiting for passage through the North (bottom left) and South (bottom right) entrances of the Suez Canal. The COSMO-SkyMed radar image above shows the *Ever Given* still stranded on March 25th. Credits: Copernicus Sentinel 1 data, © ESA, CC BY-SA 3.0 IGO; COSMO-SkyMed image © ASI, processed and distributed by e-GEOS.

avoid issues of navigational safety such as accidents and collisions, and help to identify suspicious activities to increase security and border protection. With anomaly detection systems that provide an enhanced maritime situa-

tional picture in real or near real time, these situations can be automatically detected, alerting operators by means of e.g., graphical user interface, e-mail, or warnings sent through the user's national system.

Therefore, the maritime surveillance domain continuously requires advances in the anomaly detection branch as well as in other research fields such as sensor performance and data fusion, providing a wide interesting environment for the investigation and application of several statistical solutions for the behavioral analysis of anomalous vessels.

As part of this pursuit, the proposed research aims at developing a novel statistical methodology to contribute to advances in the anomaly detection strategies, in order to improve the ability to reveal, understand, anticipate and prevent illegitimate activities at sea. The work proceeds along three main lines by exploiting *i*) a prior contextual information to represent normalcy, *ii*) the employment of the OU mean reverting stochastic process to model the vessel's dynamics, which will be fully introduced in Chapter 3, and *iii*) the complete or incomplete observation of the available data.

In the first line of work, defined in Chapter 4, we seek to reveal anomalous deviations of a vessel from the nominal conditions while navigating, i.e., a change in the planned route and/or the nominal velocity profile. We also consider that, in order to hide this irregular behavior, possibly related to a suspicious activity, the vessel switches its AIS device OFF for a due time, at the end of which it tries to restore the previous nominal conditions. During this period characterized by a data gap, the decision that has to be taken is either declaring that a deviation happened or not, considering different levels of data unavailability: *i*) relying only upon two available AIS contacts, meaning the last contact before the AIS device shutdown and the first one after the AIS device reactivation, and *ii*) considering the extension to the scenario in which multiple heterogeneous contacts are available during the silent period.

The work proceeds in Chapter 5 by raising the issue that the AIS messages can be subject to intentional reporting of false information by the crew, or by external spoofers. More specifically, we address the problem of establishing whether a vessel is reporting adulterated dynamic information through AIS messages in order to hide its current planned route, in addition to possible deviation from the nominal conditions. The problem of revealing both the spoofing and stealth deviations anomalies is translated in a multiple hypothesis testing framework and counts on the support of reli-

able information from monitoring systems (coastal radars and space-borne satellite sensors).

The last line of the research work, conducted in Chapter 6, considers the worst-case scenario in terms of anomaly detection. In this regard, we move to the opponent vessel's side, and implement the least-detectable trajectory that an elusive vessel can follow during an AIS disablement. Specifically, the trajectory of the malicious vessel must turn out as stealth as possible to the anomaly detector implemented in Chapter 4, while the vessel turns the AIS device OFF, leaves the expected traffic route to approach a specific point (for instance a rendezvous point), stops there to perform some possible illegal exchanges, and finally returns to the nominal conditions reactivating the AIS. We assume that, during the data gap period, the vessel is aware of the acquisition time instants of the space-borne sensor has been observing it. Computing the worst-case in terms of surveillance capability comes in handy to point out the limitations of the proposed detection methodology, and consequently to facilitate future research works in seeking countermeasures to improve it and make it more robust.

1.5 Contributions

In Chapter 4, a novel maritime anomaly detector is proposed to reveal path deviations of ships during an intentional AIS device disablement, in order to possibly perform activities that would not be normally allowed. The problem has been treated within the statistical hypothesis testing framework, based on the GLRT for a Gaussian linear model, where the null hypothesis is that no deviation from the nominal condition occurred and the alternative hypothesis is that a deviation actually happened, during the time-frame characterized by the data gap. The hypothesis testing procedure is designed to identify changes in the OU long-run mean velocity parameter of the vessel, since any anomalous deviation will inevitably result in an unknown contribution to such a parameter. We provide closed analytical forms for the detection and false alarm probabilities of the hypothesis test. We demonstrate the effectiveness of the proposed detection strategy through experimental analysis within synthetic scenarios, where we considered the case of only two contacts available (the last contact before the AIS device shutdown and the first after the AIS device reactivation), and the use of multiple heterogeneous contacts associated with the vessel along its trajectory. The support

of multiple contacts compared to the case of only two contacts can lead to a sensible improvement of detection performance, however, scenarios in which it turns out to be degraded are also shown. Furthermore, the proposed anomaly detection strategy is successfully applied in the real-world major events discussed in Subsection 1.3.5: the anomalous fishing activities around the Galápagos EEZ, and the grounding of the container vessel *Ever Given* in the Suez Canal.

In Chapter 5, the goal is to jointly detect spoofing and/or surreptitious deviation, or, more specifically, determining whether the AIS data received from a vessel are trustworthy or not, and whether a deviation occurred or not, with the support of additional reliable information provided by surveillance systems. The proposed solution involves two detection strategies both designed in a hypothesis testing framework and based on the changes of the OU process velocity parameter. The first strategy splits the problem into two binary hypothesis tests via GLRT. Specifically, the first GLRT decides if AIS data are truthful or not, while the second decides if a deviation occurred or not, and finally the two decisions are combined through a decision rule named Coupled GLRT (C-GLRT). The second strategy is designed to tackle the multiple hypothesis test, relying on the Model Order Selection (MOS) methodology. The performance of both strategies has been analyzed in a synthetic maritime scenario by varying relevant operational parameters.

Finally, Chapter 6 formalizes the optimal opponent vessel's stealth trajectory planning as a non-convex optimization problem and considers the Kullback-Leibler (KL) divergence between the statistical hypotheses of the nominal and the anomalous trajectories as key performance measure. A computationally efficient technique, called Non-Convex Optimized Stealth Trajectory (N-COST) algorithm, is proposed to handle the resulting constrained optimization problem, where physical requirements on the mean velocity dynamics, the compliance of the trajectory with the underlying OU statistical model, the existence of a time instant allowing the rendezvous, as well as sea coast constraints are accounted for at the design stage. Interesting case studies, concerning both synthetic and real-world scenarios, are also investigated to prove the effectiveness of the proposed planning strategy.

Chapter 2

Literature review

This chapter gives a brief survey of related work on anomaly detection methodologies. The first part of the chapter lists the works that seek solutions for detecting anomalies in a wide variety of application domains. The second part focuses on the specific literature related to maritime anomaly detection, distinguishing between data mining and machine learning, and statistical approaches.

2.1 Anomaly detection methodologies

Anomaly detection typically refers to the problem of finding patterns in data not conforming to an expected behavior [50]. From image processing [53] and sensor networks [106] to biological data [133], several approaches have been established for detecting anomalies. Indeed, an anomalous traffic pattern in computer networks could mean that a hacked computer is sending out sensitive data to an unauthorized destination [126]. An anomalous mammographic image may indicate presence of malignant tumors [205]. Anomalies in credit card transaction data could indicate credit card or identity theft [18] or anomalous values from a space craft sensor could reveal a fault in a component of the space craft [89].

In cyber-physical systems the anomaly detection problem is usually referred to as *intrusion detection*. Within such a framework, the works conducted by Forti *et al.* [83–85] investigate the effects of signal attacks possibly combined with network deception attacks by injecting fake measurements on

stochastic cyber-physical systems. In particular, the random set paradigm is adopted in order to model the switching nature of the signal attack and the fake measurement injection via Bernoulli and/or Poisson random sets. The problem of jointly detecting a signal attack and estimating the system state in presence of fake measurements is then formulated and solved in the Bayesian framework.

Furthermore, the outlier trajectory detection problem is addressed by Zheng [233], that conducted a systematic survey on the major research regarding *spatial trajectories*, representing the mobility of various moving objects, such as people, vehicles, and animals. The outlier trajectories could be a taxi driver's malicious detour [52, 141] or unexpected road changes (due to traffic accidents or constructions).

Another direction is to detect traffic anomalies (rather than trajectory itself) by using many trajectories. The traffic anomalies could be caused by accidents, controls, protests, sports, celebrations, disasters and other events. Liu *et al.* [142] partition a city into disjointed regions with major roads and glean the anomalous links between two regions according to the trajectories of vehicles traveling between the two regions. They divide a day into time bins and identify for each link three features: the number of vehicles traveling a link in a time bin, proportion of these vehicles among all vehicles entering the destination region, and that departing from the origin region. The Mahalanobis distance is used to measure the extreme points (in the 3D space), which are regarded as outliers. Following the aforementioned research, Chawla *et al.* [51] proposed a two-step mining and optimization framework to detect traffic anomalies between two regions and explain an anomaly with the traffic flows passing the two regions. Pan *et al.* [172] identify traffic anomalies according to drivers' routing behavior on an urban road network. Here, a detected anomaly is represented by a sub-graph of a road network where drivers' routing behaviors significantly differ from their original patterns. They then tried to describe the detected anomaly by mining representative terms from the social media that people have posted when the anomaly was happening. Finally, Pang *et al.* [173, 174] use Global Positioning System (GPS) data from taxis to monitor the emergence of unexpected behavior in the Beijing metropolitan area, which has the potential to estimate and improve traffic conditions in advance. They adapt the likelihood ratio test to describe traffic patterns, partition a city into uniform grids, and count the number of vehicles arriving inside a grid over a time pe-

riod. The objective was to identify contiguous sets of cells and time intervals which have the largest statistically significant departure from the expected behavior (i.e., the number of vehicles).

2.2 Review of recent research works on automatic maritime anomaly detection

In recent years, anomaly detection in maritime traffic domain has attracted a significant body of research to the development of new methodologies which apply various data-driven solutions [189, 197, 210, 227].

2.2.1 Data mining and machine learning approaches

In the following we present some data mining and machine learning approaches that have been recently proposed for automatic anomalous maritime vessel behavior detection.

Pallotta *et al.* [169] proposed an approach based on an incremental variant of the DBSCAN algorithm to learn maritime traffic routes from vessel AIS data to detect anomalous vessel behaviors and predict vessels' future positions. The proposed approach is called TREAD. The same TREAD approach was used by Pallotta and Jousselme [170] to perform track association to vessel traffic routes for hierarchical maritime anomaly detection. In their approach, a tool implementing the TREAD approach is first used to extract vessel traffic routes from historical AIS data of areas of interest. Secondly, for associating and classifying tracks (i.e., partial trajectories), the same authors have adopted a hierarchical reasoning, where new tracks are first associated to existing routes based on their positional information only, and "OFF-route" vessels are detected as vessels not following an existing route. Then, for on-route vessels (vessels following an existing route) further anomalies are detected such as speed anomaly and heading anomaly.

The works by Ardeguas *et al.* [22, 23] proposed a methodology to automatically generate graph-based maritime networks, inferring vessel behaviors revealed in historical traffic data to represent maritime traffic patterns. TREAD is adopted to extract entry/exit gates, ports, and routes. The methodology consists of *i)* a hierarchical graph-based extraction algorithm and *ii)* an automatic maritime lane learning method. Similarly, both ports and entry/exit gates are learned using an incremental DBSCAN algorithm.

Vespe *et al.* [218] proposed a vector-based representation of maritime traffic network, where trajectories can be thought of as a set of lines (sea lanes) connecting nodes (turning points/ports and offshore platforms/entry/exit points), extracted using the k-means algorithm. The proposed approach utilizes AIS data to incrementally mine motion patterns using unsupervised learning without any specific a priori contextual knowledge or labelled traffic data.

Nevell [162] developed a methodology to automatically detect and model the ocean traffic using a node-sparse network, where each landmass in the World is characterized by a closed polygon area and the Great Circle links connect between pairs of nodes that are not blocked by any landmass. This work focuses on the traffic network design, meanwhile Bayesian methods are applied to estimate the probabilities for journey destinations and then used to detect kinematic anomalies like changes of destination, and inconsistent and unexpected routings. Likewise, a similar representation was proposed by Lane *et al.* [127] to construct maritime traffic network, where the waypoints near landmasses are considered as network nodes and ocean voyages are represented with Great Circle routes. This work conceived that an overall threat is indicated by a sequence of the individual behaviors, and introduced five specific anomalies.

Zhen *et al.* [232] use the k-medoids algorithm to cluster ship trajectories and a Bayesian network to detect abnormal vessels.

As introduced in Chapter 1, Coscia *et al.* [61,62] and Millefiori *et al.* [153] proposed the OU processes to establish the graph-based representation of the maritime traffic, called MTG. In the MTG representation, the graph vertices represent clusters of waypoints, which are connected together by a network of navigational sealanes, i.e., graph edges. Multiple OU processes are adopted to define vessel motion dynamics which are combined with a change detection procedure to discover regions where vessels recurrently exhibit sailing state changes. The detected changes are then clustered in groups of similar elements to become the nodes of the graph using DBSCAN, which are robust against outliers, such as abnormal or nonstandard behaviors. Such methods work in an unsupervised way to automatically extract a graph-based model of maritime traffic routes, are computationally efficient, and fit well with big data processing models and paradigms. The established MTG is supportive for long-term prediction of future positions of ships at sea and allows for global surveillance of ships at sea.

The work by Wang *et al.* [224] has performed vessel route anomaly detection using DBSCAN considering Speed and Direction (DBSCANSD) [140], and parallel meta-learning algorithms on the Hadoop MapReduce environment cluster. According to their experiment, the accuracy and execution time of the proposed vessel route anomaly detector improve linearly with the number of nodes in the cluster.

Lei [135] proposed a framework for maritime trajectory modelling and anomaly detection. First, the author defines three outlying features of abnormal movement behaviors: *spatial*, *sequential*, and *behavioral* features extracted from the sequential occurrence in trajectory data collection, and then builds a maritime trajectory model for anomaly detection. Finally, the author develops an anomaly detection algorithm based on the learned trajectory model, in which an indicator is used to evaluate suspicious behavior and scores trajectory behavior according to the defined outlying features.

De Vries and Van Someren [69] apply different trajectory alignment kernels (Dynamic Time Warping and Edit Distance) with one-class SVM for detecting the outlying trajectories. However, this trajectory-based method is not applicable for real-time AIS surveillance.

To reduce the rate of false alarm in anomalous maritime vessel behavior detection, Radon *et al.* [180] have proposed an approach in two phases. In the first phase, the normal vessel movement patterns are extracted from historical vessel tracks (trajectories within a particular origin and destination) using trajectory partition, the DBSCAN, and line segment clustering algorithms. In the second phase, potential anomalies are detected and then real anomalies are filtered out from potential anomalies based on contextual verification. In the same line of false alarm rate reduction, the works by Laxhammar and Falkman [130,132,186] and by Smith *et al.* [199] have used an anomalous vessel behavior detection approach based on conformal predictors [221] known as *conformal anomaly detection*. The conformal anomaly detection approach assumes that all the training data are generated by the same probability distribution and makes no further assumptions about the distribution. For each new test data, the likelihood that it is coming from the same probability distribution as the training data is estimated. If this likelihood is below the specified confidence level, then the new object is declared as an anomaly. The advantage of the conformal anomaly detection approach is that the rate of false alarm is well-calibrated and it is at most equal to the specified confidence level.

Eitienne *et al.* [74] extract trajectories of mobile objects following the same itinerary from the spatio-temporal database and cluster them. A statistical analysis on this set of trajectories lead to spatio-temporal patterns such as the main route and spatio-temporal channel followed by most of trajectories of the set. Using these patterns, unusual situations can be detected.

Soleimani *et al.* [204] have proposed a new method for finding anomalies in ships' movements. The proposed method analyses the trajectory of ships from a geometrical perspective. The trajectory of the ship is compared with a near-optimal path that is generated by a graph search algorithm. To make this comparison the proposed method extracts some scale-invariant features from the real trajectory and from the optimal movement pattern, and then it compares the two sets of features to generate an abnormality score.

Zhao and Shi [231] present a method that determines the parameters of the DBSCAN algorithm through statistical analysis, and then use the results of clustering as the traffic patterns to train a recurrent neural network; the neural network is then applied as a vessel trajectory predictor to conduct real-time maritime anomaly detection. While Zhong *et al.* [234] use a random forest machine learning algorithm to classify vessels using AIS data streams.

Xu *et al.* [228] used a back propagation neural network to predict vessel trajectories. Daranda [68] built a three-layer back propagation neural network model to learn the clustered turning points and used it to predict marine traffic. Handayani *et al.* [107] presented the use of SVM to establish the vessel behavior classification model for anomaly detection.

Finally, Singh and Heymann [198] use real-world AIS data to analyze the possibility of successful detection of various anomalies in the maritime domain; the work proposes a multi-class artificial neural network-based anomaly detection framework to classify intentional and non-intentional AIS ON-OFF switching anomalies.

Ristic *et al.* [188] designed a grid-based approach to describe the maritime traffic pattern, where the target surveillance area is discretized as non-overlapping grids. The Poisson point process is adopted to model the number of AIS messages that fall inside each grid. The velocity distribution is estimated by KDE. The established maritime knowledge is then applied for anomaly detection using decisive thresholds to differentiate the anomalous states from a range of acceptable values. The feasibility of the method is verified with the real-world AIS data collected for Sydney Harbour.

Osekowska *et al.* [164, 165] proposed a grid-based approach that employs

the concept of artificial potential fields to express maritime traffic patterns. The general idea is to aggregate all the charge potentials that are related to the AIS records over a geographical grid to characterize the typical traffic patterns. The potential can be regarded as a function of the values of traffic density, course and velocity distribution along grid-based locations. The method is used to support maritime traffic visualization, route planning and anomaly detection. The authors also provided the discussion on grid optimization and implemented a web-based anomaly detection system, called STRAND (Seafaring TRansport ANomaly Detection).

The works by Rhodes *et al.* [184,185] presented a neural network classifier known as fuzzy Adaptive Resonance Theory Map to evaluate the behaviors of vessels. Bomberger *et al.* [37] developed a method based on associative learning and neural networks to predict future vessel behaviors and detect abnormal vessels.

Lei *et al.* [136] proposed a grid-based framework of Maritime Traffic Route Discovery (MTRD) to generate the knowledge of movement pattern, which is able to contribute to maritime traffic management, situation awareness in maritime surveillance and anomaly detection. MTRD comprises of three modules: *i) AIS Pattern Mining* module to identify the trajectories with similar movement behavior, *ii) Pattern Summarization* module to recapitulate the generated patterns and *iii) Traffic Route Generation* module to derive the pattern-aware routes.

Finally, Venskus *et al.* [217] proposed an adaptive self-learning algorithm based on a bio-inspired virtual pheromone method for the detection of anomalous maritime traffic. The proposed algorithm first learns normal vessel traffic pattern grids from vessel movement data provided by the surveillance data fusion subsystem, then detects a potential incident where the moving vessel path does not match with learned pattern grids. The authors have verified that the algorithm provides rapid self-learning and fast adaptation characteristics.

2.2.2 Statistical approach

Laxhammar [129] uses the GMM and a greedy version of the EM algorithm to extract normalcy by clustering. The work conducetd by Ristic *et al.* [187] proposes to use an adaptive KDE for estimating unknown probability densities and modelling arbitrary sea lanes to derive a normal traffic model; then particle filtering is used to predict the positions of vessels based on the

derived density. In the anomaly detection phase, the anomaly detector is sequentially applied to the incoming data. The value of new incoming point's density is calculated under the null hypothesis (no anomaly) and this value is then compared with a detector parameter related to false alarms for deciding the new point's abnormality. A comparison between the two approaches above is given in Laxhammar *et al.* [131], demonstrating that the anomaly detection results from both models are not satisfactory. As the two models detect the anomalous segments at a significant distance from the point where anomaly behavior happens (three kilometer and four kilometer respectively) while an expected effective anomaly detector should detect such behaviors at a shorter distance.

Osekowska *et al.* [166] reported a maritime traffic modeling study characterizing the fluctuations of maritime traffic patterns. Such fluctuations represent the unforeseen shifts in statistical properties of traffic pattern over time, based on one year's AIS vessel tracking data. The outcomes offer insights into regular and reoccurring drifts as well as irregularities existed in the traffic data.

Anneken *et al.* [21] used a non-simulative anomalies dataset based on real-world data, to provide an evaluation and comparison of maritime anomaly detection algorithms, specifically the GMM and the KDE.

The method proposed by Kowalska and Peel [123] foresees a data-driven non-parametric Bayesian model, based on Gaussian Processes, to model normal shipping behavior; this model is learned from AIS data and uses an Active Learning paradigm to select an informative sub-sample of the data to reduce the computational complexity of training; the resulting model allows a measure of normality to be calculated for each newly-observed transmission.

The works by Smith *et al.* [200,201] have introduced an anomaly detection technique using a combination of Gaussian Processes, extreme value theory and divergence measurement, to identify anomalous vessel behavior in both streaming and batch data.

Mascaro *et al.* [149] addressed the anomalous vessel behavior detection issue using Bayesian Networks. By training Bayesian Networks with real-world AIS data combined with additional real-world contextual information such as weather and time, as well as vessel interactions, they have taught two Bayesian Network-based models at two different time scales, namely in the form of the time series and track summary models. As a training tool,

the authors used a machine learner named CaMML (Causal discovery via Minimum Message Length). The authors have found that adding further real-world attributes helped to create a better model of normalcy.

To preserve port safety by classifying the movements of different ships in the scene, Castaldo *et al.* [49] have applied Bayesian Networks to the analysis of vessel behavior and ship-to-ship interactions in port areas.

Statistical solutions, such as the extended Kalman filter and particle filter, have been used by Perera *et al.* [177] to reconstruct trajectories of vessels; whereas Bayesian networks were applied in Hruschka *et al.* [110] for tackling missing values (anomalies) in a dataset for prediction and classification purposes.

The work by Guerriero *et al.* [105] attempts to identify the AIS ON-OFF switching, through the introduction of the notion of channel memory and the application of hidden Markov models, while Mazzarella *et al.* [151] address the same problem exploring a solution that exploits the received signal strength indicator available at the AIS base stations.

The work by Balduzzi *et al.* [25] presents a method of spoofing a vessel by injecting invalid data into AIS gateways in order to trigger fake collision warning alerts, and potentially make surrounding vessels alter their course. This approach makes it possible for attackers to alter any information broadcast by existing AIS.

Ray *et al.* [182] propose an architecture capable of identifying spoofing messages in AIS after those have entered the database system, by using a methodology based on integrity and quality assessment of the AIS messages. Whereas, Katsilieris *et al.* [117] focus on the problem of trustworthiness of the received AIS data with the help of radar measurements and information from the tracking system.

Furthermore, an innovative statistical framework that combines contextual information to define the nominal behavior at sea with the OU model to represent the vessel dynamics has been recently proposed by Forti *et al.* [79, 81, 82, 88]. Here, the formulation of maritime anomaly detection is based on a probabilistic joint anomaly detection and tracking methodology for sequential detection of maritime anomalous deviations and simultaneous vessel tracking where new contacts (e.g., from AIS, radar, SAR) are periodically available based on surveillance coverage and reporting frequencies. This is motivated by the fact that maritime anomaly detection would ideally be performed in real-time, as new observations become available.

2.2.3 Rule-based and hybrid approaches

Although most of the recent works on anomalous vessel detection are data-driven, some rule-based and hybrid approaches have been proposed.

Maggi *et al.* [143] have demonstrated the use of process mining for analyzing the behavior of vessels, and for detecting anomalies, i.e., deviations from the normal behavior. To model the behavior of vessels, the authors have used the *Declare language*, which is a constraint-based language that can specify complex behavior in terms of a couple of restrictions. Declare combines a formal semantics for analysis purposes with a graphical representation for users.

Brax *et al.* [42] presented a Multi-Agent System (MAS) dedicated to anomalous behavior detection and alerts triggering in the maritime surveillance area. Anomalies are detected based on the implementation of maritime regulation.

Vandecasteele *et al.* [215] proposed a semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behavior analysis. For maritime anomaly detection, the authors have used spatial ontology for modelling a maritime operator's knowledge.

Shahir *et al.* [195, 196] combined data-driven machine learning methods with maritime background domain knowledge to detect anomalous vessel interaction patterns. In Kazemi *et al.* [120] a hybrid framework for anomaly detection based on the use of open data (contextual information and background knowledge) is proposed. Based on the proposed anomaly detection framework and the algorithms implementing the expert rules, the Open Data Anomaly Detection System (ODADS) was developed.

Chapter 3

Ornstein-Uhlenbeck stochastic mean-reverting process

This chapter is merely introductory and aims at presenting a novel method for predicting long-term target states based on mean-reverting stochastic processes, that will be used in the next chapters. The dynamics of non-maneuvering targets, such as vessels under way in open sea, is traditionally modeled with a white noise random process on the velocity, usually referred to as nearly-constant velocity model. This model is shown to be unrealistic for a significant portion of maritime ship traffic, when two consecutive contacts are several hours distant in time, since vessels under way tend to adjust their speed continuously around a desired value. A method to predict long-term vessel states has been developed based on the OU mean-reverting stochastic process, leading to a revised target state equation and to a completely different time scaling law for the related uncertainty, which in the long term is shown to be orders of magnitude lower than nearly-constant velocity model assumption. [1]

[1] This chapter has been implemented courtesy of L. M. Millefiori and originally published as: L. M. Millefiori, P. Braca, K. Bryan and P. Willett, "Modeling Vessel Kinematics using a Stochastic Mean-Reverting Process for Long-Term Prediction" in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2313-2330, October 2016 [154].

3.1 Introduction

Ship traffic monitoring is a foundation for many maritime security domains, and modern monitoring system specifications and requirements reflect the need for an extended and continuous ability to track vessels beyond territorial waters and over several sensor coverage areas. However, vessels in open seas are seldom continuously observed by monitoring sensors and even the data coming from self-reporting systems is often highly intermittent.

The problem of long-term vessel state estimate and prediction is therefore crucial. Unfortunately, this issue has been overlooked in the target tracking literature, and only few works partially address the problem, e.g., see [169,187,218], while most of the literature is focused on maneuvering target models, e.g., see [138, 139, 223]. The analysis of real-world self-reported data [i.e., largely AIS] shows that a significant portion of the vessels in open seas maneuver very seldom. In the literature, non-maneuvering target dynamics are modeled with a velocity that is perturbed by a white noise process. This model is often referred as Nearly Constant Velocity (NCV) [29, 138] and has been successfully used in several target tracking applications, such as radar [145] and sonar [41] where the prediction step always refers to the very near future, generally one sensor time-scan ahead. The NCV model is adopted also in [187] for anomaly detection and motion prediction.

The work in [154] proposes a novel method for the long-term prediction of target states based on the OU stochastic process, which leads to a revised target state equation and to a completely different time scaling law for the related uncertainty. This novel formulation reduces of order of magnitude the uncertainty region of the predicted position with respect to the models available in literature.

This aspect is crucial for several applications. In SaR operations, a smaller uncertainty region implies a smaller search region, which can significantly improve the probability of success for search cases. For instance, let us consider a vessel having an accident in a region with intermittent AIS coverage (e.g., open sea); the position of the accident is consequently unknown. Not having any information other than its last observed position and the time of accident (e.g., time of SOS message), the only possibility is to hypothesize that the ship had been moving from the last observation to the position of the accident in a straight line. It is important to notice that this assumption has to be made whether the traditional or the proposed approach

is taken, equally; the important difference is that using the proposed method the search area (the uncertainty region) would increase linearly –instead of quadratically– in proportion to the time from last report, for a given level of confidence. In the counter-piracy framework [102], vessels can keep their AIS transmitters off for several hours when sailing in dangerous areas, in order not to be detected by pirates: the use of the proposed modeling would provide a good clue about their position even when their last broadcast message is several hours old. Also the association of SAR vessel detections with AIS contacts would greatly benefit from the proposed modeling in terms of reduction of ambiguities because of the narrower gating region, especially in open sea where AIS coverage might be poor. From another standpoint, a smaller uncertainty region of the long-term prediction may reduce the probability of missing a vessel of interest in a high-resolution SAR acquisition, that has to be planned several hours ahead of the platform passing above the target.

The OU model is popular in various and heterogeneous scientific fields, spanning from physics [72,94,128,212] to finance [30,55], and biology [36,211]; but is much less popular within the tracking community [208]. The OU stochastic process was first introduced in physics [72, 128, 212] to describe the velocity of a Brownian particle under the influence of friction. It can be seen as a modified Wiener process so that there is a tendency of the walk to move back towards a central location, with a greater attraction when the process is further away from the center. More recently, OU processes have been used in finance, e.g., [30], to model the instantaneous variance of stock price volatility, and more recently in the context of high-frequency trading [55]. In biology, the OU process has been adopted as a baseline model for animal movement [36]. In [211] the authors considered a spatial neuron model in which the membrane potential satisfies an equation with an input current which is a dynamic random process of the OU type. In the tracking literature the OU model has been discussed mostly notably by Coraluppi and Carthel in [56–60], where the stability of the OU, and the so-called Mixed Ornstein-Uhlenbeck (MOU) processes are studied.

However, all of these works, including [208], deal with *zero-mean-reverting* processes, i.e., the typical velocities are null, being the aim not the long-term target state prediction, but rather the short-term characterization of the target dynamics. Indeed, in [56] authors are more interested in the boundedness of the target components, and for this reason the MOU is constructed in or-

der to be a zero-mean-reverting process for both position and velocity.[2]

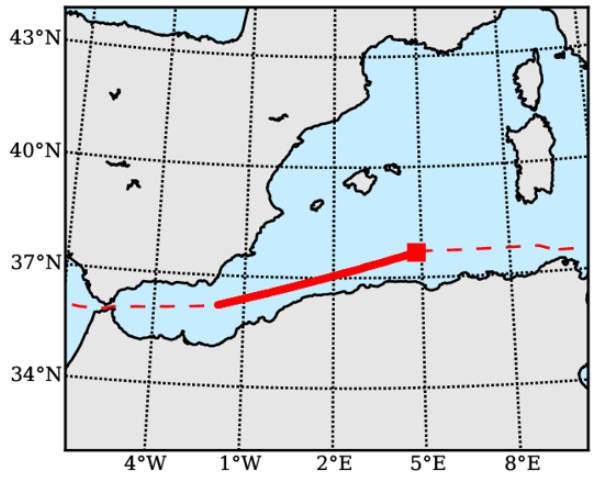


Figure 3.1: Trajectory of the target under consideration: a tanker vessel of 270 meters in length and 48 in breadth sailing northeastwards from the Gibraltar Strait to East Mediterranean. The dashed line represents the entire trajectory of the vessel, but only the highlighted portion has been considered, in order for the non-maneuverability hypothesis to be satisfied. The red square denotes the last contact in the sequence considered.

Different from [56–60,208], the work in [154], from which this chapter has been mainly drawn, focuses on the long-term prediction of non-maneuvering vessels, such as those under way in open sea. The thesis in [154] is that the NCV model might be implausible for a significant portion of maritime ship traffic, as vessels under way tend to continuously adjust their speeds around a desired operating point. This fundamental difference between the OU and NCV models is well explained intuitively by the case of study shown in 3.1 and 3.2, where a real-world vessel is sailing in the Mediterranean Sea with a desired speed. In the uppermost plot of Fig. 3.2 the actual velocity of the vessel is reported, while the lowermost plots contain one hundred realizations of the velocity using simulated OU and NCV models having the

[2]In [56–60, 208], the OU is a process in which the target position follows an OU dynamic, while the Integrated Ornstein-Uhlenbeck (IOU) is a process in which the target velocity follows an OU dynamic while the position is integrated. We follow a different nomenclature, used in the statistic literature [30], in which velocities are OU and then positions are IOU.

same initial point as the real trajectory. The velocity samples drawn from the OU process are bounded around the average value, as indeed are the actual velocity samples, while in the simulated NCV case the velocity apparently tends to diverge. In other words, in the long-term the NCV model does not seem to be well-suited for the representation of the vessel velocity.

Supported by real-world vessel traffic data, the main result presented in [154] is that mean-reverting processes can be used to model non-maneuvering vessel movement. Specifically, [154] provides evidence that the vessel velocity is well-described by a OU stochastic process, and consequently the vessel position by a IOU process. As a consequence, after an initial transient, the vessel position is mathematically equivalent to a Brownian particle motion. It is also shown that the popular NCV model that is commonly adopted in the target tracking literature is not well-suited for the characterization of the uncertainty of the long-term target state prediction. While it is sufficiently accurate for short-term predictions (meaning: typically the case for traditional target tracking applications) the NCV model can overestimate the actual uncertainty of long-term predictions, even to orders of magnitude.

3.2 Vessel dynamic models

This section describes the stochastic models of the vessel (or target) dynamics. The four-dimensional target state at time $t \in \mathbb{R}_0^+$ is indicated with

$$\boldsymbol{s}(t) \triangleq [\boldsymbol{x}(t),\, \dot{\boldsymbol{x}}(t)]^{\mathrm{T}}\,, \tag{3.1}$$

where $\boldsymbol{x}(t)$ and $\dot{\boldsymbol{x}}(t)$ denote the target position and velocity, respectively, in a two-dimensional Cartesian reference system

$$\boldsymbol{x}(t) \triangleq [x(t),\, y(t)] \;\;,\;\; \dot{\boldsymbol{x}}(t) \triangleq [\dot{x}(t),\, \dot{y}(t)]\,. \tag{3.2}$$

The choice to define the target state in the Cartesian coordinates is standard in the target tracking literature, e.g., see [29]. In this formulation of the problem, (x, y) can be either the Universal Transverse Mercator (UTM) coordinate system, or the rotated coordinate along the target trajectory as usually assumed in the knowledge-based tracking, e.g., see [34, 46, 220].

The target dynamics are generally modeled by a set of linear stochastic differential equation (SDE)s [138]:

$$\mathrm{d}\boldsymbol{s}(t) = \mathbf{A}\,\boldsymbol{s}(t)\,\mathrm{d}t + \mathbf{D}\,\boldsymbol{u}(t)\,\mathrm{d}t + \mathbf{B}\,\mathrm{d}\boldsymbol{\omega}(t), \tag{3.3}$$

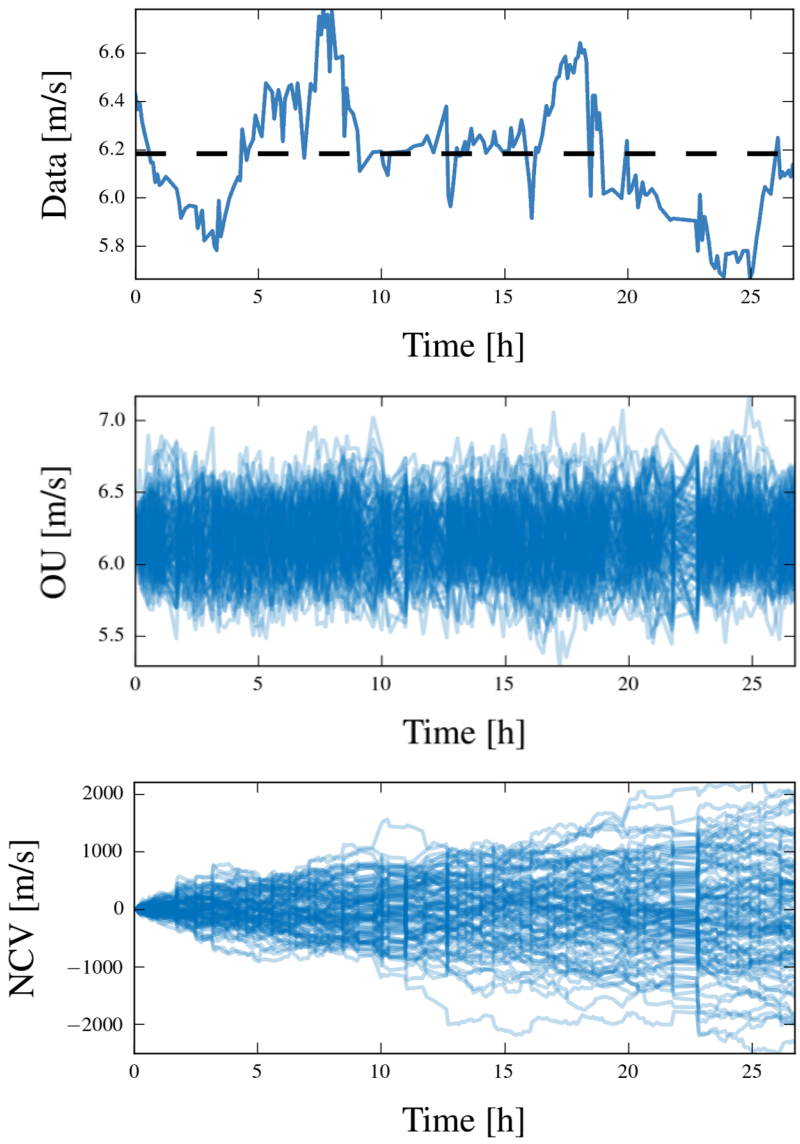


Figure 3.2: Single-component velocity of the real-world vessel trajectory illustrated in 3.1: the uppermost plot shows the actual velocity samples; the middle plot is for a simulated OU process; the lowermost plot illustrates a simulated NCV process. Simulated processes use as initial state the first velocity sample of the real-world trajectory; the process parameters are estimated from the data series shown at the top of this figure.

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ are constant matrices, $\boldsymbol{u}(t)$ is a deterministic function, and $\boldsymbol{\omega}(t)$ is a standard bi-dimensional Wiener process. The SDE can be solved by the use of Itô calculus [163].

Given the state of a target $\boldsymbol{s}(t_0)$ observed at the time t_0, the prediction of its state at the time t is carried out by the optimal Bayesian estimator:

$$\begin{aligned}\boldsymbol{s}\left(t|t_0\right) &\triangleq \left[x\left(t|t_0\right), y\left(t|t_0\right), \dot{x}\left(t|t_0\right), \dot{y}\left(t|t_0\right)\right]^{\mathrm{T}} \\ &= \mathbb{E}\left[\boldsymbol{s}(t)|\boldsymbol{s}(t_0)\right],\end{aligned} \tag{3.4}$$

where $\mathbb{E}[\cdot]$ indicates the expected value operator. The estimator is highly dependent on the underlying motion model described by the SDE (3.3).

It is worth mentioning that in the tracking literature the target state observation is typically affected by noise. The work in [154], and reported in this chapter, assumes to observe directly the target positional state and focuses only on the case of non-maneuvering vessels while under way.

3.2.1 NCV model

One of the most popular target motion models, commonly adopted in the scientific target tracking literature, is the NCV model [138], where (3.3) has the form

$$\mathrm{d}\boldsymbol{s}(t) = \mathbf{A}\,\boldsymbol{s}(t)\,\mathrm{d}t + \mathbf{B}\,\mathrm{d}\boldsymbol{\omega}(t), \tag{3.5}$$

with

$$\mathbf{A} = \begin{bmatrix}\mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix}\mathbf{0} \\ \boldsymbol{\Sigma}\end{bmatrix}, \tag{3.6}$$

being $\mathbf{I}$ the bi-dimensional identity matrix, $\mathbf{0}$ the bi-dimensional null matrix and $\boldsymbol{\Sigma}$ a bi-dimensional matrix defining the noise process. In practice, the equation for the target dynamics relies on the fact that, for non-maneuvering vessels $\ddot{\boldsymbol{x}}(t) \approx [0, 0]^{\mathrm{T}}$, i.e., there is a "small" effect on $\dot{\boldsymbol{x}}(t)$ that accounts for unpredictable modeling errors [138].

3.2.2 OU model

For the OU model the SDE has a slightly different form, with an additional term that accounts for the mean-reverting tendency of the velocity:

$$\mathrm{d}\boldsymbol{s}(t) = \mathbf{A}\,\boldsymbol{s}(t)\,\mathrm{d}t + \mathbf{D}\,\boldsymbol{v}\,\mathrm{d}t + \mathbf{B}\,\mathrm{d}\boldsymbol{\omega}(t), \tag{3.7}$$

where $\boldsymbol{v} = [v_x, v_y]^{\mathrm{T}}$, and $\boldsymbol{\omega}(t)$ is a standard bi-dimensional Wiener process. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ are defined as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -\boldsymbol{\Theta} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Sigma} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Theta} \end{bmatrix}, \tag{3.8}$$

where $\boldsymbol{\Sigma}$ is a bi-dimensional matrix defining the noise process and $\boldsymbol{\Theta}$ is a bi-dimensional matrix quantifying the mean-reversion effect, meaning the rate at which the target will tend back to the desired velocity after a perturbation; its diagonal terms refer to the x and y components, while the off-diagonals quantify the coupling effect. The diagonal terms of $\boldsymbol{\Theta}$ represent the mean reversion effect along the x and y components, respectively, while the off-the-diagonal elements are representative of the coupling effect between them. Assuming that $\boldsymbol{\Theta}$ is diagonalizable and has positive eigenvalues, an affine transformation can be found that projects the matrix $\boldsymbol{\Theta}$ onto another space, i.e., $\boldsymbol{\Theta} = \mathbf{R}\boldsymbol{\Gamma}\mathbf{R}^{-1}$, where $\mathbf{R}$ is the matrix whose columns contain the eigenvectors of $\boldsymbol{\Theta}$ and $\boldsymbol{\Gamma}$ is a diagonal matrix whose elements are the corresponding eigenvalues. This idea is further expanded in Appendix B.1, where the general solution to the coupled problem is also provided.

Equation (3.7) has the form of a Langevin dynamic [128] and can be solved in closed form by using Itô calculus [94, 163], see details in Appendix B.1. The $\dot{\boldsymbol{x}}(t)$ process is said to be of OU type [30, 212], and correspondingly, we say that $\boldsymbol{x}(t)$ is an IOU process [30]. The parameters v_x and v_y in $\boldsymbol{v} = [v_x, v_y]^{\mathrm{T}}$ play a key role in the proposed model because they represent the *typical* velocities along x and y, respectively, of the vessel on the trajectory under consideration. Roughly speaking, the velocity of the process tends to drift over time towards its long-term mean; and the mean-reversion tendency is stronger when the velocity is further away from that long-term mean. In [56, 59, 208], the OU zero-mean-reverting process is introduced, in which the parameters v_x and v_y are both null. Furthermore the Singer model [138] provides for the target acceleration to be modeled as an OU zero-mean-reverting process. However, in both cases the aim is not the long-term target state prediction, but rather the short-term characterization of the target dynamics.

3.2.3 Prediction procedure

The solution of the SDE provides for the target state prediction $\boldsymbol{s}(t|t_0)$ and the related variance, which we will take as a measure of the prediction

uncertainty. This subsection describes the prediction procedure in the two cases that the OU and NCV models are assumed for the target velocity.

NCV model

Assuming the NCV model for the velocity of the target, the optimal prediction, given the initial target state $\boldsymbol{s}(t_0)$, is the following [29]

$$\boldsymbol{s}(t|t_0) = \boldsymbol{F}(t-t_0)\boldsymbol{s}(t_0), \tag{3.9}$$

where $\boldsymbol{F}(t)$ is often referred to as the state transition matrix and is given by

$$\boldsymbol{F}(t) = \begin{bmatrix} \mathbf{I} & t\,\mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{3.10}$$

According to [60], the covariance matrix of the estimator is provided by the solution provided by Itô calculus and is given by

$$\mathrm{Cov}\left[\boldsymbol{s}(t)|\boldsymbol{s}(t_0)\right] = \begin{bmatrix} \frac{(t-t_0)^3}{3} & \frac{(t-t_0)^2}{2} \\ \frac{(t-t_0)^2}{2} & t-t_0 \end{bmatrix} \otimes \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}}, \tag{3.11}$$

where

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}. \tag{3.12}$$

A slightly different process noise assumption of the NCV (see [29]), would lead to the same estimator (3.9) but with a higher scaling law of the errors, i.e., proportional to $(t-t_0)^4/4$ instead of $(t-t_0)^3/3$ for the position and proportional to $(t-t_0)/2$ instead of $(t-t_0)$ for the velocity. In common tracking applications, where the time interval is fixed, the two models are quite similar because the parameters σ_x and σ_y can be differently tuned in order to compensate for the different scaling laws. Instead, in [154], even if the parameters were tuned differently, the two models would still behave differently in the long-term.

OU model

The case in which the velocity of the target follows the OU model is here specifically considered. It is assumed hereafter, for the sake of clarity, that $\mathbf{R} = \mathbf{I}$, so that $\boldsymbol{\Theta} = \boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$, with $\boldsymbol{\gamma} = [\gamma_x, \gamma_y]^{\mathrm{T}}$. The general solution is provided in Appendix B.1.

The optimal prediction, given the initial target state, is provided by the first moment of the SDE solution [30, 94] and, for the velocity, we have

$$\dot{\boldsymbol{x}}\left(t|t_0\right) = \boldsymbol{v} + \begin{bmatrix} \mathrm{e}^{-\gamma_x(t-t_0)} & 0 \\ 0 & \mathrm{e}^{-\gamma_y(t-t_0)} \end{bmatrix} \left(\dot{\boldsymbol{x}}(t_0) - \boldsymbol{v}\right). \tag{3.13}$$

Proceeding similarly for the target position, which is an IOU process, the following expression can be derived

$$\boldsymbol{x}\left(t|t_0\right) = \boldsymbol{x}(t_0) + (t-t_0)\boldsymbol{v} + \begin{bmatrix} \frac{1-\mathrm{e}^{-\gamma_x(t-t_0)}}{\gamma_x} & 0 \\ 0 & \frac{1-\mathrm{e}^{-\gamma_y(t-t_0)}}{\gamma_y} \end{bmatrix} \left(\dot{\boldsymbol{x}}(t_0) - \boldsymbol{v}\right). \tag{3.14}$$

The optimal prediction can be rearranged in the matrix form

$$\boldsymbol{s}\left(t|t_0\right) = \boldsymbol{\Phi}(t-t_0, \boldsymbol{\gamma})\, \boldsymbol{s}(t_0) + \boldsymbol{\Psi}(t-t_0, \boldsymbol{\gamma})\, \boldsymbol{v}, \tag{3.15}$$

where $\boldsymbol{\Phi}(t, \boldsymbol{\gamma})$ is the analog of the state transition matrix and $\boldsymbol{\Psi}(t, \boldsymbol{\gamma})$ is often called control input function. Their definitions are respectively provided by equations (B.6) and (B.8) in Appendix B.

The estimator covariance is provided in equation (B.12) in Appendix B. The variance terms are reported here:

$$\mathbb{E}\left[\left(x(t|t_0) - x(t)\right)^2 |\boldsymbol{s}(t_0)\right] = \frac{\sigma_x^2}{\gamma_x^3} f\left(\gamma_x(t-t_0)\right), \tag{3.16}$$

$$\mathbb{E}\left[\left(y(t|t_0) - y(t)\right)^2 |\boldsymbol{s}(t_0)\right] = \frac{\sigma_y^2}{\gamma_y^3} f\left(\gamma_y(t-t_0)\right), \tag{3.17}$$

$$\mathbb{E}\left[\left(\dot{x}(t|t_0) - \dot{x}(t)\right)^2 |\boldsymbol{s}(t_0)\right] = \frac{\sigma_x^2}{\gamma_x} g\left(\gamma_x(t-t_0)\right), \tag{3.18}$$

$$\mathbb{E}\left[\left(\dot{y}(t|t_0) - \dot{y}(t)\right)^2 |\boldsymbol{s}(t_0)\right] = \frac{\sigma_y^2}{\gamma_y} g\left(\gamma_y(t-t_0)\right), \tag{3.19}$$

where $f(t)$ and $g(t)$ are the prediction position and the velocity error *normalized* variance, defined in (B.15) and (B.16) in Appendix B.

Remark 3.2.1. *Interestingly, it is easy to show that when $\gamma_{x,y} = 0$ the two models are equivalent because they have the same dynamic expressions. Furthermore, by applying l'Hôpital's rule to (3.15) when $\gamma_{x,y} \to 0$, then $\boldsymbol{\Phi}(t-t_0, \boldsymbol{\gamma}) \to \boldsymbol{F}(t-t_0)$, while $\boldsymbol{\Psi}(t-t_0, \boldsymbol{\gamma}) \to \boldsymbol{0}$, and again, by applying l'Hôpital's rule the OU covariance (B.12) degenerates in the NCV covariance (3.11).*

Summing up, the SDE parameters in the NCV case are reduced to the process noise standard deviations $\boldsymbol{\sigma} = [\sigma_x, \sigma_y]^{\mathrm{T}}$. While, under the OU assumption on the target velocity, beyond $\boldsymbol{\sigma}$, we have two more parameters: the long-run mean velocity $\boldsymbol{v} = [v_x, v_y]^{\mathrm{T}}$, and the reversion rate $\boldsymbol{\gamma} = [\gamma_x, \gamma_y]^{\mathrm{T}}$.

In [154], a complete description of the procedure adopted to estimate the aforementioned SDE parameters for every given trajectory is provided. Moreover, the work conducted in [154] illustrates the validation methods used against the NCV and OU models, based on the capability of the model not only to predict future target states but also to quantify the uncertainty of the prediction.

3.3 Model validation using real-world vessel traffic data

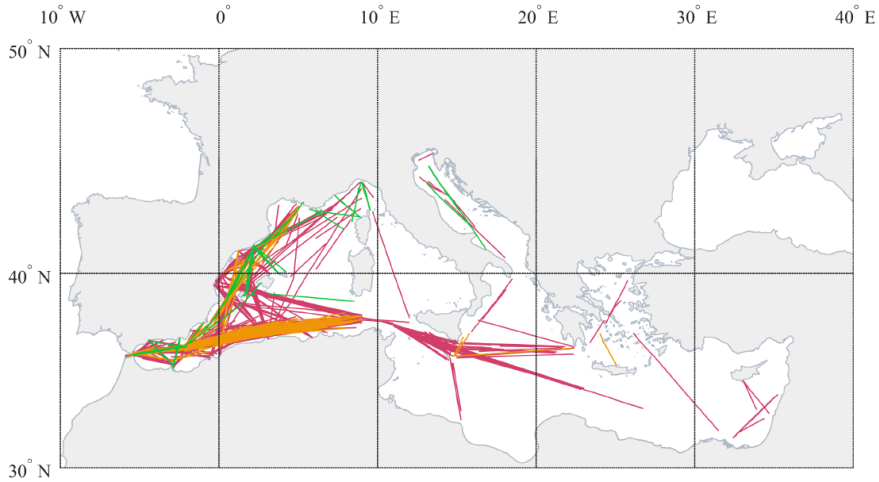


Figure 3.3: Quasi-rectilinear trajectories under consideration in the validation study; the color indicates the traffic category: pink for the cargo traffic, orange for the tanker and green for the passenger. These trajectories represent a subset of a data set of AIS messages collected over the Mediterranean Sea from multiple AIS networks during two months of 2014.

This section reports evidence, provided in [154], that the uncertainty of long-term state predictions of non-maneuvering vessels fits better to the OU (for the velocity) and IOU (for the position) models than what does the NCV

model. This evidence is based on the analysis of a significant record of all the maritime traffic in the Mediterranean Sea, collected by NATO Science and Technology Organization (STO)-Centre for Maritime Research and Experimentation (CMRE). Specifically, the data set consists of the entirety of AIS messages broadcast by vessels navigating in the Mediterranean Sea in two months of 2014 and collected by a network of satellite and terrestrial receivers.

Since the models described in Section 3.2 are valid under the assumption of a non-maneuvering vessel, a preprocessing phase is conceived to break every observed trajectory into "linear" piecewise parts where the target has essentially no process noise.[3] This initial step provides a set of observed trajectories of non-maneuvering vessels of each traffic category considered (cargo, tanker and passenger vessels). Fig. 3.3 illustrates the vessel trajectories that will be used afterwards in this section to perform the analysis of prediction errors, evaluating the actual prediction error variance against the theoretical models described in Section 3.2.

The prediction procedure is repeated for all the trajectories and for the different motion models, providing a collection of prediction errors relative to the target position and velocity, as described in [154]. The OU motion model for the velocity and its integrated version for the position, as described in Subsection 3.2.3, are characterized by three parameters for each coordinate: the noise level $\boldsymbol{\sigma}$, the desired speed $\boldsymbol{v}$, and the reversion rate $\boldsymbol{\gamma}$, which basically represents how quickly the target tends to restore its desired speed after a perturbation. The NCV motion model has instead just the noise level $\boldsymbol{\sigma}$. These process parameters are important to the specific realization but, more importantly, they are not known a priori and therefore have to be estimated [154]. The estimation of the process parameters is not error-free but, instead, introduces additional error to the prediction error. However, the following analysis shows a good match between the real-data and the theoretical curves meaning that the trajectory are sufficiently long to guarantee a parameter estimation with a negligible error.

[3]Being on a linear track does not necessarily imply that the vessel is not-maneuvering because, especially for long trajectories, it may have changed its velocity quite significantly. However, it can safely assumed that a vessel under way on a linear track is not a maneuvering target.

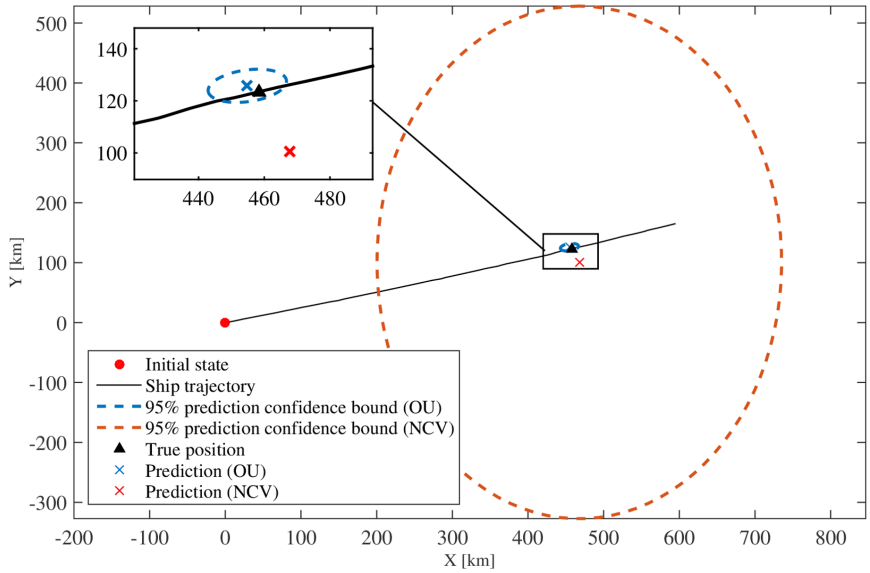


Figure 3.4: Single target scenario: comparison of the uncertainty of the predicted positional target state under the assumptions of NCV and OU model. The illustration reports the highlighted portion of the trajectory in Fig. 3.1 that has been converted into projected coordinates. The predictions of the target position after approximately 20 hours from the initial state are indicated with cross markers. The related 95%-confidence prediction covariance ellipses are plotted with dashed lines.

3.3.1 Single-target scenario

In this subsection a single trajectory is analyzed. A tanker vessel of 270 meters in length and 48 in breadth is sailing north-eastwards from the Gibraltar Strait to East Mediterranean, see Fig. 3.1. A prediction from the initial point of the trajectory, located in the origin, up to 20 hours is performed for both OU and the NCV and then compared with the true vessel position. Figure 3.4 reports the vessel trajectory, the OU and NCV predictions and their related uncertainty (95^{th} percentile) provided by both models. This case study shows that even if the NCV prediction is reasonably close to the true vessel position, its uncertainty is disproportionately large. On the other hand, the OU prediction is not only closer to the true vessel position with respect to the NCV prediction, but also has a significantly smaller uncertainty (orders of magnitude).

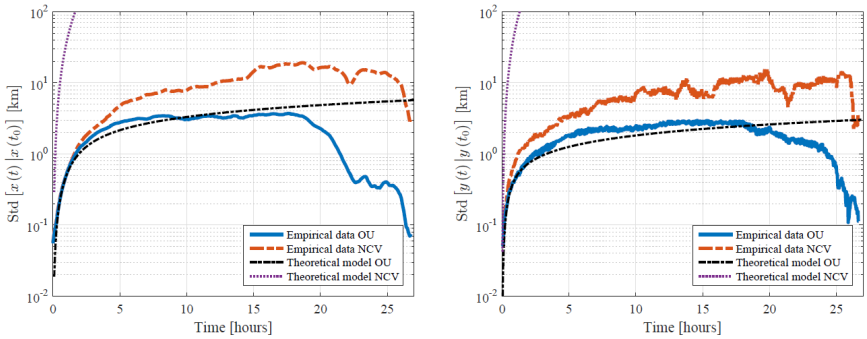


Figure 3.5: Single target scenario: unnormalized standard deviation of the prediction error on the target position assuming the OU or NCV motion model, for the two components of the positional target state (x on the left and y on the right). Theoretical curves are reported with dashed lines. The target is given by the tanker vessel whose trajectory is reported in Fig. 3.1.

The empirical standard deviation of the prediction errors along x and y with respect to the prediction time horizon is reported in Fig. 3.5 and compared with the theoretical curves. The empirical standard deviation curves are computed based on the data set related to this single trajectory, see details in [154]. The empirical curves are more accurate for short-term horizons (< 5 hours) because they are computed with more samples with respect to the larger time horizon. For instance, the last point of these curves has just a single sample related to the prediction from the initial point to the last point of the trajectory. It is worth noticing that the OU and NCV prediction errors are equal in first part (< 1 hour), while in the second part they separate and the NCV exhibits a larger error. Furthermore, as already discussed, it is easy to recognize that the NCV model in the very short-term (< 10 minutes) fits sufficiently well the data while in the long time the NCV theoretical curve significantly diverges from the empirical curve. The OU model instead follows the data until 15 hours, after this point the empirical curve is not statistically significant because of the lack of samples.

3.3.2 Multi-target scenario

To keep the notation lighter, let us define a generic coordinate $u = \{x, y\}$ and a generic velocity $\dot{u} = \{\dot{x}, \dot{y}\}$, along its corresponding long-term mean $v = v_{x,y}$, reversion rate $\gamma = \gamma_{x,y}$ and noise term $\sigma = \sigma_{x,y}$.

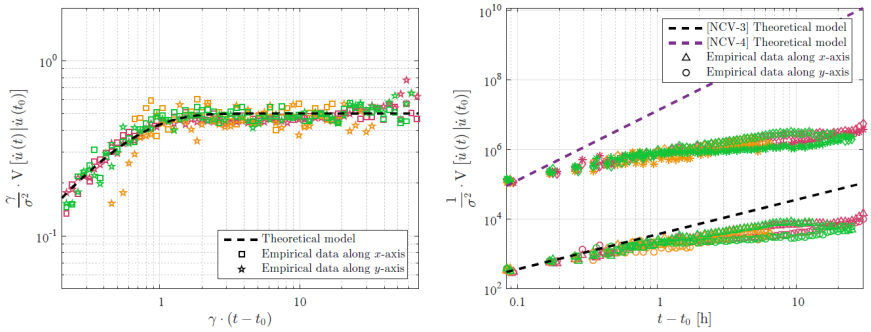


Figure 3.6: Normalized variance of the prediction error on the target velocity assuming the OU (left) or NCV (right) target motion model, for the three traffic categories: cargo (pink), tanker (green), and passenger (orange). Components along the x-axis and y-axis are identified using different markers.

Velocity

Figure 3.6 shows the prediction error variance on the target velocity over the prediction horizon, for the cargo, tanker and passenger traffic categories. In each subfigure marker plots illustrate the empirical data, i.e., the actual prediction error variance observed on the target velocity, whereas dashed lines represent the theoretical model. The aim of these figures is to compare how well the hypothesized models reproduce the evolution of the prediction uncertainty over the prediction horizon. In each of the aforementioned figures, the leftmost graphs are under the OU model hypothesis, while those on the far right relate to the NCV models, in both cases $p = 3$ and $p = 4$ (see discussion in Subsection 3.2.3 and [154]).

In Fig. 3.6 it is easy to recognize that while the NCV models fit the empirical curve only in the short time prediction the OU instead fits the whole evolution. It is worth mentioning that in the long time prediction the fundamental difference between the two model becomes clear: the NCV uncertainty diverges, as for a Brownian motion, while the OU reaches an asymptotic level provided by $g(t) \to 1/2$, see (B.16). The asymptotic unnormalized OU uncertainty is instead given by $\sigma^2_{x,y}/2\gamma_{x,y}$, basically it is proportional to the noise variance $\sigma^2_{x,y}$ and to the reversion rate $1/\gamma_{x,y}$. As already explained in Section 3.2, the NCV model is equivalent to the OU for $\gamma_{x,y} \to 0$, this explains from another point of view why the NCV uncertainty diverges.

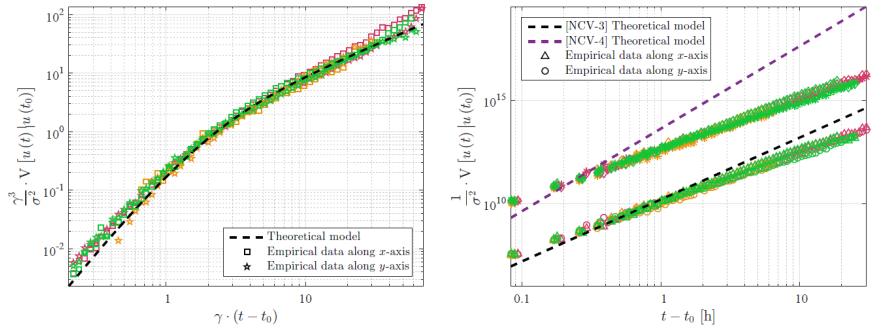


Figure 3.7: Normalized variance of the prediction error on the target position assuming the OU (left) or NCV (right) target motion model, for the three traffic categories: cargo (pink), tanker (green), and passenger (orange). Components along the x-axis and y-axis are identified using different markers.

Position

Figure 3.7 illustrates how the prediction error variance on the target position varies over the prediction horizon for the cargo, tanker and passenger traffic categories. As before, marker plots have been used for the empirical data and dashed lines for the theoretical models.

Like the OU model for the target velocity, the IOU is apparently more appropriate to model the uncertainty growth over the prediction horizon, especially for long-term target state prediction, where the actual prediction error variance on the position of the target is several orders of magnitude below the value one would expect according to both versions of the NCV.

As opposed to the velocity curve, the position uncertainty diverges for both IOU and NCV. However, the divergence rates are different, specifically the normalized IOU uncertainty grows as t while the NCV grows as t^p with $p = 3$ or $p = 4$. The growth rate for the unnormalized IOU uncertainty is ruled by σ^2/γ^3, as opposed to the asymptotic variance of the velocity, this being proportional to the cube of the reversion rate.

3.4 Conclusion

In this introductory chapter the problem of issuing long-term predictions of future target states has been presented, with specific focus on the modeling of the related uncertainty. Two different stochastic motion models have

been hypothesized, the first being the well-known Nearly Constant Velocity (NCV) model, and the second the less commonly studied – at least in the tracking literature – Ornstein-Uhlenbeck (OU) mean-reverting stochastic model. A description and comparison of the two theoretical models has been provided and it is accompanied by an extensive validation study on a real-world data set which is representative of a significant portion of the maritime traffic in the Mediterranean Sea. Experimental results confirm that OU stochastic processes may be used to model the motion of non-maneuvering vessels while under way. Its major advantage over the more traditional NCV model is that the variance of the predicted position grows linearly with the prediction horizon, resulting a prediction uncertainty that is much more contained in larger time scales.

Chapter 4

Detection of maritime anomalous deviations

In this chapter we present a novel anomaly detection procedure based on the OU mean-reverting stochastic process. The considered maritime anomaly is the deviation of a vessel from a planned route, with a possible change of the nominal velocity. In order to hide this behavior, possibly related to a suspicious activity, the vessel switches its AIS device OFF *for a given time, after which it tries to revert to the previous nominal conditions. During this period the decision that has to be taken is either declaring that a deviation happened or not, relying only upon two available AIS contacts. Furthermore, we also consider the extension to the scenario in which multiple contacts (e.g., radar) are available during the silent period. A proper statistical hypothesis test that builds on the changes in the OU process long-term velocity parameter of the vessel is the core of the proposed approach and enables the solution of the anomaly detection problem. We provide closed analytical forms for the detection and false alarm probabilities of the hypothesis test, and we demonstrate the effectiveness of the proposed detection strategy through experimental analysis within synthetic and real-world scenarios.* [1]

[1] This chapter has been mainly published as "Detecting Anomalous Deviations From Standard Maritime Routes Using the Ornstein-Uhlenbeck Process," in IEEE Transactions on Signal Processing, Dec. 2018 [64].

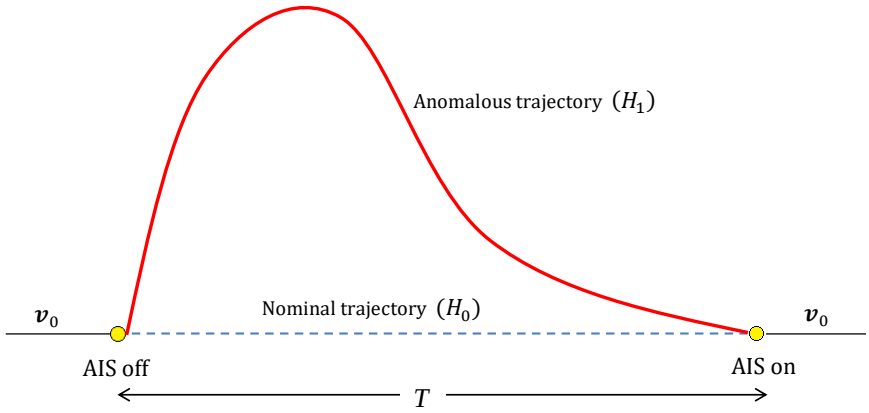


Figure 4.1: A vessel turns the AIS transponder OFF and follows a different trajectory from the nominal one, changing its velocity.

4.1 Introduction

In this chapter, we study the anomaly detection problem roughly depicted in Fig. 4.1, where a vessel deviates from its planned route, changing the nominal velocity $\boldsymbol{v}_0$. The vessel attempts to hide the deviation by switching its AIS transponder OFF for a time period T. After the deviation, the vessel would then try to revert back to the planned route and to the original nominal velocity $\boldsymbol{v}_0$. The decision that needs to be taken is whether a deviation happened or not, relying upon the available contacts (AIS, radar, etc.).

During the period of data gap the vessel might have been loitering or drifting for an unspecified (and unknown) amount of time, perhaps to encounter other ships; all actions commonly classified as anomalous [214]. As also documented in [115], a practical example could be a mother ship transporting a drug shipment from its place of origin to the waters off the country of the ultimate destination, where the drug is then transferred to a second vessel waiting in a pre-established location, which eventually brings it ashore.

Unlike other works [105, 117, 123, 127, 169, 187, 218], here the anomaly detection problem is addressed relying on a hypothesis test that builds on the changes of the OU process long-run mean velocity parameter, previously introduced in Chapter 3. Such a strategy will be tested against a trajectory of a real transshipment incident [17].

As previously discussed in Chapter 3, a ship motion model based on the

OU process has been shown to be more realistic than other conventional kinematic models for the behavior of the real-world commercial maritime traffic [61, 154, 155, 219]. In this framework, the OU model turns out to be a valuable tool when vessel information is not available, providing an accurate estimation of a ship's position and velocity, even after several hours.

The proposed detection strategy is investigated assuming to have multiple contacts available before and after the possible anomaly. We take into consideration the case where only two contacts are available, for example the last contact before the AIS device shutdown and the first one after the AIS device reactivation, which is of particular interest for real-world applications. Then, we extend our analysis to the use of multiple heterogeneous contacts associated with the vessel along its trajectory. This situation could be represented by a scenario where multiple radar contacts are available in addition to AIS contacts.

It is worth mentioning that, different from the AIS data, which contain vessel labeling information, other (especially non-collaborative) sensors (e.g. radar) suffer from the measurement-origin uncertainty [219]. In this chapter we assume that the association of contacts to the vessel of interest is solved in a preliminary stage, see e.g., [28, 219]. The possible association error, relevant when several multiple targets are close to each other (uncommon scenario in open sea), is neglected and left to future investigation.

The use of multiple contacts compared to the case of only two contacts can lead to a sensible improvement of detection performance. However, counterintuitively, there are scenarios in which the detection performance is degraded. Mathematical conditions and physical insights related to such scenarios are provided.

4.2 Problem formulation

Let us consider a vessel of interest, represented by a point in a bidimensional space, which is following its planned route. As already specified in Chapter 3, the position and velocity of the vessel are expressed in Cartesian coordinates, resulting from the projection of the geographic coordinates reported by the on-board AIS transponder.

Let us suppose that AIS data is unavailable for a time T after a given instant, due to a lack of communications from the ship (because of limited sensor coverage, interference, etc.) or an intentional shutdown of the AIS

transponder. In this scenario, two hypotheses can be envisioned:

1. Hypothesis H_0: the vessel is expected to navigate along a planned trajectory according to a piecewise OU model characterized by a sequence of N_0 long-run mean nominal velocities, comprised in the $2N_0$-dimensional vector $\boldsymbol{v}_0 = \text{col}\{\boldsymbol{v}_{0,n}\}_{n=1}^{N_0}$ of nominal velocities, identifying an N_0-section nominal path. Precisely, the nominal velocity $\boldsymbol{v}_0$ is assumed to be a known deterministic parameter of the OU process, which may be available from contextual information such as historical patterns extracted in the form of MTG;

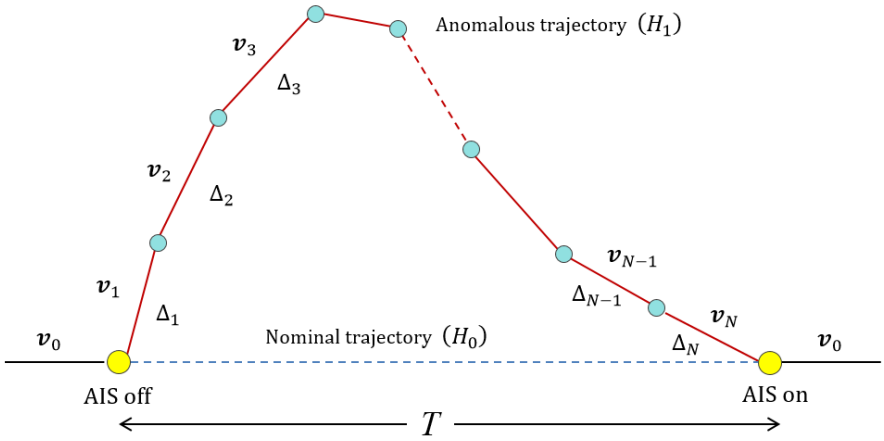
2. Hypothesis H_1: the vessel moves away from the nominal condition once the AIS transponder has been shut down. We assume that, during the time it *went dark*, the vessel is assumed to move away from the nominal condition by following a piecewise OU model with N unknown long-run mean velocities, comprised in the $2N$-dimensional vector $\boldsymbol{v}_{1:N} = \text{col}\{\boldsymbol{v}_n\}_{n=1}^{N}$. It identifies an N-section path, where, in general, $N \neq N_0$. In this chapter we assume $N = N_0$.

At the end of the time interval T the AIS device is switched back ON and the vessel keeps on moving under the nominal condition, as shown in Fig. 4.1.

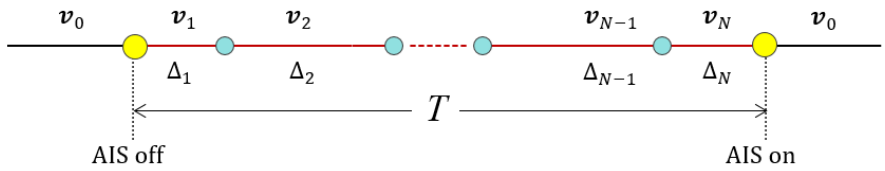
The sequence of the unknown long-run mean velocities $\boldsymbol{v}_{1:N}$, corresponding to the N-section path, is shown in Fig. 4.2. In particular, Fig. 4.2(a) represents the case where the vessel varies both trajectory and velocity, whereas in Fig. 4.2(b) only the velocity is altered. In both cases, considering the set of time instants $t_1 < ... < t_n < ... < t_N$, the period taken to cover the n-th section corresponds to the difference $\Delta_n \triangleq t_n - t_{n-1}$ and the sequence of these time intervals is denoted with $\mathcal{D}_N \triangleq \{\Delta_n\}_{n=1}^{N}$. The time period T when the AIS is disabled can be expressed as the sum of all the different time intervals Δ_n, so that $T = \sum_{n=1}^{N} \Delta_n$.

The considered problem amounts to determining, in the absence of AIS data and without any other information during the time interval T, whether the vessel has been following the planned trajectory at the nominal velocity $\boldsymbol{v}_0$ or not, by means of a *composite hypothesis test formulation* designed to identify changes in the velocity parameter.

In the first instance, the problem is studied just relying only upon two contacts available at the instants $t_0 = 0$ and $t_N = T$, respectively. Then, the problem is extended to the case where multiple contacts are available, each one located in a generic point along the N-section path.



(a) The vessel changes both trajectory and velocity during the AIS shutdown period.



(b) The vessel only alters its velocity during the AIS shutdown period.

Figure 4.2: Sequence of long-run mean velocities and time intervals characterizing the N-section path under hypothesis H_1.

4.3 Statistical representation of data based on the OU Process

As presented in Chapter 3, the OU model [56,58,154,212], has been validated against a real-world commercial maritime traffic dataset [154] and has been shown to enable a more accurate representation of the target state in the long-term when ships are not maneuvering.

It is worth mentioning that the NCV model could also be adopted in anomaly detection [187], but it does not turn out to be suitable for the specific problem at hand, since it lacks a parameter that represents a target's desired (cruise) velocity (see Section 3.2). This effectively makes the NCV incompatible with our detection strategy, that is instead based on the

changes of the long-run mean velocity. In such a situation, the OU appears clearly better suited, even when the anomalous maneuvering trajectory is considered, since this can be seen as a sequence of non-maneuvering paths, each one modeled by an OU process, and therefore as a piecewise OU model.

The SDE for the target motion model (3.7) is only suitable to represent a non-maneuvering target, i.e., whose long-run mean velocity does not change in time. However the model can be easily extended to the case of waypoint navigation [61], that is relevant to our application, being the navigation mode of substantially all the commercial maritime traffic. Along the navigational legs, the long-run mean velocity of the target obeys, by all means, the OU process; conversely, the maneuver is represented by a change of the long-run mean velocity parameter.

Under these assumptions, using hereafter sub-scripted indexes to denote time dependency, i.e., $\boldsymbol{s}_n = \boldsymbol{s}(t_n)$ and $\boldsymbol{\omega}_n = \boldsymbol{\omega}(t_n)$, where $\boldsymbol{s}(t)$ and $\boldsymbol{\omega}(t)$ are defined in (3.7), the target state at time t_n, given the target state at the previous time t_{n-1}, can be written as

$$\boldsymbol{s}_n = \boldsymbol{\Phi}(\Delta_n, \boldsymbol{\gamma})\boldsymbol{s}_{n-1} + \boldsymbol{\Psi}(\Delta_n, \boldsymbol{\gamma})\boldsymbol{v}_n + \boldsymbol{\omega}_n, \tag{4.1}$$

where $\boldsymbol{v}_n$ is the long-run mean velocity in the time interval $[t_{n-1}, t_n]$ and $\boldsymbol{\omega}_n$ is a zero-mean Gaussian random vector with covariance $\mathbf{C}(\Delta_n)$, whose complete expression is reported in (B.12) in Appendix B.1. The state transition matrix $\boldsymbol{\Phi}(\Delta_n, \boldsymbol{\gamma})$, and the control input function $\boldsymbol{\Psi}(\Delta_n, \boldsymbol{\gamma})$ have been introduced in Chapter 3 and are respectively defined in (B.6) and (B.8).

Omitting hereafter $\boldsymbol{\gamma}$ for clarity, the target state at time t_N can be expressed recursively as

$$\begin{aligned}
\boldsymbol{s}_N &= \boldsymbol{\Phi}(\Delta_N)\boldsymbol{s}_{N-1} + \boldsymbol{\Psi}(\Delta_N)\boldsymbol{v}_N + \boldsymbol{\omega}_N \\
&= \boldsymbol{\Phi}(\Delta_N)\left[\boldsymbol{\Phi}(\Delta_{N-1})\boldsymbol{s}_{N-2} + \boldsymbol{\Psi}(\Delta_{N-1})\boldsymbol{v}_{N-1} + \boldsymbol{\omega}_{N-1}\right] \\
&\quad + \boldsymbol{\Psi}(\Delta_N)\boldsymbol{v}_N + \boldsymbol{\omega}_N \\
&= \boldsymbol{\Phi}(T)\boldsymbol{s}_0 + \boldsymbol{\Psi}(\Delta_N)\boldsymbol{v}_N + \boldsymbol{\omega}_N + \sum_{n=1}^{N-1}\left[\prod_{i=n+1}^{N}\boldsymbol{\Phi}(\Delta_i)\right]\left[\boldsymbol{\Psi}(\Delta_n)\boldsymbol{v}_n + \boldsymbol{\omega}_n\right],
\end{aligned} \tag{4.2}$$

where we exploited the property of the state transition matrix by which $\boldsymbol{\Phi}(\Delta_1)\boldsymbol{\Phi}(\Delta_2)...\boldsymbol{\Phi}(\Delta_n) = \boldsymbol{\Phi}(\Delta_1 + \Delta_2 + ... + \Delta_n)$, $\forall n = 1, ..., N$, that can be derived by inspection from (B.6).

4.3.1 Two contacts available

Let us start by considering the case in which two measurements are available from the vessel. Specifically, we denote with $\boldsymbol{m}_0$ and $\boldsymbol{m}$ the two measurements respectively available at time $t_0 = 0$ and time t_N

$$\boldsymbol{m}_0 = \boldsymbol{s}_0 + \boldsymbol{n}_0, \quad \text{and} \quad \boldsymbol{m} = \boldsymbol{s}_N + \boldsymbol{n}, \tag{4.3}$$

where $\boldsymbol{n}_0$ and $\boldsymbol{n}$ are independent zero-mean Gaussian noises with covariance matrices $\mathbf{C}_{\boldsymbol{n}_0}$ and $\mathbf{C}_{\boldsymbol{n}}$, respectively. Clearly, the measurement noise is independent of the OU process noise.

Even if the distribution of $\boldsymbol{m}_0$ does not affect the hypothesis test, meaning that $\boldsymbol{s}_0$ has the same distribution under both hypotheses, such information is important because $\boldsymbol{s}_0$ represents the starting point of the kinematic terminating in $\boldsymbol{s}_N$. Given the linearity of equations (4.3) and exploiting (4.2), we can use the following data vector to avoid the dependence on $\boldsymbol{s}_0$ in $\boldsymbol{m}$

$$\begin{aligned}\boldsymbol{y} &\triangleq \boldsymbol{m} - \boldsymbol{\Phi}(T)\boldsymbol{m}_0 \\ &= \boldsymbol{\Psi}(\Delta_N)\boldsymbol{v}_N + \boldsymbol{\omega}_N + \sum_{n=1}^{N-1}\left[\prod_{i=n+1}^{N}\boldsymbol{\Phi}(\Delta_i)\right][\boldsymbol{\Psi}(\Delta_n)\boldsymbol{v}_n + \boldsymbol{\omega}_n] + \boldsymbol{n} - \boldsymbol{\Phi}(T)\boldsymbol{n}_0.\end{aligned} \tag{4.4}$$

In Appendix C.1 we show that the process noise is distributed as a zero-mean Gaussian with covariance matrix $\mathbf{C}(T)$, as

$$\boldsymbol{\omega}(T) \triangleq \sum_{n=1}^{N-1}\left[\prod_{i=n+1}^{N}\boldsymbol{\Phi}(\Delta_i)\right]\boldsymbol{\omega}_n + \boldsymbol{\omega}_N \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}(T)\right), \tag{4.5}$$

so that the data vector $\boldsymbol{y}$ can be recast as follows

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{\omega}(T) + \boldsymbol{n} - \boldsymbol{\Phi}(T)\boldsymbol{n}_0. \tag{4.6}$$

Consequently, the anomaly detection problem previously introduced in Section 4.2 can be described by the following Gaussian composite hypothesis testing problem:

$$\begin{cases} H_0: \ \boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{\mu}_0, \mathbf{C}_{\boldsymbol{y}}\right) \\ H_1: \ \boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{\mu}, \mathbf{C}_{\boldsymbol{y}}\right) \end{cases} \tag{4.7}$$

with covariance matrix given by

$$\mathbf{C}_{\boldsymbol{y}} \triangleq \mathbf{C}(T) + \mathbf{C}_{\boldsymbol{n}} + \boldsymbol{\Phi}(T)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T)^{\mathrm{T}}. \tag{4.8}$$

The mean vectors are distinguished under the null and the alternative hypotheses:

Under hypothesis H_1

The mean term under hypothesis H_1 is derived from (4.4) as

$$\begin{aligned}\boldsymbol{\mu} &= \mathbb{E}\left[\boldsymbol{y}|H_1\right] \\ &= \sum_{n=1}^{N-1}\left[\prod_{i=n+1}^{N}\boldsymbol{\Phi}(\Delta_i)\right]\boldsymbol{\Psi}(\Delta_n)\boldsymbol{v}_n + \boldsymbol{\Psi}(\Delta_N)\boldsymbol{v}_N \\ &= \mathbf{H}\,\boldsymbol{v}_{1:N},\end{aligned} \tag{4.9}$$

where, in this case, $\mathbf{H}$ is a $4 \times 2N$ matrix incorporating the state transition matrices (B.6) and the control input functions (B.8), dependent on the number of sections N and the time interval sequence $\mathcal{D}_N$:

$$\mathbf{H} \triangleq \underbrace{\left[\boldsymbol{\Phi}\left(\sum_{i=2}^{N}\Delta_i\right)\boldsymbol{\Psi}(\Delta_1), \ldots, \boldsymbol{\Phi}(\Delta_N)\boldsymbol{\Psi}(\Delta_{N-1}), \boldsymbol{\Psi}(\Delta_N)\right]}_{4\times 2N}. \tag{4.10}$$

Furthermore, we will make a distinction between the case where the number of sections N and the sequence of the time intervals $\mathcal{D}_N$ are known quantities, and the case where N and $\mathcal{D}_N$ are instead unknown. Then the mean term $\boldsymbol{\mu}$ (4.9) can be expressed as

$$\boldsymbol{\mu} = \mathbb{E}\left[\boldsymbol{y}|H_1\right] = \begin{cases} \mathbf{H}\,\boldsymbol{\theta} & \text{if } N, \mathcal{D}_N \text{ known} \\ \boldsymbol{\theta} & \text{if } N, \mathcal{D}_N \text{ unknown} \end{cases} \tag{4.11}$$

Under the condition of known parameters, $\mathbf{H}$ is a known matrix and $\boldsymbol{\theta} = \boldsymbol{v}_{1:N}$ is the $2N$-dimensional unknown parameter vector. Under the alternative condition, $\boldsymbol{\theta} = \boldsymbol{\mu}$ represents a global unknown parameter.

Under hypothesis H_0

Let us define the $2N$-dimensional vector $\boldsymbol{\theta}_0 \triangleq \boldsymbol{v}_0$. Then, under the null hypothesis H_0, the mean term is given by

$$\boldsymbol{\mu}_0 = \mathbb{E}\left[\boldsymbol{y}|H_0\right] = \begin{cases} \mathbf{H}\,\boldsymbol{\theta}_0 & \text{if } N, \mathcal{D}_N \text{ known} \\ \boldsymbol{\theta}_0 & \text{if } N, \mathcal{D}_N \text{ unknown.} \end{cases} \tag{4.12}$$

Note that, given $\boldsymbol{v}_0$, $\boldsymbol{\mu}_0$ is always known, even if N and $\mathcal{D}_N$ are unknown.

4.3.2 Multiple contacts available

Let us now assume that, in addition to the contacts in t_0 and T, a number of sporadic measurements (AIS, radar, SAR, etc.) is available at any time during the period between t_0 and $t_0 + T$. We hence assume that we have a stack of $K+1$ measurements, as $\operatorname{col}\{\boldsymbol{m}_k\}_{k=0}^{K}$, where the k-th measurement is given by

$$\boldsymbol{m}_k = \boldsymbol{s}(T_k) + \boldsymbol{n}_k, \tag{4.13}$$

and the measurement noise terms $\boldsymbol{n}_k$ are assumed to be independent and identically distributed according to a zero-mean Gaussian with covariance $\mathbf{C}_{\boldsymbol{n}_k}$. The k-th measurement is available at time $T_k = \frac{p_k}{N}T$, where $p_k \in [0, N]$ is by definition a fraction of the interval $[0, N]$ representing the time location of the contact with respect to the N piecewise OU velocities, as shown in Fig. 4.3, where p_k is located at some point along the n-th section of the path.

Accordingly, it is possible to consider the $4K$-dimensional data vector $\boldsymbol{y} = \operatorname{col}\{\boldsymbol{y}_k\}_{k=1}^{K}$, where the k-th component is defined as

$$\begin{aligned}\boldsymbol{y}_k &= \boldsymbol{m}_k - \boldsymbol{\Phi}(T_k)\boldsymbol{m}_0 \\ &= \boldsymbol{\mu}_k + \boldsymbol{\omega}(T_k) + \boldsymbol{n}_k - \boldsymbol{\Phi}(T_k)\boldsymbol{n}_0 \sim \mathcal{N}\left(\boldsymbol{\mu}_k, \mathbf{C}_{\boldsymbol{y}_k}\right).\end{aligned} \tag{4.14}$$

Similarly to (4.5) we can compute the OU process noise $\boldsymbol{\omega}(T_k) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(T_k))$ at time T_k, which is given by

$$\begin{aligned}\boldsymbol{\omega}(T_k) = &\boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\delta_k) \sum_{n=1}^{\lfloor p_k \rfloor - 1} \left[\prod_{i=n+1}^{\lfloor p_k \rfloor} \boldsymbol{\Phi}(\Delta_i) \right] \boldsymbol{\omega}(\Delta_n) \\ &+ \boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\delta_k)\boldsymbol{\omega}(\Delta_{\lfloor p_k \rfloor}) + \boldsymbol{\omega}(\Delta_{\lceil p_k \rceil}\delta_k),\end{aligned} \tag{4.15}$$

where $\delta_k = p_k - \lfloor p_k \rfloor$, with $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denoting respectively the ceiling and the floor functions. Consequently, the k-th component $\boldsymbol{y}_k$ is Gaussian with covariance matrix given by

$$\mathbf{C}_{\boldsymbol{y}_k} \triangleq \mathbf{C}(T_k) + \mathbf{C}_{\boldsymbol{n}_k} + \boldsymbol{\Phi}(T_k)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_k)^{\mathrm{T}}, \tag{4.16}$$

assuming the independence of the noise terms in (4.14).

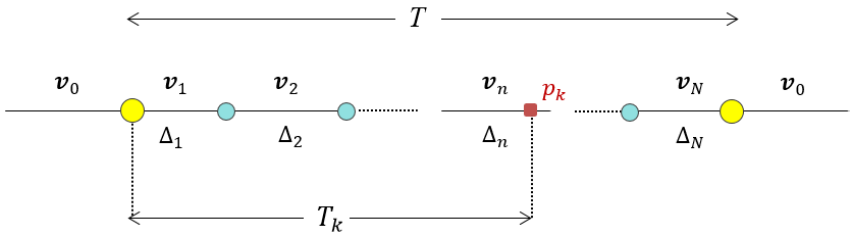


Figure 4.3: The parameter p_k identifies the time location of the k-th contact along the N-section path covered by the vessel.

The mean vector $\boldsymbol{\mu}_k$ is defined as

$$\begin{aligned}\boldsymbol{\mu}_k &\triangleq \boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\delta_k)\left\{\sum_{n=1}^{\lfloor p_k \rfloor -1}\left[\prod_{i=n+1}^{\lfloor p_k \rfloor}\boldsymbol{\Phi}(\Delta_i)\right]\boldsymbol{\Psi}(\Delta_n)\boldsymbol{v}_n + \boldsymbol{\Psi}(\Delta_{\lfloor p_k \rfloor})\boldsymbol{v}_{\lfloor p_k \rfloor}\right\} \\ &\quad + \boldsymbol{\Psi}(\Delta_{\lceil p_k \rceil}\delta_k)\boldsymbol{v}_{\lceil p_k \rceil}, \\ &= \overline{\mathbf{H}}_k\, \boldsymbol{v}_{1:N} \end{aligned} \tag{4.17}$$

which is derived in the same way as the mean term (4.9) in the case of only two contacts. $\overline{\mathbf{H}}_k$ is the $4 \times 2N$ model matrix which includes the state transition matrices (B.6) and the control input functions (B.8) related to the k-th radar contact, whose expression is

$$\overline{\mathbf{H}}_k = \left[\underbrace{\boldsymbol{\Pi}_k}_{4\times 2\lceil p_k \rceil}, \underbrace{\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}}_{4\times 2(N-\lceil p_k \rceil)}\right], \tag{4.18}$$

with

$$\begin{aligned}&\boldsymbol{\Pi}_k = \\ &[\boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\delta_k)\boldsymbol{\Phi}\left(\sum_{i=2}^{\lfloor p_k \rfloor}\Delta_i\right)\boldsymbol{\Psi}(\Delta_1), \dots \\ &\dots, \boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\delta_k)\boldsymbol{\Phi}\left(\sum_{i=\lceil p_k \rceil -1}^{\lfloor p_k \rfloor}\Delta_i\right)\boldsymbol{\Psi}(\Delta_{\lceil p_k \rceil -2}), \boldsymbol{\Phi}(\Delta_{\lceil p_k \rceil}\beta_k)\boldsymbol{\Psi}(\Delta_{\lceil p_k \rceil -1}), \\ &\boldsymbol{\Psi}(\Delta_{\lceil p_k \rceil}\beta_k)], \end{aligned} \tag{4.19}$$

and $\beta_k = 1$ if $\delta_k = 0$, or $\beta_k = \delta_k$ otherwise. Specifically, the null matrix appearing in (4.18) cancels the long-run mean velocities contributing later than the time T_k of the k-th contact. Notice that, for $K = 1$ and $p_k = N$, $\overline{\mathbf{H}}_k$ is just the matrix (4.10) found in the previous case where only two contacts are available.

The full data vector $\boldsymbol{y}$ is therefore characterized as in (4.7) with covariance matrix given by

$$\mathbf{C}_{\boldsymbol{y}} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1K} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{K1} & \mathbf{C}_{K2} & \cdots & \mathbf{C}_{KK} \end{bmatrix}, \tag{4.20}$$

where, $\forall i, j = 1, \ldots, K$, (see Appendix C.2)

$$\mathbf{C}_{ij} = \begin{cases} \mathbf{C}(T_i) + \mathbf{C}_{\boldsymbol{n}_i} + \boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_i)^{\mathrm{T}} & \text{if } i = j \\ \mathbf{C}(T_i)\boldsymbol{\Phi}(T_j - T_i)^{\mathrm{T}} + \boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_j)^{\mathrm{T}} & \text{if } i < j, \end{cases} \tag{4.21}$$

and $\mathbf{C}_{ji} = \mathbf{C}_{ij}^{\mathrm{T}}$.

As for the two-contacts case, the mean vectors of the full data vector $\boldsymbol{y}$ are distinguished under the null and the alternative hypotheses. Indeed, $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_0$ are still expressible as (4.11) and (4.12), respectively, with the only difference that, in this case, $\mathbf{H}$ is the overall model matrix of size $4K \times 2N$ made up of the model matrices $\overline{\mathbf{H}}_k$ in (4.18) as

$$\mathbf{H} = \begin{bmatrix} \overline{\mathbf{H}}_1 & \ldots & \overline{\mathbf{H}}_k & \ldots & \overline{\mathbf{H}}_K \end{bmatrix}^{\mathrm{T}}. \tag{4.22}$$

4.4 Detection of anomalous deviations from the expected route in the presence of different levels of data unavailability

Summarizing the results from the previous section, the anomaly detection problem described in Section 4.2 is addressed in a binary Gaussian hypothesis test framework as:

$$\begin{cases} H_0 : \ \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}_0, \mathbf{C}_{\boldsymbol{y}}), \\ H_1 : \ \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C}_{\boldsymbol{y}}), \end{cases} \tag{4.23}$$

with the same covariance matrix under the two hypotheses, given in (4.8) in the case of only two contacts available, and in (4.20) in the case of $K+1$ contacts. Instead, the mean term depends on the hypothesis as:

$$\boldsymbol{\mu}_0 = \begin{cases} \mathbf{H} \text{ known}: & \mathbf{H}\boldsymbol{\theta}_0 = \mathbf{H}\,\boldsymbol{v}_0, \\ \mathbf{H} \text{ unknown}: & \boldsymbol{\theta}_0 = \mathbf{H}\,\boldsymbol{v}_0, \end{cases} \tag{4.24}$$

$$\boldsymbol{\mu} = \begin{cases} \mathbf{H} \text{ known}: & \mathbf{H}\boldsymbol{\theta} = \mathbf{H}\boldsymbol{v}_{1:N}, \\ \mathbf{H} \text{ unknown}: & \boldsymbol{\theta} = \mathbf{H}\boldsymbol{v}_{1:N}, \end{cases} \tag{4.25}$$

with model matrix $\mathbf{H}$ given by (4.10) in the case of only two contacts available, and (4.22) in the case of multiple contacts.

The anomaly detection strategy developed throughout this chapter for the hypothesis testing problem at hand, is based on the GLRT approach that can be easily traced back to the GLRT for Gaussian linear model [179]. By denoting with $L_{\boldsymbol{\theta}}(\boldsymbol{y})$ the test function, the GLRT can be expressed as

$$L_{\boldsymbol{\theta}}(\boldsymbol{y}) = \arg\max_{\boldsymbol{\theta}} \{\ln[p_{\boldsymbol{\theta}}(\boldsymbol{y})]\} - \ln[p_{\boldsymbol{\theta}_0}(\boldsymbol{y})] \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{4.26}$$

where $p_{\boldsymbol{\theta}_0}$ is the probability distribution under the simple null hypothesis H_0 by which there have been no changes in the vessel velocity during the AIS transponder shutdown; conversely, $p_{\boldsymbol{\theta}}$ is the probability distribution, depending on the unknown parameter $\boldsymbol{\theta}$, under the composite alternative hypothesis H_1 by which the vessel has changed its velocity during the AIS transponder shutdown. With a minor abuse of notation, the threshold will be always identified by τ from now on.

The test performance is defined in terms of false alarm probability P_{FA}, i.e., the probability that the test statistic exceeds the threshold under H_0 and detection probability P_{D}, i.e., the probability that the test statistic exceeds the threshold under H_1. By letting the threshold τ vary, the P_{FA} and P_{D} values define a curve in the $(P_{\text{FA}}, P_{\text{D}})$ plane named the Receiver Operating Characteristic (ROC). Since the hypothesis testing problem at hand is addressed via the GLRT for Gaussian linear model, the test statistics under the two hypotheses H_0 and H_1 are characterized, respectively, by a central and a non-central Chi-squared distributions [119], both with d degrees of freedom. In such a way, the detection performance is described by

$$P_{\text{FA}} = Q_{\chi^2_d}(\tau), \quad \text{and} \quad P_{\text{D}} = Q_{\chi'^2_d(\lambda)}(\tau), \tag{4.27}$$

where $Q_{\chi^2_d}$ and $Q_{\chi'^2_d(\lambda)}$ are the right tail probabilities of the central and non-central Chi-squared distributions, respectively. The parameter $\lambda = \lambda(\boldsymbol{\theta})$, denoted as *non-centrality parameter* [119], is twice the KL-divergence between the actual distributions under the hypotheses H_0 and H_1. In other words, $\lambda(\boldsymbol{\theta})$ represents a measure of how much the hypotheses H_0 and H_1 are distant from each other, meaning that the two hypotheses become more disjoint as $\lambda(\boldsymbol{\theta})$ increases. Therefore, the anomaly detection performance of the clairvoyant receiver, obtained assuming known the ship velocity profile under the two hypotheses, is ruled by $\lambda(\boldsymbol{\theta})$, which is also identified as *anomaly distance* in such a context.

In the following we develop the detection strategy for the case where multiple contacts are available, making a distinction between the case where N and $\mathcal{D}_N$ are unknown and the one where such parameters are known. Finally the issue arising from this strategy will be highlighted and investigated.

4.4.1 Scenario 1: N and $\mathcal{D}_N$ unknown parameters

In this case, we have $\boldsymbol{\mu} = \boldsymbol{\theta} = \mathbf{H}\,\boldsymbol{v}_{1:N}$ under H_1 and $\boldsymbol{\mu}_0 = \boldsymbol{\theta}_0 = \mathbf{H}\,\boldsymbol{v}_0$ under H_0. The GLRT is derived from (4.26), as follows

$$L_{\boldsymbol{\theta}}(\boldsymbol{y}) = \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{4.28}$$

where $\widehat{\boldsymbol{\theta}}$ represents the Maximum Likelihood Estimate (MLE) for the parameter $\boldsymbol{\theta}$, and, in Appendix C.3, it is shown that $\widehat{\boldsymbol{\theta}} = \boldsymbol{y}$.[2]

As previously anticipated, the test statistics under the two hypotheses H_0 and H_1 are characterized, respectively, by central and non-central Chi-squared distributions, both with $d = 4K$ degrees of freedom, corresponding to the dimension of the known parameter $\boldsymbol{\theta}_0$. The detection performance is therefore given by (4.27) with the following non-centrality parameter

$$\begin{aligned} \lambda(\boldsymbol{\theta}) &= (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \\ &= (\boldsymbol{v}_{1:N} - \boldsymbol{v}_0)^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H} (\boldsymbol{v}_{1:N} - \boldsymbol{v}_0)\,. \end{aligned} \tag{4.29}$$

[2] In this case, it is preferable to directly estimate the global unknown parameter $\boldsymbol{\theta}$, instead of jointly estimating N, $\mathcal{D}_{\mathcal{N}}$ and $\boldsymbol{v}_{1:N}$, which would require a numerical solution ot the GLRT with no available closed-form performance expression.

4.4.2 Scenario 2: N and $\mathcal{D}_N$ known parameters

In this case $\boldsymbol{\theta} = \boldsymbol{v}_{1:N}$ under H_1 and $\boldsymbol{\theta}_0 = \boldsymbol{v}_0$ under H_0. Similarly to what seen in the previous case, the GLRT has the following form:

$$L_{\boldsymbol{\theta}}(\boldsymbol{y}) = \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H} \left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{4.30}$$

where $\widehat{\boldsymbol{\theta}}$ is the MLE of the parameter $\boldsymbol{\theta}$, given by (see Appendix C.3)

$$\widehat{\boldsymbol{\theta}} = \left(\mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \boldsymbol{y}, \tag{4.31}$$

when $\mathbf{H}$ is a full rank matrix. Therefore it is possible to achieve a direct estimate of the vector incorporating the velocities assumed by the vessel under the hypothesis H_1 instead of an estimate of the global parameter as shown in the previous case[3]. The detection performance is given by (4.27), but in the current case the number of degrees of freedom for both central and non-central Chi-squared distributions is $d = 2N$, corresponding to the dimension of the known parameter $\boldsymbol{\theta}_0$, and the non-centrality parameter is given by (4.29).

On the other hand, when $\mathbf{H}$ is ill-conditioned, a problem of matrix inversion arises in the MLE expression (4.31). This is identified as a *rank deficiency* problem that can be approached by using the Singular Value Decomposition (SVD) [96] of matrix $\mathbf{H}$, given by

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{Q}^{\mathrm{T}}, \tag{4.32}$$

where $\mathbf{U}$ is a $4K \times 4K$ unitary matrix, $\mathbf{S}$ is a $4K \times 2N$ rectangular diagonal matrix with non-negative real numbers on the diagonal, corresponding to the singular values of $\mathbf{H}$, and $\mathbf{Q}$ is a $2N \times 2N$ unitary matrix.

The central idea is to replace $\mathbf{H}$ by its rank reduced version, and this is referred to as truncating the SVD [209]. It is shown that $\rho \triangleq \mathrm{rank}(\mathbf{H}) = \mathrm{rank}(\mathbf{S})$ and it is equal to the number of non-zero singular values of $\mathbf{H}$. Therefore ρ provides the effective size of the matrices involved, meaning that $\mathbf{U}\mathbf{S}\mathbf{Q}^{\mathrm{T}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{S}}\widetilde{\mathbf{Q}}^{\mathrm{T}}$, where $\widetilde{\mathbf{U}}$, of size $4K \times \rho$, and $\widetilde{\mathbf{Q}}^{\mathrm{T}}$, of size $\rho \times 2N$, are respectively the matrices of the left-singular vectors and of the right-singular vectors of $\mathbf{H}$, while $\widetilde{\mathbf{S}}$ is a $\rho \times \rho$ diagonal matrix. In such a way, by using the reduced rank version of $\mathbf{H}$, it is possible to get

$$\mathbf{H}\boldsymbol{\theta} = \widetilde{\mathbf{U}}\widetilde{\mathbf{S}}\widetilde{\mathbf{Q}}^{\mathrm{T}}\boldsymbol{\theta} = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}, \tag{4.33}$$

[3] When $\mathbf{H}$ holds a rank equal to the number of its rows, we fall back into the case where $\mathbf{H}$ is unknown, so that $\mathbf{H}\widehat{\boldsymbol{\theta}} = \boldsymbol{y}$, as shown in [209].

where $\widetilde{\boldsymbol{\theta}} \triangleq \widetilde{\mathbf{S}}\widetilde{\mathbf{Q}}^{\mathrm{T}}\boldsymbol{\theta}$ is a ρ-dimensional unknown vector resulting from a process of rotation and scaling of the vector $\boldsymbol{\theta}$. In the same way, under hypothesis H_0, we obtain $\mathbf{H}\boldsymbol{\theta}_0 = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}_0$.

By applying the SVD (4.33) to the GLRT (4.30) and considering the following *Cholesky Decomposition* [96]

$$\mathbf{C}_{\boldsymbol{y}}^{-1} = \widetilde{\mathbf{C}}^{\mathrm{T}}\widetilde{\mathbf{C}}, \tag{4.34}$$

with $\widetilde{\mathbf{C}}$ lower triangular matrix with positive elements on the diagonal, the reformulation of the GLRT is achieved as

$$\begin{aligned} L_{\widetilde{\boldsymbol{\theta}}}(\boldsymbol{y}) &= \left(\widehat{\widetilde{\boldsymbol{\theta}}} - \widetilde{\boldsymbol{\theta}}_0\right)^{\mathrm{T}} \widetilde{\mathbf{U}}^{\mathrm{T}}\widetilde{\mathbf{C}}^{\mathrm{T}}\widetilde{\mathbf{C}}\widetilde{\mathbf{U}} \left(\widehat{\widetilde{\boldsymbol{\theta}}} - \widetilde{\boldsymbol{\theta}}_0\right) \\ &= \left(\widehat{\widetilde{\boldsymbol{\theta}}} - \widetilde{\boldsymbol{\theta}}_0\right)^{\mathrm{T}} \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}} \left(\widehat{\widetilde{\boldsymbol{\theta}}} - \widetilde{\boldsymbol{\theta}}_0\right) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \end{aligned} \tag{4.35}$$

with $\widetilde{\mathbf{H}} \triangleq \widetilde{\mathbf{C}}\widetilde{\mathbf{U}}$. Derived as done for (4.31), the MLE of parameter $\widetilde{\boldsymbol{\theta}}$ is given by

$$\widehat{\widetilde{\boldsymbol{\theta}}} = \left(\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\right)^{-1} \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{C}}\boldsymbol{y}, \tag{4.36}$$

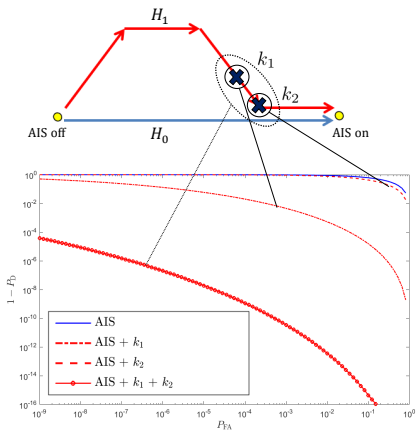
where the problem of matrix inversion does not arise since $\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}$ is a full-rank matrix of size $\rho \times \rho$ by construction, and, therefore, invertible. In this case, the number of degrees of freedom is $d = \rho$, corresponding to the dimension of the known parameter $\widetilde{\boldsymbol{\theta}}_0$, while the non-centrality parameter is immutably given by (4.29).

4.5 Experimental results

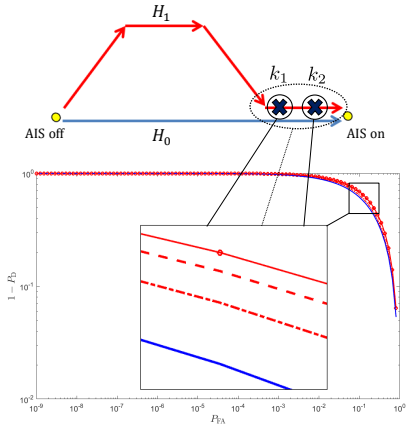
In this subsection we present some experimental results by applying the GLRTs (4.28) and (4.35) derived in Section 4.4 to both synthetic and real-world scenarios. In particular, the proposed anomaly detection strategy will be applied to the real-world scenarios introduced in Chapter 1 concerning the illicit rendezvous nearby the Galápagos Marine Reserve, and the *Ever Given* grounding in the Suez Canal.

4.5.1 Analysis of a synthetic scenario

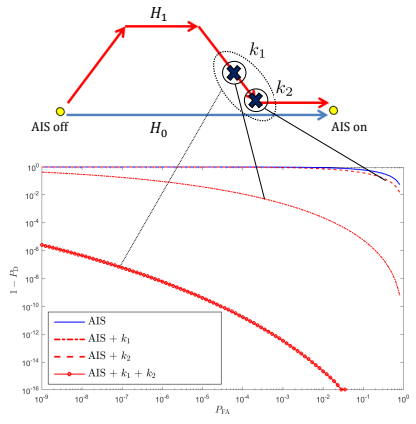
The performance of the tests (4.28) and (4.35) will be examined for a situation of concern where multiple contacts are considered. We set a config-



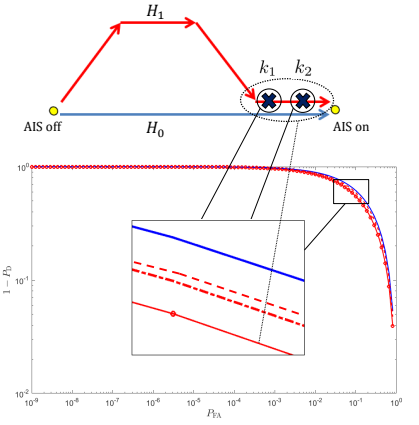
(a) Case study 1.(a)



(b) Case study 1.(b)



(c) Case study 2.(a)



(d) Case study 2.(b)

Figure 4.4: Panel (a) shows the GLRT (4.28) performance for case study (a). Panel (b) shows the GLRT (4.28) performance for case study (b). Panel (c) shows the GLRT (4.35) performance for case study (a). Panel (d) shows the GLRT (4.35) performance for case study (b).

uration for the analyses of a synthetic scenario, where a vessel is navigating under nominal conditions along a straight route with constant velocity $\boldsymbol{v}_0 = [8\ \ 0]^{\mathrm{T}}$ m/s and at some point it turns its AIS device off for a time period $T = 12$ hours. Under hypothesis H_1 the vessel is supposed to follow an N-section path and the time intervals are assumed all equal, so that $\Delta_n = T/N,\ \forall n = 1, ..., N$. The reversion rate of the underlying OU dynamic model is set as $\gamma_x = \gamma_y = 0.9 \cdot 10^{-2}$ and $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}} = \sigma^2\mathbf{I}$, with noise level $\sigma^2 = 10^{-2}$. The noise covariance matrix is set as $\mathbf{C}_{\boldsymbol{n}_k} = \mathrm{diag}(50^2, 50^2, 1, 1)$ $\forall k = 1, ..., K$, while $\mathbf{C}_{\boldsymbol{n}_0}$ is assumed to be negligible.

A comparison between the tests (4.28) and (4.35) derived in Section 4.4 is highlighted here, where we distinguish two specific case studies depending on the time location of multiple contacts. Specifically we assume to observe two AIS contacts (the last one before, and the first one just after the AIS device shutdown) and two radar contacts located along the path covered by the vessel, which, under hypothesis H_1, consists of $N = 4$ sections characterized by the following sequence of long-run mean velocities $|\boldsymbol{v}_1| = 7.9$, $|\boldsymbol{v}_2| = 8.6$, $|\boldsymbol{v}_3| = 7.8$, $|\boldsymbol{v}_4| = 7.8$. The two radar contacts are denoted with k_1 and k_2 and the corresponding time locations with p_{k_1} and p_{k_2}. This means that the contacts are observed at $T_{1,2} = \dfrac{p_{k_{1,2}}}{N}\, T$.

Performance curves are provided in terms of missed detection probability, $1 - P_{\mathrm{D}}$, versus false alarm probability, P_{FA}, (see Fig. 4.4). Specifically, for both case studies, while the solid line is related to the detection procedure performed by using the only AIS contribution, the dot-dashed and dashed lines represent the test performance achieved by using AIS data and the measurement provided by the radar contacts k_1 and k_2, respectively. Furthermore performance related to the combined use of both radar contacts and AIS is denoted with an o-marker line.

Case study (a)

The performance of the GLRTs (4.28) and (4.35) for the case study (a) is depicted in Fig. 4.4(a) and (c) respectively. The two radar contacts are located in $p_{k_1} = 2.7$ and in $p_{k_2} = 3$. We can easily verify that, in both cases, the improvement obtained by using contact k_2 is not significant since it is located where the vessel is very close to the expected position as if the deviation never happened. Whereas, the use of k_1, located where the anomalous behavior and the nominal condition are significantly distant, provides a remarkable improvement. Finally the combined use of both radar contacts

Table 4.1: Non-centrality parameter and degrees of freedom for radar contacts located in $p_{k_1} = 2.7$, $p_{k_2} = 3$.

		AIS	AIS + k_1	AIS + k_2	AIS + k_1 + k_2
$\lambda(\boldsymbol{\theta})$		3.3072	50.9995	7.4007	134.4314
d	1.(a)	4	8	8	12
	2.(a)	4	6	6	6

Table 4.2: Non-centrality parameter and degrees of freedom for radar contacts located in $p_{k_1} = 3.3$, $p_{k_2} = 3.7$.

		AIS	AIS + k_1	AIS + k_2	AIS + k_1 + k_2
$\lambda(\boldsymbol{\theta})$		3.3072	4.1040	3.6086	4.1040
d	1.(b)	4	8	8	12
	2.(b)	4	4	4	4

improves the performance with respect to the use of a single radar contact. We observe that in both cases performance improves with increasing number of radar contacts, however as discussed in Subsection 4.4, given that $\lambda(\boldsymbol{\theta})$ is equal for both detectors, GLRT (4.35) exhibits better performance because d is smaller comparing to the case of GLRT (4.28) (see also the deepening in Appendix C.4). The values of $\lambda(\boldsymbol{\theta})$ and d for cases 1.(a) and 2.(a) are reported in Table 4.1.

Case study (b)

Performance of GLRT (4.28) and (4.35) for the case study (b) is depicted in Fig. 4.4(b) and (d) respectively, where the two radar contacts are now located at $p_{k_1} = 3.3$ and $p_{k_2} = 3.7$, that is, along a part of the path where the anomalous behavior is very close to the nominal condition, and therefore the two hypotheses are difficult to distinguish. In this case, compared to the case of AIS information only, the performance of the GLRT (4.28) gets worse with increasing number of radar contacts. This phenomenon is explained in Appendix C.4, and basically it is the case in which adding new contacts in location where H_1 is close to H_0 has the effect of adding mostly noise to the decision statistic. On the other hand, GLRT (4.35) shows increasingly

better performance, even though the improvement is quite small because of the radar contacts contain limited information for discriminating the two hypotheses. The values of $\lambda(\boldsymbol{\theta})$ and d for cases 1.(b) and 2.(b) are reported in Table 4.2.

4.5.2 Analysis of a real-world scenario: Illicit rendezvous nearby the Galápagos Marine Reserve

The detection strategy proposed in Chapter 4.4 has been applied to a real-world example of anomalous behavior provided in Fig. 1.5, where the about five-month AIS track of a cargo vessel is shown.[4] The vessel navigates with a nominal speed of about 5 m/s in the waters of the Pacific Ocean [17].

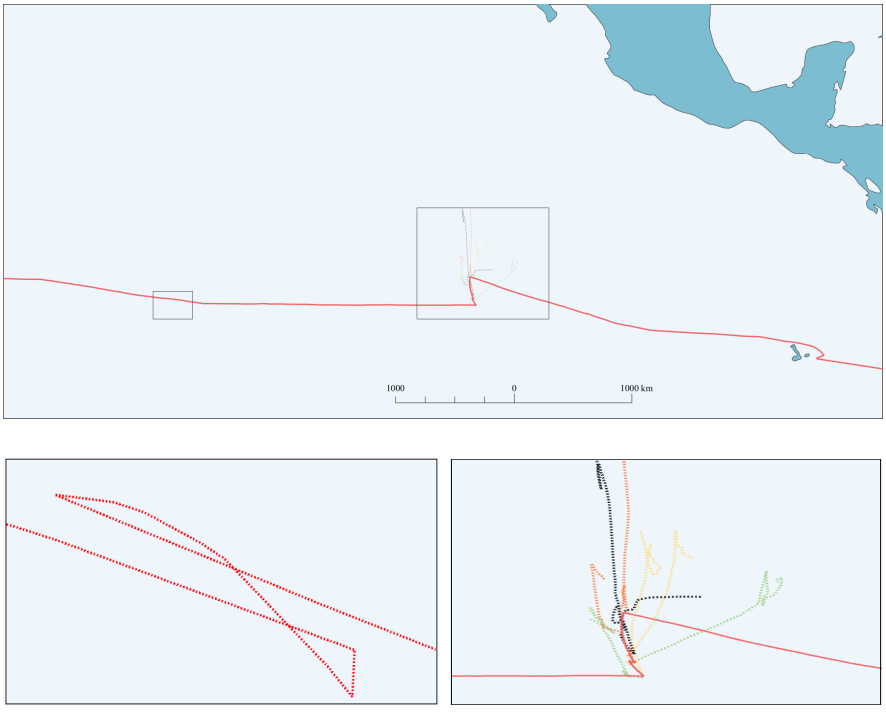
Nearby the Galápagos EEZ two area of interest (AoI) characterized by an anomalous behavior can be identified by simple visual inspection of Fig. 4.5. In both AoIs the vessel shuts the engines down and starts drifting, with an apparent deviation from its expected route. While there is no apparent reason for the drifting in the first AoI, the deviation in the second AoI is clearly performed to rendezvous with four tuna longliners at about 1700 miles away from Galápagos. Each fishing vessel spends about 12 hours moving along with the vessel at a distance of about 30 m, which indicates the boats were likely tied up. This behavior suggests a substantial transfer of cargo was possible [17].

For the application of the detection strategy, we estimated the OU parameters in the path section immediately preceding the AoI where the deviation from the nominal conditions actually happens. Specifically, the estimated values of the process noise $\boldsymbol{\sigma}$, the long-run mean velocity $\boldsymbol{v}_0$, and the reversion rate $\boldsymbol{\gamma}$ are reported in Table 4.3.

First AoI

The anomalous behavior in the first AoI is short and not very apparent. According to the AIS track, the vessel seems to shut down the engines and drift for about 14 hours, then it goes back to the planned route as it can be seen in the close-up of Fig. 4.6 (a). Such an anomaly is reflected on the velocity components shown in Fig. 4.6 (e), and it is more visible in Fig. 4.6 (c)

[4] This subsection has been published as "Maritime Anomaly Detection Based on Mean-Reverting Stochastic Processes Applied to a Real-World Scenario," in 21st International Conference on Information Fusion, July 2018 [65].

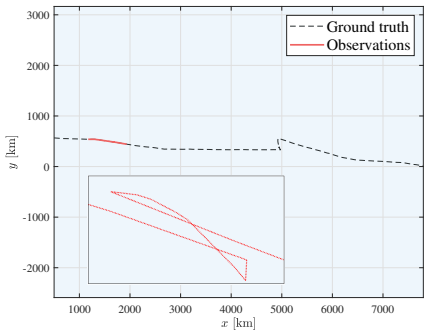


(a) First AoI (b) Second AoI

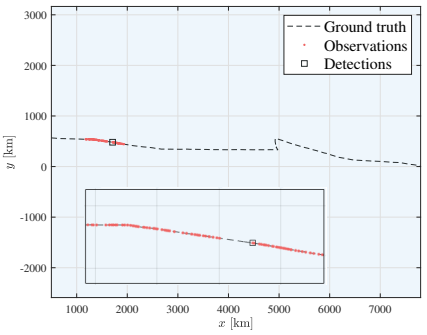
Figure 4.5: The track of the cargo vessel (red) exhibiting anomalous behaviors along the expected route. Panel (a) shows the first AoI, where the cargo vessel drifts with no apparent reason. Panel (b) shows the second AoI, where the cargo vessel is evidently suspected of several rendezvous with four fishing vessels (dotted tracks).

where, during the deviation time, the decision statistic (4.28) grows by several orders of magnitude with respect to nominal readings, exceeding all the values of the threshold (i.e., for different values of the false alarm probability: $P_{\mathrm{FA}} \in \left\{10^{-4}, 10^{-6}, 10^{-8}\right\}$).

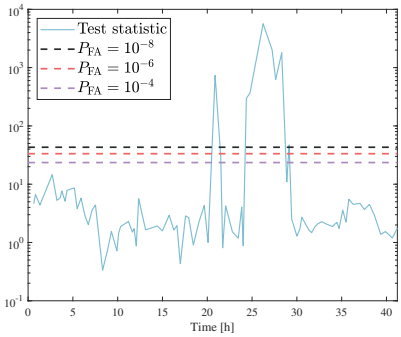
In order to test the proposed detection strategy, we simulated a data gap in the AIS track corresponding to the anomaly time frame, as depicted in Fig. 4.6 (b) with the corresponding velocity gaps shown in Fig. 4.6 (f). The detector (4.28) correctly reveals the deviation from the nominal condition using a threshold selected with $P_{\mathrm{FA}} = 10^{-6}$, as shown in Fig. 4.6 (d).



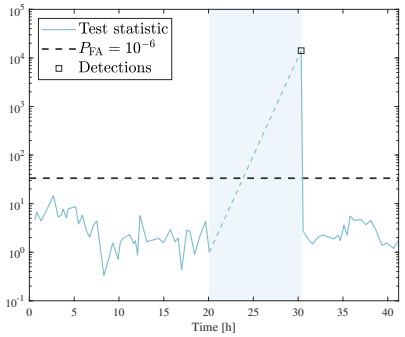
(a) Complete AIS track (1st AoI).



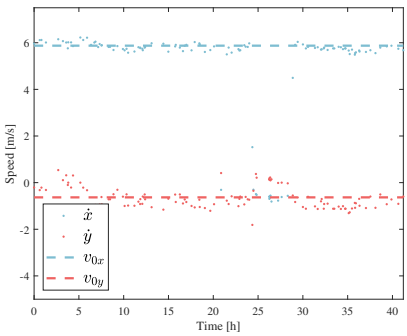
(b) Simulated gaps (1st AoI).



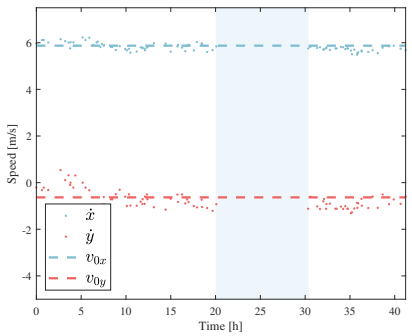
(c) Test statistic (4.28) related to the complete AIS track (1st AoI).



(d) Test statistic (4.28) related to the incomplete AIS track (1st AoI).

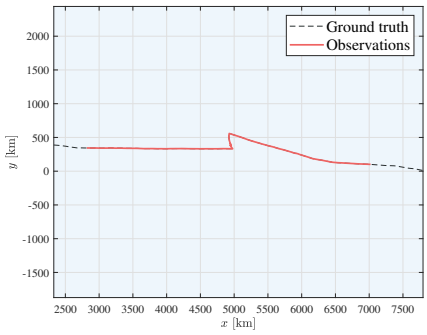


(e) Velocity components related to the complete AIS track (1st AoI).

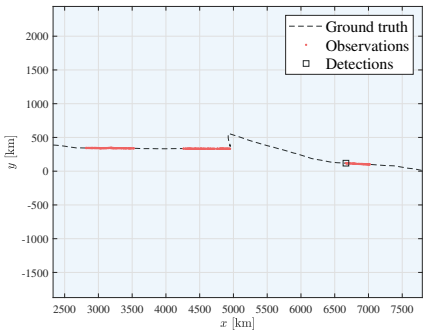


(f) Velocity components related to the incomplete AIS track (1st AoI).

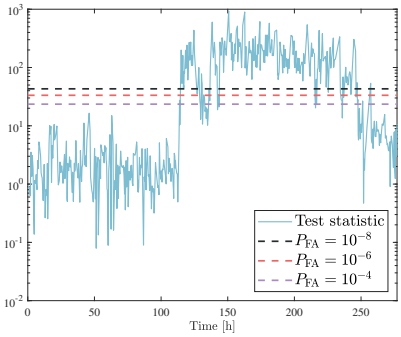
Figure 4.6



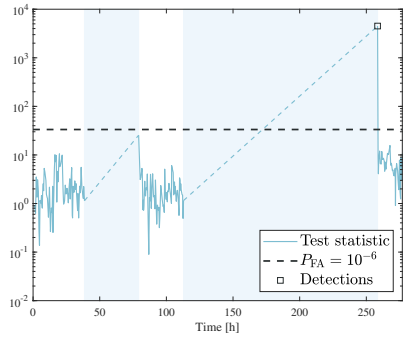
(a) Complete AIS track (2nd AoI).



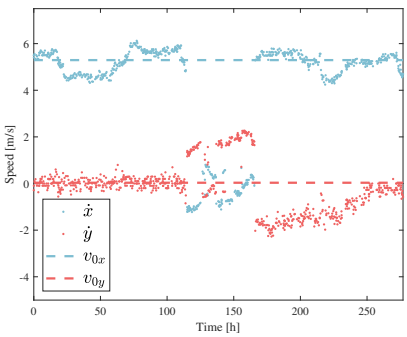
(b) Simulated gaps (2nd AoI).



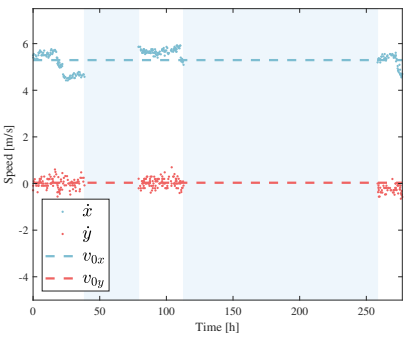
(c) Test statistic (4.28) related to the complete AIS track (2nd AoI).



(d) Test statistic (4.28) related to the incomplete AIS track (2nd AoI).



(e) Velocity components related to the complete AIS track (2nd AoI).



(f) Velocity components related to the incomplete AIS track (2nd AoI).

Figure 4.7

Table 4.3: OU parameters estimates

		First AoI	Second AoI
$\boldsymbol{\gamma}$	γ_x	5.89×10^{-3}	2.30×10^{-4}
	γ_y	8.49×10^{-4}	4.19×10^{-3}
$\boldsymbol{\sigma}$	σ_x	2.83×10^{-2}	1.13×10^{-2}
	σ_y	1.84×10^{-2}	2.23×10^{-2}
$\boldsymbol{v}_0$	v_{0_x}	5.8743	5.2931
	v_{0_y}	-0.6320	0.0331

Second AoI

The AIS track of the ship in the second observation window is shown in Fig. 4.7 (a) and it does indeed reveal a deviation from the normal route during a time frame of about 5 days; Fig. 4.7 (c) displays the test statistic (4.28), which exceeds the threshold (plotted for different values of the false alarm probability in the same range considered for the first region analysis) corresponding to the deviation from the nominal condition. In particular, the deviation from the nominal velocity is visible in Fig. 4.7 (e) where the velocity components show an apparent change in that specific time frame.

As done for the first AoI, the detection strategy is tested with simulated gaps in AIS data, as shown in Fig. 4.7 (b), with the corresponding velocity gaps shown in Fig. 4.7 (f). The first data gap occurs in a section of the trajectory where there is no deviation from the nominal conditions, while the second one occurs where the deviation actually happens. From the application of the detector (4.28) with $P_{\text{FA}} = 10^{-6}$, the deviation can be properly detected while no detection is correctly declared in the first gap, as shown in Fig. 4.7 (d).

4.5.3 Analysis of a real-world scenario: *Ever Given* grounding in the Suez Canal

In this subsection, we show how the anomaly detection tools developed in Section 4.4 could have been successfully applied to the real-world maritime scenario of the *Ever Given* using the available sequence of AIS records, and contextual information defining the expected nominal behavior of navigation.[5] In the specific, the GLRT (4.30) is sequentially run between two consecutive AIS records of the *Ever Given* to decide whether or not a deviation from the nominal behavior happened within that specific time period. Our results reveal an anomalous behavior of the ship, detected 19 minutes before the grounding, which could have helped avoid the accident that caused such a negative impact on maritime traffic and global trade.

Configuration of model parameters and nominal behavior

The five available AIS data points let us consider four time windows of duration T_j, $j = 1, 2, 3, 4$, in each of which the vessel is assumed to move according to a piecewise OU model. Specifically, under H_0, the vessel is expected to move along a planned trajectory according to a piecewise OU model characterized by a sequence $\boldsymbol{v}_{0_j} = \text{col}\{\boldsymbol{v}_{n,j}^0\}_{n=1}^{N_j^0}$ of nominal velocities, identifying an N_j^0-section nominal path within the j-th time frame (see Fig. 4.8). Alternatively, under H_1, the vessel is assumed to move away from the nominal condition by following a sequence $\boldsymbol{v}_j = \text{col}\{\boldsymbol{v}_{n,j}\}_{n=1}^{N_j}$ of N_j unknown velocities.

Given the characteristics of the navigation area, the expected nominal behavior of the *Ever Given* inside the canal as shown in Fig. 4.8 was defined based on

- the trajectory path to be followed considering its geographical constraints;
- speed of transit limits put in place by the SCA [10].

In particular, we considered a constant nominal speed $v_0 = 4$ m/s, below the limit of 4.44 m/s (16 km/h) for ordinary vessels, and consistent with an average transit time of 14 hours to navigate a canal with a length of 193.3

[5] This subsection has been published as "Maritime Anomaly Detection in a Real-World Scenario: Ever Given Grounding in the Suez Canal," in IEEE Transactions on Intelligent Transportation Systems, 2021 [86].

Figure 4.8: Sequence of OU velocities and time intervals characterizing the nominal dynamic behavior of the *Ever Given* inside the Suez Canal given the available sequence of AIS contacts.

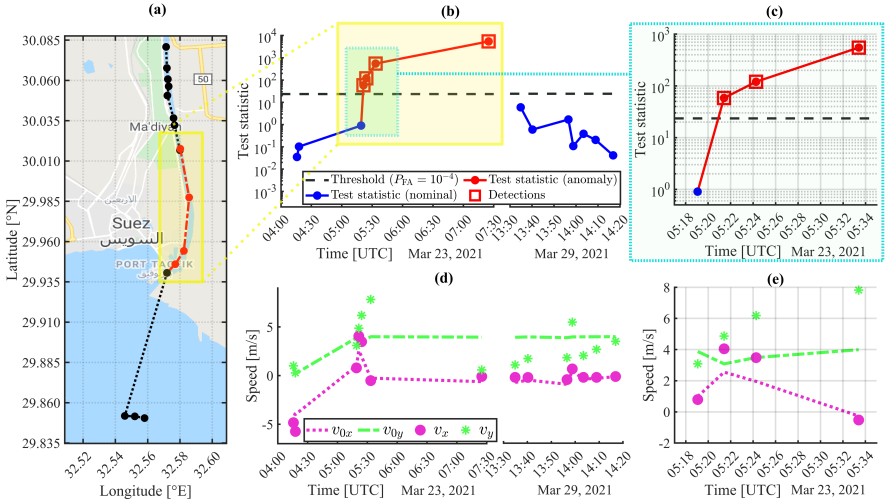


Figure 4.9: (a) *Ever Given* trajectory reported by AIS data collected from 23-03-21 to 29-03-21. Please note that the AIS trajectory is obtained by simply connecting the available contacts (circles) with a dotted line. (b) Test statistic of the anomaly detector (4.30) over the period pre- and post-grounding of the *Ever Given*. (c) Close-up of the evolution of the anomaly test statistic during the navigation of the ship inside the Suez Canal until the time of the last AIS contact before grounding (i.e., from 05:19 to 05:33 UTC, 23-03-21). (d) Velocity components of the ship along the x- and y-axis over the period pre- and post-grounding. (e) Close-up of the velocity components during the navigation of the ship inside the canal before grounding.

km. Note that the uncertainty about the nominal speed v_0 is taken into account through a suitable choice of the nominal OU process noise standard deviation $\boldsymbol{\sigma}$. Then, the x and y components of the nominal velocity vectors $\boldsymbol{v}^0_{n,j}$, $n = 1, \ldots, N^0_j$, are obtained using the nominal speed and course angle computed for each section of the planned trajectory path as shown in Fig. 4.8 for the trajectory segment inside the canal and before grounding. The same nominal speed is applied to the entire AIS trajectory shown in Fig. 4.9(a).

In addition, the parameters of the underlying OU process were selected following the statistical features estimated for the cargo traffic category in [154]. However, since this configuration of OU parameters is typically used in open-water navigation where the motion of vessels involves larger distances and longer time frames for which a steady-state behavior of the process can be assumed, for this case in the canal we have applied an adjustment factor $\alpha = 200$ to the reversion rate, to account for a shorter time scale with faster

transient dynamics. The same factor was applied to the noise variance so as to keep the same ratio $\boldsymbol{\sigma}^2/\boldsymbol{\gamma}$, shown to be proportional to the asymptotic OU uncertainty [154]. In particular, we set the reversion rate and process noise level as $\bar{\boldsymbol{\gamma}} = [6.28 \times 10^{-4}, 5.63 \times 10^{-4}]$, $\bar{\boldsymbol{\sigma}} = [9.04 \times 10^{-3}, 9.17 \times 10^{-3}]$ outside the canal, and $\boldsymbol{\gamma} = \alpha\bar{\boldsymbol{\gamma}}$, $\boldsymbol{\sigma}^2 = \alpha\bar{\boldsymbol{\sigma}}^2$ inside the canal.

Anomaly detector (4.30) application

The maritime anomaly detection strategy presented in Section 4.4 is applied to the AIS data set of the *Ever Given*, reported from 23-03-21 to 29-03-21, in order to detect an anomaly in its motion before grounding, which occurred at 05:40 UTC, 23-03-21. The AIS track of interest, shown in Fig. 4.9(a), spans over three time periods:

1. the period when the ship is navigating outside the canal,
2. the period when the ship is inside the canal until the grounding,
3. the period post-grounding.

The first AIS contacts starting from 04:13 UTC, 23-03-21 are outside the canal when the ship starts moving on March 23rd towards the South entrance. Then, five key contacts are received and highlighted in our AoI in Fig. 4.9(a), of which four, at 05:19, 05:21, 05:24, and 05:33 UTC, respectively, are before grounding. The fifth data point in the AoI represents the first contact available (07:24 UTC) once the *Ever Given* is stranded. Then, we reported seven AIS contacts of the ship moving northbound after being finally freed on March 29th at 13:04 UTC.

The hypothesis test described in Section 4.4 is run between two consecutive AIS contacts to determine whether or not deviation occurred from the nominal behavior during that time window. The red segments in Fig. 4.9(a) show an anomalous behavior detected just after entrance to the canal at the second AIS contact inside (05:21 UTC, 23-03-21), and again at two other contacts (05:24, 05:33 UTC) immediately before the grounding. The last detection of an anomaly occurred at the first AIS contact (07:24 UTC), which is after the accident. Otherwise, no anomaly is reported before sailing through the Suez Canal or after the accident. Figure 4.9(b) displays the test statistic (4.30) of the anomaly detector over the period pre- and post-grounding, which is shown to exceed the threshold within four intervals of the AIS track with false alarm probability $P_{\mathrm{FA}} = 10^{-4}$ corresponding to the deviation from

the nominal condition. The velocity components of the ship along the x- and y-axis (longitude and latitude, respectively) are displayed for the same time period in Fig. 4.9(d). These components are obtained using the projections of geographical position coordinates in a two-dimensional Cartesian reference system. Figure 4.9(c) shows a close-up of the evolution of decision statistic during the navigation of the ship corresponding to the four contacts inside the Suez Canal and before grounding where a deviation from the nominal behavior was detected in three time intervals. From Fig. 4.9(c), note that the velocity component in the x-direction (orange $\circ$-markers), despite constraints due to the geographical characteristics of the passage, shows a clear alteration from the nominal behavior (orange dot line) at two time points (05:21 and 05:24 UTC) where the velocity is shown to be abnormally higher than expected. This alteration of the velocity along the x direction could be an indication of swerving, that combined with the increased speed along the y direction, could have contributed to the cause of the *Ever Given*'s grounding.

An intuitive analysis on the operating principle behind the anomaly detection strategy (4.30) is provided in Fig. 4.10, where we can see the predicted nominal behavior of the ship (blue solid line), and the predicted position (blue triangle) with related uncertainty (blue ellipse), corresponding to the decision test run for the three contacts inside the canal and before grounding as shown in Fig. 4.9(c). Each prediction in Fig. 4.10 is obtained through the OU dynamic model (4.1) or, equivalently, (4.2) given the nominal parameters, and the time frame between the current and last AIS observation, with the last contact as initial state. The magenta dotted lines in Fig. 4.10 represent the actual deviation between the nominal prediction of the ship and the observed position. Intuitively, an anomaly behavior between two observations is detected through the GLRT (4.30) whenever the actual position of the ship available from the AIS contact falls outside the uncertainty region associated with the predicted nominal position. In practice, this is due to the difference between the nominal and estimated OU velocity parameters within the time window of interest. This difference in velocity is evident in Fig. 4.9(e) for the three AIS contacts at 05:21, 05:24, and 05:33 UTC, which for this reason fall outside the predicted uncertainty region of the nominal position as shown in Fig. 4.10. This, in turn, makes the test statistic exceed the threshold, hence revealing three anomalies as shown in Fig. 4.9(c).

Based on our results derived from real-world data of the *Ever Given*,

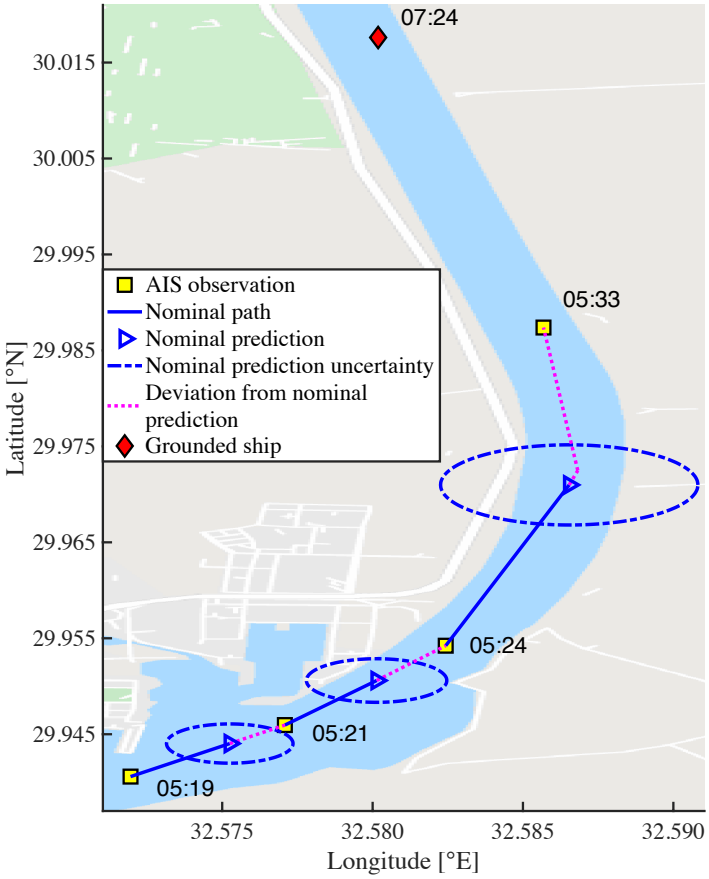


Figure 4.10: Predicted nominal path of the ship (blue solid line) and predicted nominal position (blue triangle) with related uncertainty (blue ellipse) compared to the actual AIS observations (yellow squares) received inside the canal before grounding.

one could see how the ship traveling at a speed statistically higher than the nominal one to be adopted in the Suez Canal, would have triggered the anomaly detector 19 minutes before the grounding.

4.6 Conclusion

In this chapter the maritime anomaly detection problem has been studied assuming an OU mean-reverting stochastic motion model for the vessel dy-

namics. The aim was to reveal a possible deviation of the vessel under consideration from its nominal conditions, during an AIS device disablement, relying on a hypothesis test based on the generalized likelihood ratio decision statistic that builds on the changes in the OU process long-term velocity parameter.

For the sake of clarity the radar model does not consider false alarms and association error with other vessels, leading to a closed form expression for the detector and detection performance expressed as central chi-squared under the nominal condition and non-central chi-squared if anomaly occurs.

A detailed description of the proposed detection strategy, built by exploiting multiple contacts has been provided, presenting both synthetic and real-world analyses.

As confirmed by the numerical analysis, the joint use of radar and AIS information compared to the case of only AIS can lead to a remarkable improvement of detection performance, while it has been shown that under certain conditions detection performance not only does not improve, but actually deteriorates. Such conditions depend on the timing of radar contacts with respect to the differences between nominal and anomalous trajectories.

Moreover, two cases of a real anomalous trajectory have been processed by exploiting the proposed detection strategy allowing to assess its performance. Indeed, the proposed anomaly detection strategy has been applied to the real-world scenarios of the illegal rendezvous nearby the Galápagos Marine Reserve and the *Ever Given* grounding in the Suez Canal. Through our results we show how automatic processing of the sequence of AIS records available from the involved vessel emphasizes the importance of anomaly detection systems in real-world maritime situations.

Chapter 5

Joint detection of anomalous deviations from the expected route and spoofing of dynamic AIS data

The AIS can be subject to intentional reporting of false information, or "spoofing". In this chapter we address the problem of establishing whether a vessel is reporting adulterated dynamic information through AIS messages in order to hide its current planned route and a possible deviation from the nominal route. Multiple hypothesis testing suggests a framework to enlist reliable information from monitoring systems (coastal radars and spaceborne satellite sensors) in support of detection of both anomalies, spoofing and stealth deviations. The proposed solution involves the derivation of anomaly detection rules based on the GLRT and the MOS methodologies. We tested the effectiveness of the proposed anomaly detection strategy for different case studies within an operational scenario with simulated data.[1]

[1]This chapter has been published as "Malicious AIS Spoofing and Abnormal Stealth Deviations: A Comprehensive Statistical Framework for Maritime Anomaly Detection," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 66, no. 24, pp. 6474-6487, Jun. 2021 [66], and as "Maritime Anomaly Detection of Malicious Data Spoofing and Stealth Deviations from Nominal Route Exploiting Heterogeneous Sources of Infor-

5.1 Introduction

This chapter takes into consideration that the AIS's transmission can be easily jammed or manipulated, opening keys to malicious actors to explore attack possibilities [25]. For example, the information contained in AIS messages can be entered manually by the crew of the vessel and maliciously altered, reporting false dynamic information and ruining the reliability of AIS data [117, 124].

We propose to jointly detect dynamic AIS data spoofing [2] and/or surreptitious deviation, where the problem is complicated by the fact that both the deviated trajectory and the spoofed trajectory are *unknown* and *generic*. The test determining whether the AIS data received from a vessel are trustworthy or not, and whether a deviation occurred or not, is supported by additional reliable information provided by surveillance systems such as coastal radars or space-borne sensor imagery [202,203]. In the case that the AIS data are indeed trustworthy, they can be safely used in the data fusion algorithms, e.g., for enhancing the tracking accuracy [91]. If the AIS data are considered to be spoofed, then their fusion with other data should be avoided and an anomaly flag has to be raised.

Building on the work in Chapter 4, in this chapter, we propose a coupled anomaly detection strategy based on a multiple hypothesis testing framework addressed using two approaches. The first approach is based on the GLRT [179] and comprises two GLRTs: the first one decides if AIS data are truthful or not, the second one decides if a deviation occurred or not; finally, the two decisions are combined in a decision rule named C-GLRT. The second approach is based on the MOS methodology [44, 207, 235], which applies an appropriate (complexity-aware) penalty term to the maximized log-likelihood based on the statistical model hypothesized for the vessel kinematics. Common application examples embrace the order selection of Autoregressive Moving Average (ARMA) models, the variable selection in statistical regression models, the channel order estimation in wireless communications, and the order selection of models used to track maneuvering targets in radar systems [54, 70, 116, 190, 226].

mation," in *2021 24th International Conference on Information Fusion*, [67].

[2] In this chapter, *spoofing* is intended as the intentional falsification pursued by the vessel's crew.

5.2 Problem formulation

In this section we formalize the problem, starting with an operational scenario concerning a forbidden area. After, we formalize the detection problem in its most general mathematical form.

Let us consider an area wherein specific types of vessels are not supposed to enter, e.g., an area where fishing is forbidden. In this scenario, a ship equipped with an AIS transponder is expected to follow the planned route sailing around the forbidden area, as depicted in Fig. 5.1(a) with a green dotted line. Assuming the presence of data gaps within a given time window of duration T, four hypotheses can be envisioned:

1. Hypothesis H_0: the vessel respects the nominal condition sailing around the forbidden zone, as in Fig. 5.1(a); it is assumed to move in accordance with a piecewise OU model with a sequence of N_0 long-run mean nominal velocities, comprised in the $2N_0$-dimensional vector $\boldsymbol{v}_0 = \mathrm{col}\left\{\boldsymbol{v}_{0,n}\right\}_{n=1}^{N_0}$, which identifies the nominal N_0-section path. As in Chapter 4, $\boldsymbol{v}_0$ is assumed to be the deterministic parameter of the OU process made available from contextual information;

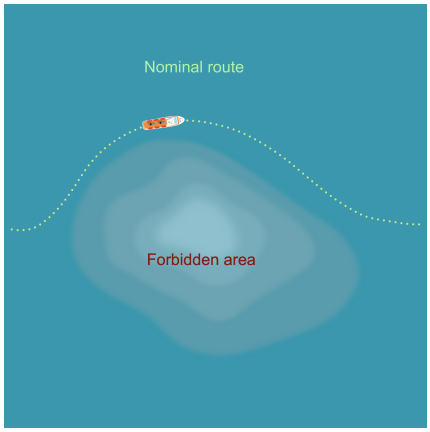
2. Hypothesis H_1: the vessel deviates from the nominal trajectory entering the forbidden zone, as in Fig. 5.1(b); along the deviation, described by the red dotted line in Fig. 5.1, the vessel is assumed to move in accordance with a piecewise OU model with a sequence of N_1 unknown long-run mean velocities, collected in the $2N_1$-dimensional vector $\boldsymbol{v}_R = \mathrm{col}\left\{\boldsymbol{v}_{r,n}\right\}_{n=1}^{N_1}$;

3. Hypothesis H_2: the vessel enters the forbidden zone while reporting that it is bypassing it, as in Fig. 5.1(c); and

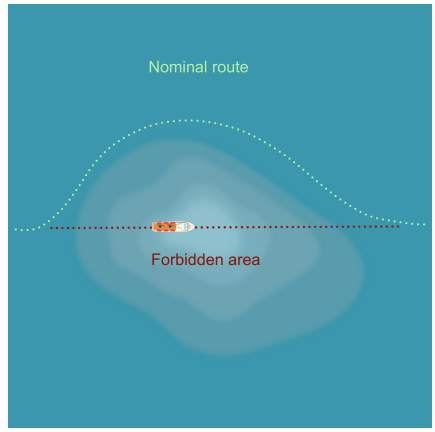
4. Hypothesis H_3: the vessel is just transmitting spoofed AIS messages without an effective deviation from the nominal condition,[3] as in Fig. 5.1(d).

Under hypotheses H_2 and H_3, in the attempt to hide its anomalous behavior, the vessel is assumed to report AIS data generated accordingly to an OU

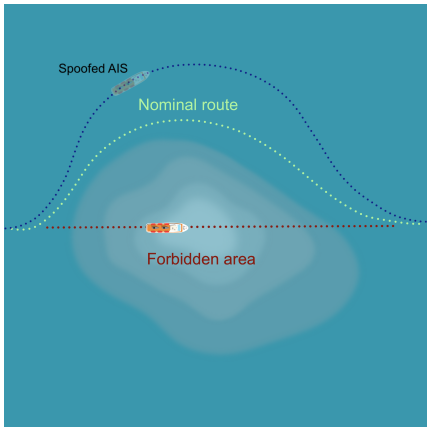
[3]Hypothesis H_3 could be representative of a situation where the vessel at stake navigates under the nominal condition but reports a false position (e.g., spoofing the trajectory of another vessel, as pictorially represented by the blue dotted line in Fig. 5.1(d)). In this way, it seems that the vessel is actually avoiding the forbidden zone. Such a strategy could be adopted by the vessel for hiding a possible illicit rendezvous in the proximity of the nominal route.



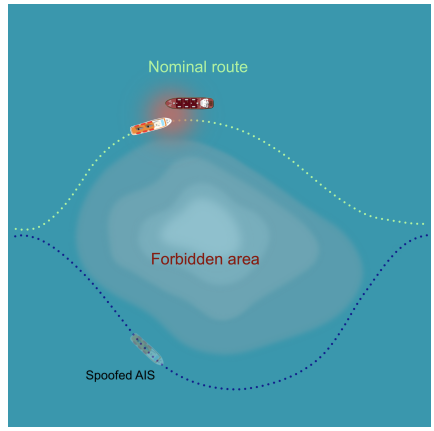
(a) Hypothesis H_0: nominal condition.



(b) Hypothesis H_1: deviation.



(c) Hypothesis H_2: deviation and spoofing.



(d) Hypothesis H_3: spoofing.

Figure 5.1: (a) Hypothesis H_0: the vessel navigates following the nominal route and avoids the forbidden area. (b) Hypothesis H_1: the vessel deviates from the nominal route to broach the forbidden area. (c) Hypothesis H_2: the vessel deviates from the nominal route broaching the forbidden area and transmitting spoofed AIS messages. (d) Hypothesis H_3: the vessel keeps the nominal route transmitting a different position from the actual one in order to mask a rendezvous with another vessel.

process with a $2N_2$-dimensional unknown spoofed velocities vector $\boldsymbol{v}_A = \text{col}\{\boldsymbol{v}_{a,n}\}_{n=1}^{N_2}$, which identifies the spoofed N_2-section path.

Assuming $N = N_0 = N_1 = N_2$ and considering the set of time instants $0 = t_0 < t_1 < \ldots < t_N = T$, the period taken to cover the n-th section in each

of the different N-section path corresponds to the difference $\Delta_n = t_n - t_{n-1}$, $n = 1, \ldots, N$.

Therefore, the multiple hypothesis testing problem addressed in this chapter can be outlined as

$$\begin{cases} H_0: & \boldsymbol{v}_A = \boldsymbol{v}_R = \boldsymbol{v}_0 \\ H_1: & \boldsymbol{v}_A = \boldsymbol{v}_R \neq \boldsymbol{v}_0 \\ H_2: & \boldsymbol{v}_A \neq \boldsymbol{v}_R \neq \boldsymbol{v}_0 \\ H_3: & \boldsymbol{v}_A \neq \boldsymbol{v}_R = \boldsymbol{v}_0, \end{cases} \tag{5.1}$$

and we need to determine which one of the four hypotheses occurred during the considered time window T by means of

- A piecewise OU mean-reverting stochastic motion model to properly describe the dynamic of the vessel (see Chapter 3);
- The awareness of potentially untrustworthy AIS position data;
- The support of trustful measurements associated with the vessel along its trajectory provided by a surveillance system, such as a radar, or a satellite sensor.

The proposed detection strategies will be fully derived in Section 5.4 assuming, without loss of generality, that both N and the sequence $\mathcal{D}_N$ are known to the detectors (as considered in Subsection 4.4.2).

5.3 Statistical representation of target kinematics and sensor observations

We assume there are $K_A + 1$ AIS measurements and $K_R + 1$ measurements provided by a surveillance system, such as a radar, available during the time period of duration T. Recalling (4.13), the generic AIS measurement and the one provided by the surveillance system are respectively defined as

$$\begin{aligned} \boldsymbol{m}_{k_a} &= \boldsymbol{s}(T_{k_a}) + \boldsymbol{n}_{k_a} \quad \text{with } k_a = 0, \ldots, K_A, \\ \boldsymbol{m}_{k_r} &= \boldsymbol{s}(T_{k_r}) + \boldsymbol{n}_{k_r} \quad \text{with } k_r = 0, \ldots, K_R. \end{aligned} \tag{5.2}$$

The measurement noise terms $\boldsymbol{n}_{k_a}$ and $\boldsymbol{n}_{k_r}$ are assumed to be independent and identically distributed according to a zero-mean Gaussian with covariance matrices $\mathbf{C}_{\boldsymbol{n}_{k_a}}$ and $\mathbf{C}_{\boldsymbol{n}_{k_r}}$, respectively. The generic measurements

in (5.2) are available at times $T_{k_a} = \frac{p_a}{N}T$ and $T_{k_r} = \frac{p_r}{N}T$ where $p_{a,r} \in [0, N]$ is a fraction of the interval $[0, N]$ representing the time location of the contact with respect to the N piecewise OU velocities.

As done in Subsection 4.3.2, we consider the data vectors

$$\boldsymbol{y}_A = \mathrm{col}\left\{\boldsymbol{y}_{k_a}\right\}_{k_a=1}^{K_A} \quad \text{and} \quad \boldsymbol{y}_R = \mathrm{col}\left\{\boldsymbol{y}_{k_r}\right\}_{k_r=1}^{K_R}, \tag{5.3}$$

respectively built from the AIS measurements and the ones provided by the surveillance system, with

$$\boldsymbol{y}_{k_a} = \boldsymbol{m}_{k_a} - \boldsymbol{\Phi}(T_{k_a})\boldsymbol{m}_0, \tag{5.4}$$

$$\boldsymbol{y}_{k_r} = \boldsymbol{m}_{k_r} - \boldsymbol{\Phi}(T_{k_r})\boldsymbol{m}_0, \tag{5.5}$$

where $\boldsymbol{m}_0$ is the first AIS contact that can be assumed reliable (or the first measurement provided by the surveillance system), and corresponds to the beginning of the considered observation window.

Just like in Section 4.3 (see (4.14)), the data vectors $\boldsymbol{y}_A$ and $\boldsymbol{y}_R$ in (5.3) can be expressed as follows

$$\boldsymbol{y}_A = \mathbf{H}_A\,\boldsymbol{v}_A + \boldsymbol{\omega}_A \sim \mathcal{N}\left(\mathbf{H}_A\,\boldsymbol{v}_A, \mathbf{C}_{\boldsymbol{y}_A}\right), \tag{5.6}$$

$$\boldsymbol{y}_R = \mathbf{H}_R\,\boldsymbol{v}_R + \boldsymbol{\omega}_R \sim \mathcal{N}\left(\mathbf{H}_R\,\boldsymbol{v}_R, \mathbf{C}_{\boldsymbol{y}_R}\right), \tag{5.7}$$

where $\mathbf{H}_A$ and $\mathbf{H}_R$ are the model matrices encompassing the state transition matrices (B.6) and the control input functions (B.8), whose components can be found in (4.22). The terms $\boldsymbol{\omega}_A$ and $\boldsymbol{\omega}_R$ are the noise terms of the OU process, which are distributed according to a zero-mean Gaussian with covariance $\mathbf{C}_{\boldsymbol{y}_A}$ and $\mathbf{C}_{\boldsymbol{y}_R}$, respectively, whose components are given in (4.20) and (4.21).

The data vectors $\boldsymbol{y}_A$ and $\boldsymbol{y}_R$ can be combined into a $4K$-dimensional data vector $\boldsymbol{y}$, with $K = K_A + K_R$, by encapsulating all the available information about the target trajectory. Vector $\boldsymbol{y}$ is structured as follows

$$\begin{aligned}\boldsymbol{y} = \begin{bmatrix}\boldsymbol{y}_A \\ \boldsymbol{y}_R\end{bmatrix} &= \begin{bmatrix}\mathbf{H}_A \\ \mathbf{0}\end{bmatrix}\boldsymbol{\varphi} + \begin{bmatrix}\mathbf{H}_A \\ \mathbf{H}_R\end{bmatrix}\boldsymbol{v}_R + \boldsymbol{\omega} \\ &= \mathbf{G}\,\boldsymbol{\varphi} + \mathbf{H}\,\boldsymbol{v}_R + \boldsymbol{\omega} \\ &= \mathbf{M}\,\boldsymbol{z} + \boldsymbol{\omega} \sim \mathcal{N}\left(\mathbf{M}\,\boldsymbol{z}, \mathbf{C}_{\boldsymbol{y}}\right),\end{aligned} \tag{5.8}$$

where $\boldsymbol{z} \triangleq [\boldsymbol{\varphi}^{\mathrm{T}},\ \boldsymbol{v}_R^{\mathrm{T}}]^{\mathrm{T}}$ is the unknown vector including the *spoofing parameter* $\boldsymbol{\varphi} \triangleq \boldsymbol{v}_A - \boldsymbol{v}_R$ and the actual velocity sequence $\boldsymbol{v}_R$, while

$$\mathbf{M} \triangleq [\mathbf{G} \quad \mathbf{H}] = \begin{bmatrix}\mathbf{H}_A & \mathbf{H}_A \\ \mathbf{0} & \mathbf{H}_R\end{bmatrix}, \tag{5.9}$$

is the known $4K \times 4N$ model matrix including the sub-matrices $\mathbf{H}_A$ and $\mathbf{H}_R$.

The OU process noise $\boldsymbol{\omega}$ is distributed as a zero-mean Gaussian with covariance

$$\mathbf{C}_{\boldsymbol{y}} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{y}_A} & \mathbf{C}_{\boldsymbol{y}_A,\boldsymbol{y}_R} \\ \mathbf{C}_{\boldsymbol{y}_R,\boldsymbol{y}_A} & \mathbf{C}_{\boldsymbol{y}_R} \end{bmatrix}, \tag{5.10}$$

whose components are (see also C.2)

$$\mathbf{C}_{ij} = \begin{cases} \mathbf{C}(T_i) + \mathbf{C}_{\boldsymbol{n}_i} + \boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_i)^{\mathrm{T}} & \text{if } T_i = T_j, \\ \mathbf{C}(T_i)\boldsymbol{\Phi}(T_j - T_i)^{\mathrm{T}} + \boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_j)^{\mathrm{T}} & \text{if } T_i < T_j, \\ \mathbf{C}(T_j)\boldsymbol{\Phi}(T_i - T_j)^{\mathrm{T}} + \boldsymbol{\Phi}(T_j)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_i)^{\mathrm{T}} & \text{if } T_i > T_j, \end{cases} \tag{5.11}$$

$\forall i = 1, \ldots, K_A$, and $\forall j = 1, \ldots, K_R$, and with $\mathbf{C}_{ji} = \mathbf{C}_{ij}^{\mathrm{T}}$.

5.4 Detection strategies

Two different anomaly detection strategies are developed for the detection problem described in Section 5.2, both based on the changes in the OU process long-run mean velocity parameter. The first strategy aims at determining whether the trajectory reported by the vessel via AIS is trustworthy or not by means of a composite hypothesis test based on the spoofing parameter $\boldsymbol{\varphi}$, and addressed with a GLRT approach [179]; in a similar way, a further GLRT is exploited in order to detect a possible deviation from the nominal route relying on the unknown velocity vector $\boldsymbol{v}_R$.

The second strategy is designed to tackle jointly both the spoofing and the deviation events from the nominal route. In this case, the MOS methodology [207] is used to deal with the resulting multiple hypotheses test, relying on the assumption that one of the *nested* hypotheses H_i, $i = 0, 1, 2, 3$, is true.

Here the AIS contacts are augmented by reliable ones from a surveillance system, and hence – differently from Chapter 4.4 – spoofing can be detected. Indeed, without any reliable measurement, matrix $\mathbf{H}_R$ would be null and the spoofing parameter $\boldsymbol{\varphi}$ would not be distinguishable from the deviation velocity $\boldsymbol{v}_R$.

5.4.1 Multiple hypothesis test based on coupled GLRTs

Spoofing detector

The AIS spoofing detection problem is addressed in a binary hypothesis testing framework as

$$\begin{cases} \{H_0, H_1\}: & \boldsymbol{y} = \mathbf{H}\,\boldsymbol{v}_R + \boldsymbol{\omega} \sim \mathcal{N}\left(\mathbf{H}\,\boldsymbol{v}_R, \mathbf{C}_{\boldsymbol{y}}\right) \\ \{H_2, H_3\}: & \boldsymbol{y} = \mathbf{G}\,\boldsymbol{\varphi} + \mathbf{H}\,\boldsymbol{v}_R + \boldsymbol{\omega} \sim \mathcal{N}\left(\mathbf{G}\,\boldsymbol{\varphi} + \mathbf{H}\,\boldsymbol{v}_R, \mathbf{C}_{\boldsymbol{y}}\right). \end{cases} \tag{5.12}$$

We wish to test if $\boldsymbol{\varphi} = \mathbf{0}$ as opposed to $\boldsymbol{\varphi} \neq \mathbf{0}$; hence the parameter test is

$$\begin{cases} \{H_0, H_1\}: & \boldsymbol{\varphi} = \mathbf{0} \\ \{H_2, H_3\}: & \boldsymbol{\varphi} \neq \mathbf{0}, \end{cases} \tag{5.13}$$

where $\{H_0, H_1\}$ is the hypothesis that the trajectory reported by the AIS messages is the actual one, and $\{H_2, H_3\}$ is the alternative hypothesis that the trajectory reported by the AIS messages is spoofed. Denoted with $p\left(\boldsymbol{y}; \widehat{\boldsymbol{z}} | \{H_0, H_1\}\right)$ and $p\left(\boldsymbol{y}; \widehat{\boldsymbol{z}} | \{H_2, H_3\}\right)$ the likelihood functions under the hypotheses $\{H_0, H_1\}$ and $\{H_2, H_3\}$, respectively, the generalized likelihood ratio for this problem is

$$\frac{p\left(\boldsymbol{y}; \widehat{\boldsymbol{z}} | \{H_2, H_3\}\right)}{p\left(\boldsymbol{y}; \widehat{\boldsymbol{z}} | \{H_0, H_1\}\right)} = \frac{p\left(\boldsymbol{y}; \widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{v}}_R | \{H_2, H_3\}\right)}{p\left(\boldsymbol{y}; \boldsymbol{\varphi} = \mathbf{0}, \widehat{\boldsymbol{v}}_R^{(0)} | \{H_0, H_1\}\right)}, \tag{5.14}$$

where $\widehat{\boldsymbol{z}} = [\widehat{\boldsymbol{\varphi}}^{\mathrm{T}}, \widehat{\boldsymbol{v}}_R^{\mathrm{T}}]^{\mathrm{T}}$ is the MLE of $\boldsymbol{z}$ under $\{H_2, H_3\}$, and $\widehat{\boldsymbol{v}}_R^{(0)}$ is the MLE of $\boldsymbol{v}_R$ under hypothesis $\{H_0, H_1\}$ (i.e., when subject to the constraint $\boldsymbol{\varphi} = \mathbf{0}$).

As in Subsection 4.4.2, in order to tackle ill-conditioned[4] matrix inversions, the reduced rank version of matrices $\mathbf{G}$ and $\mathbf{H}$ via SVD [209] is considered. Specifically, defining $\rho_G = \mathrm{rank}(\mathbf{G})$ and $\rho_H = \mathrm{rank}(\mathbf{H})$ the corresponding number of nonzero singular values of $\mathbf{G}$ and $\mathbf{H}$, respectively, we can get

$$\mathbf{G}\,\boldsymbol{\varphi} = \mathbf{U}_G\,\mathbf{S}_G\,\mathbf{Q}_G^{\mathrm{T}}\,\boldsymbol{\varphi} = \widetilde{\mathbf{U}}_G\,\widetilde{\mathbf{S}}_G\,\widetilde{\mathbf{Q}}_G^{\mathrm{T}}\,\boldsymbol{\varphi} = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{\varphi}}, \tag{5.15}$$

$$\mathbf{H}\,\boldsymbol{v}_R = \mathbf{U}_H\,\mathbf{S}_H\,\mathbf{Q}_H^{\mathrm{T}}\,\boldsymbol{v}_R = \widetilde{\mathbf{U}}_H\,\widetilde{\mathbf{S}}_H\,\widetilde{\mathbf{Q}}_H^{\mathrm{T}}\,\boldsymbol{v}_R = \widetilde{\mathbf{U}}_H\,\widetilde{\boldsymbol{v}}_R, \tag{5.16}$$

where $\widetilde{\mathbf{U}}_{G,H}$, of size $4K \times \rho_{G,H}$, and $\widetilde{\mathbf{Q}}_{G,H}$, of size $2N \times \rho_{G,H}$, are the matrices of the left-singular vectors and the right-singular vectors of $\mathbf{G}$ and

[4] Note that the ill-conditioning strongly depends on the number and the position of the contacts.

$\mathbf{H}$, respectively, whereas $\widetilde{\mathbf{S}}_{G,H}$ is a $\rho_{G,H} \times \rho_{G,H}$ diagonal matrix.

$$\widetilde{\boldsymbol{\varphi}} \triangleq \widetilde{\mathbf{S}}_G \, \widetilde{\mathbf{Q}}_G^{\mathrm{T}} \, \boldsymbol{\varphi}, \quad \text{and} \quad \widetilde{\boldsymbol{v}}_R \triangleq \widetilde{\mathbf{S}}_H \, \widetilde{\mathbf{Q}}_H^{\mathrm{T}} \, \boldsymbol{v}_R \tag{5.17}$$

are respectively the ρ_G-dimensional and ρ_H-dimensional unknown vectors resulting from the rotation and scaling of the vectors $\boldsymbol{\varphi}$ and $\boldsymbol{v}_R$. In addition, $\widetilde{\boldsymbol{v}}_0 \triangleq \widetilde{\mathbf{S}}_H \, \widetilde{\mathbf{Q}}_H^{\mathrm{T}} \, \boldsymbol{v}_0$ is the rotated and scaled version of the nominal velocity vector. Accordingly, the data model (5.8) can be rewritten as

$$\boldsymbol{y} = \widetilde{\mathbf{U}}_G \, \widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{U}}_H \, \widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega} = \begin{bmatrix} \widetilde{\mathbf{U}}_G & \widetilde{\mathbf{U}}_H \end{bmatrix} \begin{bmatrix} \widetilde{\boldsymbol{\varphi}} \\ \widetilde{\boldsymbol{v}}_R \end{bmatrix} + \boldsymbol{\omega} = \widetilde{\mathbf{U}} \, \widetilde{\boldsymbol{z}} + \boldsymbol{\omega}, \tag{5.18}$$

with $\widetilde{\mathbf{U}} \triangleq [\widetilde{\mathbf{U}}_G, \widetilde{\mathbf{U}}_H]$ matrix of size $4K \times (\rho_G + \rho_H)$ and $\widetilde{\boldsymbol{z}} \triangleq [\widetilde{\boldsymbol{\varphi}}^{\mathrm{T}}, \widetilde{\boldsymbol{v}}_R^{\mathrm{T}}]^{\mathrm{T}}$ column vector of size $\rho_G + \rho_H$.

Therefore, considering (5.14)d, and denoting the threshold with τ', the *spoofing GLRT* can be formulated as

$$L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) = \frac{p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{\varphi}}}, \widehat{\widetilde{\boldsymbol{v}}}_R | \, \{H_2, H_3\}\right)}{p\left(\boldsymbol{y}; \widetilde{\boldsymbol{\varphi}} = \mathbf{0}, \widehat{\widetilde{\boldsymbol{v}}}_R^{(0)} | \, \{H_0, H_1\}\right)} \underset{\{H_0,H_1\}}{\overset{\{H_2,H_3\}}{\gtrless}} \tau', \tag{5.19}$$

whose explicit form (see Appendix D.1 for the derivation) is:

$$L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) = \widehat{\widetilde{\boldsymbol{\varphi}}}^{\mathrm{T}} \mathbf{C}_{\widehat{\widetilde{\boldsymbol{\varphi}}}}^{-1} \, \widehat{\widetilde{\boldsymbol{\varphi}}} \underset{\{H_0,H_1\}}{\overset{\{H_2,H_3\}}{\gtrless}} \tau', \tag{5.20}$$

where $\widehat{\widetilde{\boldsymbol{\varphi}}} \sim \mathcal{N}\left(\widetilde{\boldsymbol{\theta}}, \mathbf{C}_{\widehat{\widetilde{\boldsymbol{\varphi}}}}\right)$ is the MLE of $\widetilde{\boldsymbol{\varphi}}$ under hypothesis $\{H_2, H_3\}$, whose expression is fully derived in Appendix D.2.

The test statistic under the two hypotheses $\{H_0, H_1\}$ and $\{H_2, H_3\}$ are respectively characterized by a central and a non-central Chi-squared distributions. The detection performance of the GLRT (5.20) is described by (4.27), where, in this case, P_{FA} is the probability that the test statistic exceeds τ' under $\{H_0, H_1\}$, and P_{D} is the probability that the test statistic exceeds τ' under $\{H_2, H_3\}$. The number of degrees of freedom is $d = \mathrm{rank}\left(\mathbf{C}_{\widehat{\widetilde{\boldsymbol{\varphi}}}}\right)$, and the non-centrality parameter $\lambda = \lambda(\widetilde{\boldsymbol{\varphi}})$ is given by

$$\lambda\left(\widetilde{\boldsymbol{\varphi}}\right) = \widetilde{\boldsymbol{\varphi}}^{\mathrm{T}} \mathbf{C}_{\widehat{\widetilde{\boldsymbol{\varphi}}}}^{-1} \, \widetilde{\boldsymbol{\varphi}}, \tag{5.21}$$

with $\widetilde{\boldsymbol{\varphi}}$ the true value of the spoofing parameter under hypothesis $\{H_2, H_3\}$ provided in (5.15). In this regard, $\lambda(\widetilde{\boldsymbol{\varphi}})$ represents the *spoofing distance*, i.e., it is a measure of how much the two hypotheses $\{H_0, H_1\}$ and $\{H_2, H_3\}$ are disjoint, indeed the higher $\lambda(\widetilde{\boldsymbol{\varphi}})$, the easier it is to detect the spoofing.

Deviation detector

In order to decide whether the deviation from the nominal condition occurred or not, we need to consider the following parameter test

$$\begin{cases} \{H_0, H_3\}: & \widetilde{\boldsymbol{v}}_R = \widetilde{\boldsymbol{v}}_0 \\ \{H_1, H_2\}: & \widetilde{\boldsymbol{v}}_R \neq \widetilde{\boldsymbol{v}}_0, \end{cases} \tag{5.22}$$

Denoting the threshold with τ'', the resulting *deviation GLRT* is given by (see Appendix D.1 for the derivation)

$$L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) = \left(\widehat{\widetilde{\boldsymbol{v}}}_R - \widetilde{\boldsymbol{v}}_0\right)^{\mathrm{T}} \mathbf{C}^{-1}_{\widehat{\widetilde{\boldsymbol{v}}}_R} \left(\widehat{\widetilde{\boldsymbol{v}}}_R - \widetilde{\boldsymbol{v}}_0\right) \underset{\{H_0,H_3\}}{\overset{\{H_1,H_2\}}{\gtrless}} \tau'', \tag{5.23}$$

where $\widehat{\widetilde{\boldsymbol{v}}}_R \sim \mathcal{N}\left(\widetilde{\boldsymbol{v}}_R, \mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_R}\right)$ is the MLE of $\widetilde{\boldsymbol{v}}_R$ under the deviation hypotheses, whose expression is derived in Appendix D.2. The detection performance of the GLRT (5.23) is still described by (4.27), with the difference that P_{FA} is the probability that the test statistic exceeds τ'' under $\{H_0, H_3\}$, and P_{D} is the probability that the test statistic exceeds τ'' under $\{H_1, H_2\}$. In this case, the number of degrees of freedom is $d = \operatorname{rank}\left(\mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_R}\right)$, and the non-centrality parameter $\lambda = \lambda(\widetilde{\boldsymbol{v}}_R)$ is given by

$$\lambda\left(\widetilde{\boldsymbol{v}}_R\right) = \left(\widetilde{\boldsymbol{v}}_R - \widetilde{\boldsymbol{v}}_0\right)^{\mathrm{T}} \mathbf{C}^{-1}_{\widehat{\widetilde{\boldsymbol{v}}}_R} \left(\widetilde{\boldsymbol{v}}_R - \widetilde{\boldsymbol{v}}_0\right), \tag{5.24}$$

with $\widetilde{\boldsymbol{v}}_R$ the true value of the velocity vector under the deviation hypotheses $\{H_1, H_2\}$. In this case, $\lambda(\widetilde{\boldsymbol{v}}_R)$ represents the *deviation distance*, i.e., it is a measure of how much the two hypotheses $\{H_0, H_3\}$ and $\{H_1, H_2\}$ are disjoint; indeed the higher $\lambda(\widetilde{\boldsymbol{v}}_R)$, the easier it is to detect the deviation.

It is worthwhile remarking that, for both the GLRTs (5.20) and (5.23), we are able to set the false alarm probability and compute the test performance in closed form.

Coupled GLRT rule

A test based on the GLRTs of spoofing and deviation, referred as C-GLRT, is proposed to decide H_i, $i = 0, 1, 2, 3$. The decisions spoofing versus no-spoofing and deviation versus no-deviation are compared to finally decide

Table 5.1: C-GLRT rule scheme

Decision in (5.20) \ Decision in (5.23)	$\{H_1, H_2\}$	$\{H_0, H_3\}$
$\{H_2, H_3\}$	H_2	H_3
$\{H_0, H_1\}$	H_1	H_0

about H_i, $i = 0, 1, 2, 3$. The C-GLRT decision rule is defined as follows

$$\begin{cases} \text{decide } H_0 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) < \tau'' \\ \text{decide } H_1 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) > \tau'' \\ \text{decide } H_2 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) > \tau'' \\ \text{decide } H_3 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) < \tau''. \end{cases} \tag{5.25}$$

The C-GLRT decides for H_i given by the intersection of the two decisions taken by the GLRTs (5.20) and (5.23), as shown in Table 5.1.

Optimal selection of the C-GLRT thresholds

It is worthwhile to note that the C-GLRT has two design parameters, namely the false alarm probabilities of the single GLRTs, that need to be selected.

Assuming equiprobable prior hypotheses, the probability of correct decision is given by

$$P_{\mathrm{C}} = \sum_{i=0}^{3} P(H_i|H_i)P(H_i) = \frac{1}{4}\sum_{i=0}^{3} P(H_i|H_i), \tag{5.26}$$

where $P_i = P(H_i|H_i)$ is the probability of deciding H_i when H_i is true, and $P(H_i) = 1/4$ is the prior probability. P_i depends on the joint decision of the two GLRTs (5.25), and, thus, on the joint selection of the thresholds τ' and τ''. Consequently, assuming one has the analytical expression of the probabilities P_i, it is possible to optimally select the two thresholds in order to maximize P_{C}. In the absence of an analytical expression of the probabilities P_i, an approximate expression for the average P_{C} is as follows.

Assume that the non-centrality parameter of the anomalous deviation $\lambda(\widetilde{\boldsymbol{v}}_R)$ is large enough such that the GLRT in (5.23) can operate with negligible errors. In other words, the conditional probabilities of correct decision can be approximated as follows:

$$\begin{array}{lcll} P(H_0|H_0) & \approx & P(L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau') & = & 1 - P_{\mathrm{FA}}(\tau'), \\ P(H_1|H_1) & \approx & P(L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau') & = & 1 - P_{\mathrm{FA}}(\tau'), \\ P(H_2|H_2) & \approx & P(L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau') & = & P_{\mathrm{D}}(\tau'), \\ P(H_3|H_3) & \approx & P(L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau') & = & P_{\mathrm{D}}(\tau'). \end{array} \tag{5.27}$$

Therefore, the correct decision event is approximated as only based on the spoofing (5.20):

$$P_{\mathrm{C}} \approx \frac{1}{2}\left(P_{\mathrm{D}}(\tau') + 1 - P_{\mathrm{FA}}(\tau')\right), \tag{5.28}$$

with $P_{\mathrm{FA}}(\tau') = Q_{\chi^2_d}(\tau')$ and $P_{\mathrm{D}}(\tau') = Q_{\chi'^2_d(\lambda)}(\tau')$. Alternatively, we can proceed if the spoofing non-centrality parameter is large enough to neglect the error in (5.20). In this case, the correct decision probability would depend on the detection and false alarm probabilities of the anomalous deviation.

In a given scenario, reported in the experimental Section 5.6, we show how it is possible to optimally select both the false alarm probabilities to maximize the probability of correct decision assuming to know the values of the non-centrality parameters $\lambda(\widetilde{\boldsymbol{\varphi}})$ and $\lambda(\widetilde{\boldsymbol{v}}_R)$.

5.4.2 Multiple hypothesis test based on the MOS rules

Here we apply the MOS methodology to the original multiple hypothesis test (5.1), previously introduced in Section 5.2:

$$\begin{cases} H_0: & \boldsymbol{y} = \mathbf{H}\,\boldsymbol{v}_0 + \boldsymbol{\omega} \\ H_1: & \boldsymbol{y} = \mathbf{H}\,\boldsymbol{v}_R + \boldsymbol{\omega} \\ H_2: & \boldsymbol{y} = \mathbf{M}\,\boldsymbol{z} + \boldsymbol{\omega} \\ H_3: & \boldsymbol{y} = \mathbf{G}\,\boldsymbol{\varphi} + \mathbf{H}\,\boldsymbol{v}_0 + \boldsymbol{\omega}. \end{cases} \tag{5.29}$$

As done in Subsection 5.4.1, in addition to the SVDs (5.15) and (5.16), we also consider the following equality

$$\mathbf{M}\,\boldsymbol{z} = \mathbf{U}_M\,\mathbf{S}_M\,\mathbf{Q}_M^{\mathrm{T}}\,\boldsymbol{z} = \widetilde{\mathbf{U}}_M\,\widetilde{\mathbf{S}}_M\,\widetilde{\mathbf{Q}}_M^{\mathrm{T}}\,\boldsymbol{z} = \widetilde{\mathbf{U}}_M\,\widetilde{\boldsymbol{z}}_M, \tag{5.30}$$

where $\widetilde{\mathbf{U}}_M$, of size $4K \times \rho_M$, and $\widetilde{\mathbf{Q}}_M$, of size $4N \times \rho_M$, are, respectively, the matrices of the left-singular vectors and the right-singular vectors of $\mathbf{M}$,

whereas $\widetilde{\mathbf{S}}_M$ is a $\rho_M \times \rho_M$ diagonal matrix. The term $\widetilde{\boldsymbol{z}}_M \triangleq \widetilde{\mathbf{S}}_M \widetilde{\mathbf{Q}}_M^{\mathrm{T}} \boldsymbol{z}$ is the ρ_M-dimensional unknown vector resulting from the rotation and scaling of vector $\boldsymbol{z}$.[5] Consequently, we can reformulate the corresponding multiple hypothesis testing (5.29) as

$$\begin{cases} H_0: & \boldsymbol{y} = \widetilde{\mathbf{U}}_H \, \widetilde{\boldsymbol{v}}_0 + \boldsymbol{\omega} \\ H_1: & \boldsymbol{y} = \widetilde{\mathbf{U}}_H \, \widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega} \\ H_2: & \boldsymbol{y} = \widetilde{\mathbf{U}}_M \, \widetilde{\boldsymbol{z}}_M + \boldsymbol{\omega} \\ H_3: & \boldsymbol{y} = \widetilde{\mathbf{U}}_G \, \widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{U}}_H \, \widetilde{\boldsymbol{v}}_0 + \boldsymbol{\omega}. \end{cases} \tag{5.31}$$

The MOS methodology looks for the best dimension of the parametric model given the observed data; its use is required because the different hypotheses in (5.31) are nested and show different numbers of involved parameters. Given the set of hypotheses H_i, with $i = 0, 1, 2, 3$, assumed mutually exclusive, the MOS refers to the multiple hypothesis testing problem of establishing which model best outlines the given data $\boldsymbol{y}$. More specifically, assuming there is a single correct hypothesis within the considered set, and since each hypothesis identifies a probability density function with a parameter vector $\boldsymbol{\varrho}_i$ of dimension $n_i \triangleq \dim(\boldsymbol{\varrho}_i | H_i)$, denoted as *model order*, then, the MOS is about finding the true dimension n_i of the model.

In its most general form, the MOS decides for hypothesis H_m if

$$\xi_i = -2 \ln p\left(\boldsymbol{y}; \widehat{\boldsymbol{\varrho}}_i\right) + \eta(n_i, K) \tag{5.32}$$

is minimized for $i = m$. The term $\ln p\left(\boldsymbol{y}; \widehat{\boldsymbol{\varrho}}_i\right)$ is the maximum log-likelihood function of the data $\boldsymbol{y}$ under hypothesis H_i, while the term $\eta_i = \eta(n_i, K)$ is an increasing function of n_i whose purpose is penalizing the overfitting of the observed data, so that $-2 \ln p\left(\boldsymbol{y}; \widehat{\boldsymbol{\varrho}}_i\right)$ decreases as such *penalty coefficient* increases yielding a trade-off between goodness of fit and simplicity [207].

Among all the different criteria to choose the penalty term η_i, we take into consideration the direct KL approach, also known as No-Name (NN) rule, the Akaike's Information Criteria (AIC), and the Bayesian Information Criteria (BIC). While the NN and AIC have their foundations in the information theory and coding [14,15,125], BIC method has its roots in major statistical inference paradigms [194].

Specifying the expression for the term $-2 \ln p\left(\boldsymbol{y}; \widehat{\boldsymbol{\varrho}}_i\right)$ in (5.32) under each hypothesis H_i, the MOS rules decide for hypothesis $H_i = \arg\min_j \xi_j$,

[5] Note that $\widetilde{\mathbf{U}}_M \neq \left[\widetilde{\mathbf{U}}_G, \, \widetilde{\mathbf{U}}_H\right]$ and $\widetilde{\boldsymbol{z}}_M \neq \widetilde{\boldsymbol{z}}$.

$j = 0, 1, 2, 3$. Taking into account the Cholesky decomposition (4.34), ξ_j is defined as follows

$$\begin{cases} \xi_0 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{H}}\,\widetilde{\boldsymbol{v}}_0\|^2 + \eta_0 \\ \xi_1 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{H}}\,\widehat{\widetilde{\boldsymbol{v}}}_R\|^2 + \eta_1 \\ \xi_2 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{M}}\,\widehat{\widetilde{\boldsymbol{z}}}_M\|^2 + \eta_2 \\ \xi_3 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{G}}\,\widehat{\widetilde{\boldsymbol{\varphi}}} - \widetilde{\mathbf{H}}\,\widetilde{\boldsymbol{v}}_0\|^2 + \eta_3, \end{cases} \tag{5.33}$$

where the MLEs,[6] based on the given hypothesis H_i, are given by

$$\begin{aligned} H_1\text{:}\quad \widehat{\widetilde{\boldsymbol{v}}}_R &= \left(\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{C}}\,\boldsymbol{y} && \text{with } \widetilde{\mathbf{H}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_H, \\ H_2\text{:}\quad \widehat{\widetilde{\boldsymbol{z}}}_M &= \left(\widetilde{\mathbf{M}}^{\mathrm{T}}\,\widetilde{\mathbf{M}}\right)^{-1}\widetilde{\mathbf{M}}^{\mathrm{T}}\,\widetilde{\mathbf{C}}\,\boldsymbol{y} && \text{with } \widetilde{\mathbf{M}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_M, \\ H_3\text{:}\quad \widehat{\widetilde{\boldsymbol{\varphi}}} &= \left(\widetilde{\mathbf{G}}^{\mathrm{T}}\,\widetilde{\mathbf{G}}\right)^{-1}\widetilde{\mathbf{G}}^{\mathrm{T}}\left(\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{H}}\,\widetilde{\boldsymbol{v}}_0\right) && \text{with } \widetilde{\mathbf{G}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_G. \end{aligned} \tag{5.34}$$

The penalty coefficient η_i is specified for each criterion and respectively related to the hypothesis H_i, $i = 0, 1, 2, 3$, in Table 5.2.

Table 5.2: MOS penalty coefficients
(Four hypotheses configuration)

	η_0	η_1	η_2	η_3
NN	0	ρ_H	ρ_M	ρ_G
AIC	0	$2\rho_H$	$2\rho_M$	$2\rho_G$
BIC	0	$\ln(4K)2\rho_H$	$\ln(4K)2\rho_M$	$\ln(4K)2\rho_G$

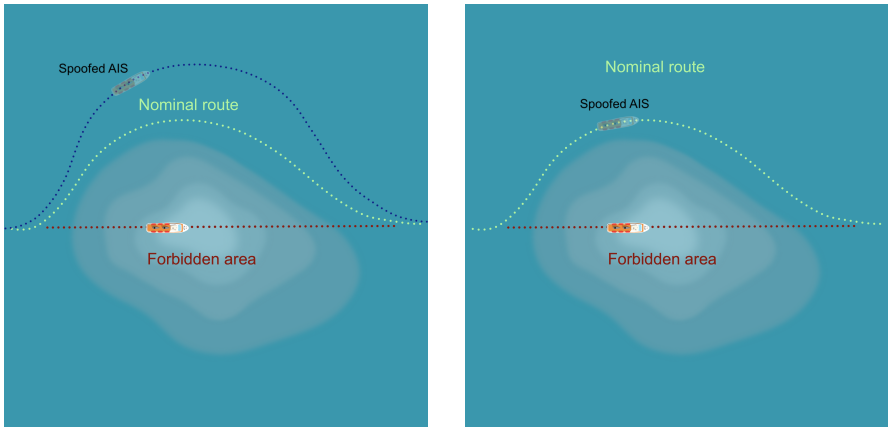
[6] Under hypothesis H_0, the nominal velocity sequence $\widetilde{\boldsymbol{v}}_0$ is deterministic and known, so that $\widehat{\widetilde{\boldsymbol{v}}}_0 = \widetilde{\boldsymbol{v}}_0$.

5.5 Extension to a five hypotheses configuration

To further investigate the combined occurrence of the AIS spoofing and the deviation from the nominal condition, it is possible to extend the multiple hypothesis test problem (5.1) to five hypotheses. Specifically, when spoofing and deviation anomalies arise simultaneously, we can take into account a particular case of hypothesis H_2, and discern two events:

- H_2: the vessel deviates from the nominal route broaching the forbidden area and transmitting spoofed AIS messages reporting a velocity profile more or less close to the nominal one, as in Fig. 5.2(a);

- H_4: The vessel deviates from the nominal route broaching the forbidden area and transmitting spoofed AIS messages reporting the nominal velocity profile, as in Fig. 5.2(b). In this last case the vessel tries to hide the real movements in the most stealthy way from the surveillance system.



(a) Hypothesis H_2: $\boldsymbol{v}_A \neq \boldsymbol{v}_0$.

(b) Hypothesis H_4: $\boldsymbol{v}_A = \boldsymbol{v}_0$.

Figure 5.2: Visual lineup between hypotheses H_2 and H_4. (a) The spoofed AIS data report a generic velocity profile, more ore less close to the nominal one, i.e., $\boldsymbol{v}_A \neq \boldsymbol{v}_0$. (b) The spoofed AIS data trace the nominal velocity profile, i.e., $\boldsymbol{v}_A = \boldsymbol{v}_0$.

The resulting extended multiple hypotheses testing problem is

$$\begin{cases} H_0: & \boldsymbol{v}_A = \boldsymbol{v}_R = \boldsymbol{v}_0 \\ H_1: & \boldsymbol{v}_A = \boldsymbol{v}_R \neq \boldsymbol{v}_0 \\ H_2: & \boldsymbol{v}_A \neq \boldsymbol{v}_R \neq \boldsymbol{v}_0 \\ H_3: & \boldsymbol{v}_A \neq \boldsymbol{v}_R = \boldsymbol{v}_0 \\ H_4: & \boldsymbol{v}_A = \boldsymbol{v}_0 \neq \boldsymbol{v}_R \end{cases} \longrightarrow \begin{cases} H_0: & \boldsymbol{y} = \mathbf{H}\,\boldsymbol{v}_0 + \boldsymbol{\omega} \\ H_1: & \boldsymbol{y} = \mathbf{H}\,\boldsymbol{v}_R + \boldsymbol{\omega} \\ H_2: & \boldsymbol{y} = \mathbf{G}\,\boldsymbol{v}_A + \mathbf{L}\,\boldsymbol{v}_R + \boldsymbol{\omega} \\ H_3: & \boldsymbol{y} = \mathbf{G}\,\boldsymbol{v}_A + \mathbf{L}\,\boldsymbol{v}_0 + \boldsymbol{\omega} \\ H_4: & \boldsymbol{y} = \mathbf{G}\,\boldsymbol{v}_0 + \mathbf{L}\,\boldsymbol{v}_R + \boldsymbol{\omega}, \end{cases} \tag{5.35}$$

where $\mathbf{L} \triangleq \mathbf{H} - \mathbf{G}$.

As previously done to avoid ill-conditioning problem in Section 5.4, we exploit the following equalities applying again the SVD method

$$\mathbf{G}\,\boldsymbol{v}_A = \mathbf{U}_G\,\mathbf{S}_G\,\mathbf{Q}_G^{\mathrm{T}}\,\boldsymbol{v}_A = \widetilde{\mathbf{U}}_G\,\widetilde{\mathbf{S}}_G\,\widetilde{\mathbf{Q}}_G^{\mathrm{T}}\,\boldsymbol{v}_A = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{v}}_A, \tag{5.36}$$

$$\mathbf{L}\,\boldsymbol{v}_R = \mathbf{U}_L\,\mathbf{S}_L\,\mathbf{Q}_L^{\mathrm{T}}\,\boldsymbol{v}_R = \widetilde{\mathbf{U}}_L\,\widetilde{\mathbf{S}}_L\,\widetilde{\mathbf{Q}}_L^{\mathrm{T}}\,\boldsymbol{v}_R = \widetilde{\mathbf{U}}_L\,\widetilde{\boldsymbol{v}}_R. \tag{5.37}$$

Defining $\rho_L = \mathrm{rank}(\mathbf{L})$, then $\widetilde{\mathbf{U}}_L$, of size $4K \times \rho_L$, and $\widetilde{\mathbf{Q}}_L$, of size $4N \times \rho_L$, are, respectively, the matrices of the left-singular vectors and the right-singular vectors of $\mathbf{L}$, whereas $\widetilde{\mathbf{S}}_L$ is a $\rho_L \times \rho_L$ diagonal matrix. The terms

$$\widetilde{\boldsymbol{v}}_A \triangleq \widetilde{\mathbf{S}}_G\,\widetilde{\mathbf{Q}}_G^{\mathrm{T}}\,\boldsymbol{v}_A \quad \text{and} \quad \widetilde{\boldsymbol{v}}_R \triangleq \widetilde{\mathbf{S}}_L\,\widetilde{\mathbf{Q}}_L^{\mathrm{T}}\,\boldsymbol{v}_R \tag{5.38}$$

are the rotated and scaled version of $\boldsymbol{v}_A$ and $\boldsymbol{v}_R$, respectively. Matrices $\widetilde{\mathbf{U}}_G$, $\widetilde{\mathbf{S}}_G$ and $\widetilde{\mathbf{Q}}_G$ have been defined in Section 5.4.

The five hypotheses testing problem (5.35) is therefore redrafted as

$$\begin{cases} H_0: & \boldsymbol{y} = \widetilde{\mathbf{U}}_H\,\widetilde{\boldsymbol{v}}_0 + \boldsymbol{\omega} \\ H_1: & \boldsymbol{y} = \widetilde{\mathbf{U}}_H\,\widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega} \\ H_2: & \boldsymbol{y} = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{v}}_A + \widetilde{\mathbf{U}}_L\,\widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega} \\ H_3: & \boldsymbol{y} = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{v}}_A + \widetilde{\mathbf{U}}_L\,\widetilde{\boldsymbol{v}}_0 + \boldsymbol{\omega} \\ H_4: & \boldsymbol{y} = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{v}}_0 + \widetilde{\mathbf{U}}_L\,\widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega}, \end{cases} \tag{5.39}$$

where $\widetilde{\boldsymbol{v}}_0 \triangleq \widetilde{\mathbf{S}}_H\,\widetilde{\mathbf{Q}}_H^{\mathrm{T}}\,\boldsymbol{v}_0$ under hypothesis H_0, $\widetilde{\boldsymbol{v}}_0 \triangleq \widetilde{\mathbf{S}}_L\,\widetilde{\mathbf{Q}}_L^{\mathrm{T}}\,\boldsymbol{v}_0$ under hypothesis H_3, and, finally, $\widetilde{\boldsymbol{v}}_0 \triangleq \widetilde{\mathbf{S}}_G\,\widetilde{\mathbf{Q}}_G^{\mathrm{T}}\,\boldsymbol{v}_0$ under hypothesis H_4.

Hierarchical GLRT decision rule

The first strategy involves three binary hypothesis test formulations structured on two different layers. In the first layer, the multiple hypothesis

testing problem (5.35) is treated as in Subsection 5.4.1, applying the spoofing detector (5.20) to decide between $\{H_0, H_1\}$ and $\{H_2, H_3, H_4\}$, and the deviation detector (5.23) to decide between $\{H_0, H_3\}$ and $\{H_1, H_2, H_4\}$. The second layer involves a third binary hypothesis test to decide what happens when spoofing and deviation may occur simultaneously,[7] where we need to distinguish between

$$\begin{cases} H_2: & \widetilde{\boldsymbol{v}}_A \neq \boldsymbol{v}_0 \\ H_4: & \widetilde{\boldsymbol{v}}_A = \boldsymbol{v}_0. \end{cases}$$

Consequently, the third additive GLRT, whose derivation is set out in Appendix D.1,

$$L_{\widetilde{\boldsymbol{v}}_A}(\boldsymbol{y}) = \left(\widehat{\widetilde{\boldsymbol{v}}}_A - \widetilde{\boldsymbol{v}}_0\right)^{\mathrm{T}} \mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_A}^{-1} \left(\widehat{\widetilde{\boldsymbol{v}}}_A - \widetilde{\boldsymbol{v}}_0\right) \underset{H_4}{\overset{H_2}{\gtrless}} \tau''' \tag{5.40}$$

is designed to establish if the velocity profile reported by the spoofed AIS data is generic, or instead traces the nominal one. $\widehat{\widetilde{\boldsymbol{v}}}_A \sim \mathcal{N}(\widetilde{\boldsymbol{v}}_A, \mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_A})$ is the MLE of $\widetilde{\boldsymbol{v}}_A$ under hypothesis H_2 (see Appendix D.2). As for the GLRTs (5.20) and (5.23), it is possible to set the false alarm probability and compute the detection performance in closed form as in (4.27). In this case P_{FA} is the probability that the test statistic exceeds τ''' under H_4, and P_{D} is the probability that the test statistic exceeds τ''' under H_2. The number of degrees of freedom characterizing the involved Chi-squared distributions is $d = \operatorname{rank}\left(\mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_A}\right)$, and the non-centrality parameter $\lambda = \lambda(\widetilde{\boldsymbol{v}}_A)$ is given by

$$\lambda\left(\widetilde{\boldsymbol{v}}_A\right) = \left(\widetilde{\boldsymbol{v}}_A - \widetilde{\boldsymbol{v}}_0\right)^{\mathrm{T}} \mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_A}^{-1} \left(\widetilde{\boldsymbol{v}}_A - \widetilde{\boldsymbol{v}}_0\right). \tag{5.41}$$

The three decisions based on GLRTs (5.20), (5.23) and (5.40) are compared to decide about H_i, $i = 0, 1, 2, 3, 4$. Such a decision rule, referred to here as the Hierarchical GLRT (H-GLRT), is defined as

$$\begin{cases} \text{decide } H_0 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) < \tau'' \\ \text{decide } H_1 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) < \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) > \tau'' \\ \text{decide } H_2 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) > \tau'' \text{ and } L_{\widetilde{\boldsymbol{v}}_A}(\boldsymbol{y}) > \tau''' \\ \text{decide } H_3 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) < \tau'' \\ \text{decide } H_4 & \text{if } L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) > \tau' \text{ and } L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) > \tau'' \text{ and } L_{\widetilde{\boldsymbol{v}}_A}(\boldsymbol{y}) < \tau'''. \end{cases} \tag{5.42}$$

Table 5.3 represents the decision made by the H-GLRT on the first layer, where H_i is selected given by the intersection of the two decisions taken by

[7] Only necessary if both (5.20) and (5.23) exceed the respective thresholds.

the GLRTs (5.20) and (5.23). The gray cell identifies the second layer of decision made by GLRT (5.40).

Table 5.3: H-GLRT rule scheme

Decision in (5.20) \ Decision in (5.23)	$\{H_1, H_2, H_4\}$	$\{H_0, H_3\}$
$\{H_2, H_3, H_4\}$	$\{H_2, H_4\}$	H_3
$\{H_0, H_1\}$	H_1	H_0

MOS decision rules

The MOS decides for hypothesis $H_i = \arg\min_j \xi_j$, $j = 0, 1, 2, 3, 4$, where ξ_j is defined as follows

$$\begin{cases} \xi_0 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{H}}\,\widetilde{\boldsymbol{v}}_0\|^2 + \eta_0 \\ \xi_1 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{H}}\,\widehat{\widetilde{\boldsymbol{v}}}_R\|^2 + \eta_1 \\ \xi_2 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{M}}\,\widehat{\widetilde{\boldsymbol{z}}}_M\|^2 + \eta_2 \\ \xi_3 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{G}}\,\widehat{\widetilde{\boldsymbol{v}}}_A - \widetilde{\mathbf{L}}\,\widetilde{\boldsymbol{v}}_0\|^2 + \eta_3 \\ \xi_4 = \|\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{G}}\,\widetilde{\boldsymbol{v}}_0 - \widetilde{\mathbf{L}}\,\widehat{\widetilde{\boldsymbol{v}}}_R\|^2 + \eta_4. \end{cases} \tag{5.43}$$

The MLEs in (5.43), based on the given hypothesis H_i, are

$$\begin{aligned} H_1\colon\quad \widehat{\boldsymbol{v}}_R &= \left(\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{C}}\,\boldsymbol{y} && \text{with } \widetilde{\mathbf{H}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_H, \\ H_2\colon\quad \widehat{\widetilde{\boldsymbol{z}}}_M &= \left(\widetilde{\mathbf{M}}^{\mathrm{T}}\widetilde{\mathbf{M}}\right)^{-1}\widetilde{\mathbf{M}}^{\mathrm{T}}\widetilde{\mathbf{C}}\,\boldsymbol{y} && \text{with } \widetilde{\mathbf{M}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_M, \\ H_3\colon\quad \widehat{\boldsymbol{v}}_A &= \left(\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}}\right)^{-1}\widetilde{\mathbf{G}}^{\mathrm{T}}\left(\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{L}}\widetilde{\boldsymbol{v}}_0\right) && \text{with } \widetilde{\mathbf{G}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_G, \\ H_4\colon\quad \widehat{\boldsymbol{v}}_R &= \left(\widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{L}}\right)^{-1}\widetilde{\mathbf{L}}^{\mathrm{T}}\left(\widetilde{\mathbf{C}}\,\boldsymbol{y} - \widetilde{\mathbf{G}}\widetilde{\boldsymbol{v}}_0\right) && \text{with } \widetilde{\mathbf{L}} = \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_L. \end{aligned}$$

The penalty coefficients are specified in Table 5.4 for each of the five hypothesis and for the three selected MOS criteria.

Table 5.4: MOS penalty coefficients
(Five hypotheses configuration)

	η_0	η_1	η_2	η_3	η_4
NN	0	ρ_H	ρ_A	ρ_G	ρ_M
AIC	0	$2\rho_H$	$2\rho_A$	$2\rho_G$	$2\rho_M$
BIC	0	$\ln(4K)2\rho_H$	$\ln(4K)2\rho_A$	$\ln(4K)2\rho_G$	$\ln(4K)2\rho_M$

5.6 Experimental results

To assess the effectiveness of the detection strategies proposed in Section 5.4, we investigate some case studies concerning an operational scenario. Specifically, we investigate the situation described in Section 5.2 and reported in Fig. 5.1.

The configuration set-up for the synthetic scenario considers a vessel navigating under nominal conditions along an N-section path with $N = 15$, moving according to a piecewise nominal velocity $\boldsymbol{v}_0$ while skirting the forbidden zone, as in Fig. 5.1(a), and plotted in Fig. 5.3 with a green line with diamond markers. The time window covers $T = 12$ hours, and the time intervals are assumed all equal, so that $\Delta_n = T/N$, $\forall n = 1, \dots, N$. Assuming the surveillance system is a radar providing trustful measurements, the vessel is supposed to transmit a number of $K_A = 4$ AIS messages whose reliability is not granted, and temporally positioned at $p_a \in \{5, 7, 9, N\}$, whereas $K_R = 2$ trusted radar measurements are associated with the vessel along its trajectory at $p_r \in \{3, 11\}$.

Under the AIS spoofing hypothesis, a fake trajectory slightly below the nominal one, represented by a blue line with o-markers in Fig. 5.3, is reported by the vessel with a velocities sequence $\boldsymbol{v}_A$, while its real deviation is plotted as a straight red path covered with a velocities sequence $\boldsymbol{v}_R$.

Finally, the reversion rate associated to the underlying OU dynamic model is set as $\gamma_x = \gamma_y = 0.9 \cdot 10^{-2}$ and $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathrm{T}} = \mathrm{diag}(\sigma^2, \sigma^2)$, with noise level $\sigma^2 = 10^{-2}$. The noise covariance matrix is set as $\mathbf{C}_{\boldsymbol{n}_{k_a,k_r}} = \mathrm{diag}(50^2, 50^2, 1, 1)$, $\forall k_a = 1, \dots, K_A$ and $\forall k_r = 1, \dots, K_R$.

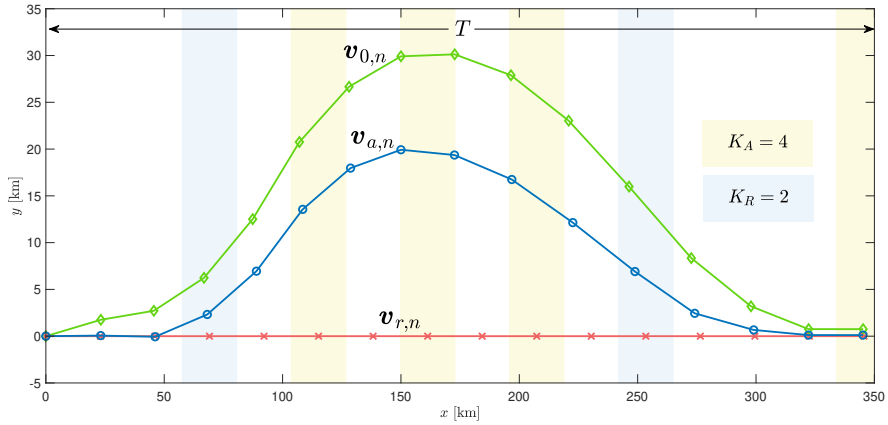


Figure 5.3: Experimental setup: nominal trajectory (green line with diamond markers), deviated trajectory (red line with x-markers), spoofed trajectory (blue line with o-markers); $K_A = 4$ AIS contacts (yellow patches) and $K_R = 2$ radar contacts (light blue patches). Vectors $\boldsymbol{v}_{0,n}$, $\boldsymbol{v}_{a,n}$, and $\boldsymbol{v}_{r,n}$ are the n-th elements of the nominal velocity $\boldsymbol{v}_0$, spoofed velocity $\boldsymbol{v}_A$ and deviating velocity $\boldsymbol{v}_R$, respectively. Each velocity is indicated on the n-th segment of the related trajectory.

5.6.1 Performance of the spoofing detector (5.20) and deviation detector (5.23)

The performance of the detectors (5.20) and (5.23) is reported in Fig. 5.5 in terms of false alarm probability P_{FA} versus missed detection probability $1 - P_{\text{D}}$ and represented through the ROC curves considering three different analyses.

In the first analysis, three different spoofed trajectories (STs) are considered with a gap of about 5 km at the largest distance from each other and from the nominal path, as shown in Fig. 5.4 (blue lines with o-markers and three different line styles). Figure 5.5(a) shows a clear performance improvement as the spoofed trajectory moves away from the nominal one in terms of spoofing detection. Figure 5.5(b) reports the ROC curve related to the deviation detector (5.23), which is clearly the same one for each of the three spoofed trajectories in Fig. 5.5(a).

In the second analysis, the performance of the detectors (5.20) and (5.23) is analyzed for three different configurations of the AIS and radar contacts:

1. $K_A = 2$ AIS contacts placed in $p_a \in \{7, N\}$, and $K_R = 2$ radar contacts

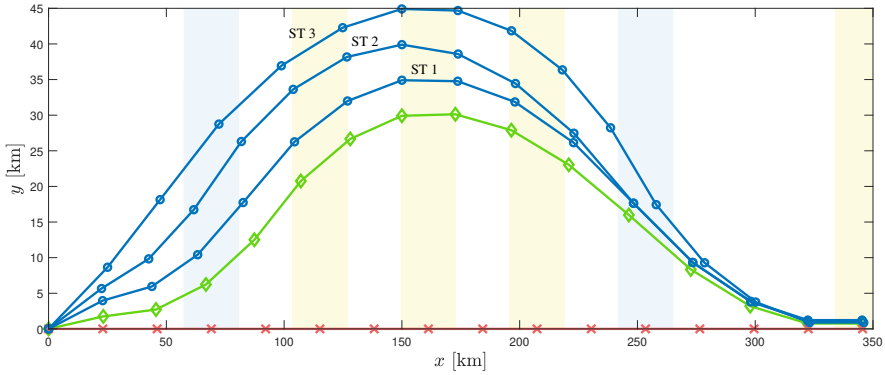


Figure 5.4: Three different spoofed trajectories (STs) (blue lines with o-markers) are considered with a gap of about 5 km at the highest distance from each other and from the nominal path (green line with diamond markers). The red curve is the deviating trajectory (red line with x-markers).

placed in $p_r \in \{2, 14\}$;

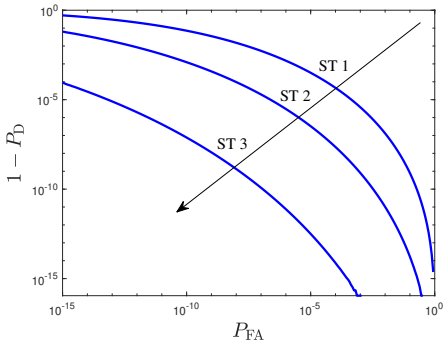
2. $K_A = 4$ AIS contacts placed in $p_a \in \{4, 7, 10, N\}$, and $K_R = 3$ radar contacts placed in $p_r \in \{3, 5, 12\}$;

3. $K_A = 5$ AIS contacts placed in $p_a \in \{4, 6, 8, 10, N\}$, and $K_R = 3$ radar contacts still placed in $p_r \in \{3, 5, 12\}$.

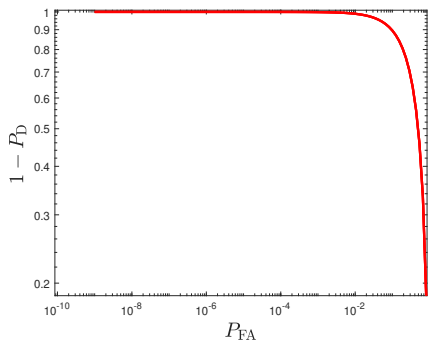
Figure 5.5(c) shows the improving performance of test (5.20) when the total number K of contacts increases, which makes the spoofing much more detectable.

In Fig. 5.5(d) there is only the ROC curve related to detector (5.23) for the first configuration of contacts because of the very high values of $\lambda(\widetilde{\boldsymbol{v}}_R)$ associated to the the increasing number of contacts in the other two configurations, which leads to numerical errors in reporting the ROC for very small values of the error probabilities. For this reason, a table of these values has been reported for completeness in Fig. 5.5(d), and also compared to the $\lambda(\widetilde{\boldsymbol{\varphi}})$ values.

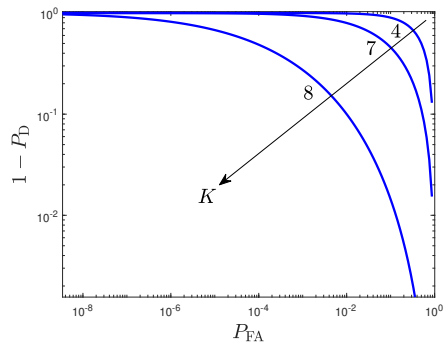
The proposed detectors are also tested against varying the time window. Specifically, T is let taking values within the set of $\{10, 15, 20\}$ hours. In Fig. 5.5(e) the missed detection probability decreases with the time window T, underlining that the longer the anomalous behavior, the easier it is to detect the AIS spoofing. The performance of the detector (5.23) shows a



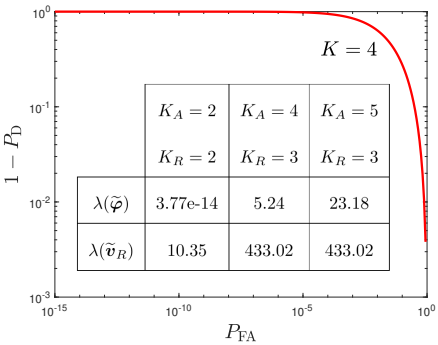
(a) Detector (5.20) performance for three different spoofed trajectories (STs).



(b) Detector (5.23) performance for the same parameter configuration in Fig. 5.5(a).

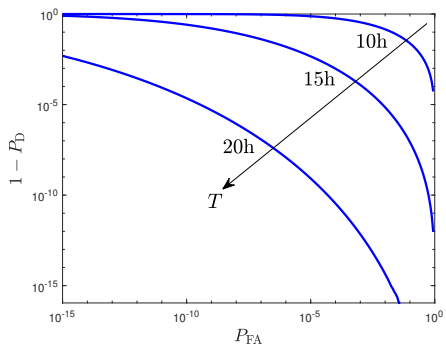


(c) Detector (5.20) performance for three different numbers and combinations of contacts K.

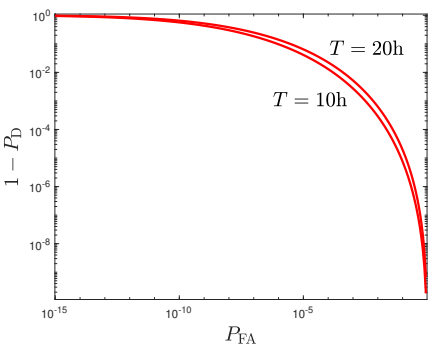


	$K_A = 2$ $K_R = 2$	$K_A = 4$ $K_R = 3$	$K_A = 5$ $K_R = 3$
$\lambda(\widetilde{\boldsymbol{\varphi}})$	3.77e-14	5.24	23.18
$\lambda(\widetilde{\boldsymbol{v}}_R)$	10.35	433.02	433.02

(d) Detector (5.23) performance for the same parameter configuration in Fig. 5.5(c).



(e) Detector (5.20) performance for three different time windows T.



(f) Detector (5.23) performance for the same parameter configuration in Fig. 5.5(e).

Figure 5.5

different behavior, i.e., $\lambda(\widetilde{\boldsymbol{v}}_R)$ is not always increasing, reaching the values of 59.0798, 411.9909, 55.5850 for $T = 10, 15$, and 20 hours, respectively. A performance improvement is observable from $T = 10$ h to $T = 15$ h, but the performance for $T = 20$ h is worse than both $T = 10$ h and $T = 15$ h. The ROC curves for $T = 10$ and 20 h are reported in Fig. 5.5(f), where the case with $T = 15$ h cannot be represented in terms of ROC since $\lambda_{\widetilde{\boldsymbol{v}}}$ is too high.

5.6.2 Comparison of MOS and C-GLRT methodologies

We will compare now the performance in terms of probability of correct decision of the C-GLRT proposed in Subsection 5.4.1 and the MOS rules NN, AIC, and BIC in Subsection 5.4.2.

We consider the scenario reported in Fig. 5.3, described at the beginning of Section 5.6. The experiment is executed for the same set of parameters used in Subsection 5.4.1. Specifically, Fig. 5.6(a)-(c) refer to three different spoofed trajectories depicted in Fig. 5.4.

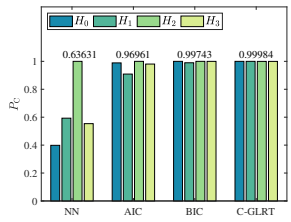
The Figures 5.6(d)-(f) refer respectively to three different configurations of the AIS and radar contacts:

1. in Fig. 5.6(d) $K_A = 2$ AIS contacts are placed in $p_a \in \{7, N\}$, and $K_R = 2$ radar contacts are placed in $p_r \in \{2, 14\}$;

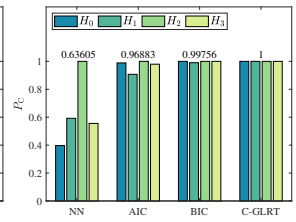
2. in Fig. 5.6(e) $K_A = 4$ AIS contacts are placed in $p_a \in \{4, 7, 10, N\}$, and $K_R = 3$ radar contacts are placed in $p_r \in \{3, 5, 12\}$;

3. in Fig. 5.6(f) $K_A = 5$ AIS contacts are placed in $p_a \in \{4, 6, 8, 10, N\}$, and $K_R = 3$ radar contacts are placed in $p_r \in \{3, 5, 12\}$.

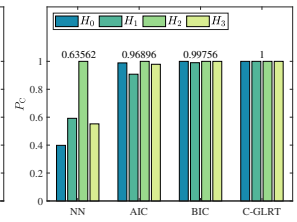
The Figures 5.6(g)-(i) refer to $T = 10, 15, 20$ hours, respectively. The probability of correct decision is computed empirically for each hypothesis being true via Monte Carlo simulation, with 10^4 trials. Clearly, the higher the probability of correct decision, the better the detector performance will be. In order to make a meaningful comparison, we can average the probability of correct decision for each hypothesis. This is equivalent to minimizing the error probability in a Bayesian fashion when the prior is uniform over the hypotheses. The averaged values are reported on the histograms for each rule in all panels in Fig. 5.6. We observe that the NN rule is outperformed in most of the scenarios. In Fig. 5.6(a)-(c) the AIC, BIC and C-GLRT seem to be equally good, being all above 0.95 in terms of probability of correct



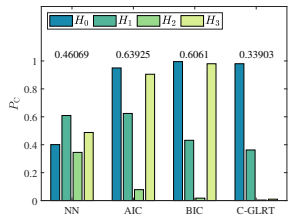
(a) Spoofed trajectory 1.



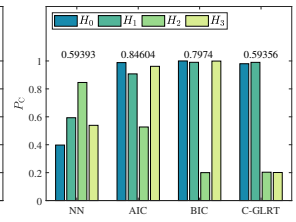
(b) Spoofed trajectory 2.



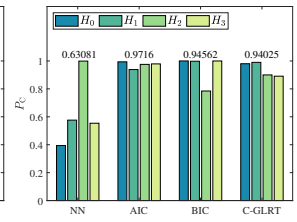
(c) Spoofed trajectory 3.



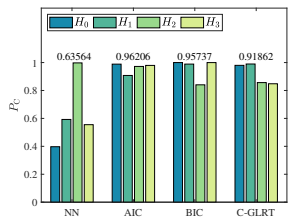
(d) $K_A = 2$, $K_R = 2$.



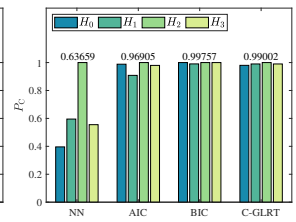
(e) $K_A = 4$, $K_R = 3$.



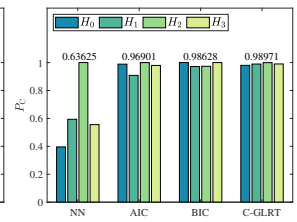
(f) $K_A = 5$, $K_R = 3$.



(g) $T = 10$ hours.



(h) $T = 15$ hours.



(i) $T = 20$ hours.

Figure 5.6: Results of the Monte Carlo simulation to compute the probability of correct decision P_C for each detector under each one of the four hypotheses. Panels (a)-(b)-(c) report performance of the three different spoofed trajectories of Fig. 5.4. Panels (d)-(e)-(f) report performance for three different configurations of AIS and radar measurements. Panels (g)-(h)-(i) report performance respectively for $T = 10, 15, 20$ hours.

decision. However, there are differences, and the AIC is slightly worse than the BIC and C-GLRT, with C-GLRT being the best detector.

In Fig. 5.6(a)-(c), we observe that MOS rules do not change significantly their performance when the spoofed trajectory is changing from (a) to (c). Our conjecture is that the MOS rules are somehow collocated in an analogous region of high P_{FA} ($> 10^{-4}$) in Fig. 5.7, where there is basically no difference when the spoofed trajectory is changing.

In Fig. 5.6(d)-(f), the detection problem is more challenging, especially when fewer data are available as in Fig. 5.6(d). However, when more observations are available, increasing both K_A and K_R as shown in Fig. 5.6(e) and (f), performance sensibly increases for all the decision strategies. In these settings, the C-GLRT is outperformed by both the AIC and BIC. Contrary to the previous cases reported in Fig. 5.6(a)-(c), here the AIC is better than the BIC. In this case, we do not have a good strategy to set both the P_{FA} of the detectors (5.20) and (5.23) of the C-GLRT, and a nominal value of $P_{\text{FA}} = 0.01$ has been used.

In Fig. 5.6(g) we observe again that the AIC and BIC are better than the C-GLRT; in Fig. 5.6(h), the C-GLRT outperforms the AIC but not the BIC, whereas, in Fig. 5.6(i), the C-GLRT outperforms both the AIC and BIC. As observed in Subsection 5.6.1, we notice that, extending the time window, the performance tends to increase; this is evident from $T = 10$ hours to $T = 15$ hours, but much less evident from $T = 15$ hours to $T = 20$ hours. Indeed, as already discussed in the scenarios reported in Fig. 5.5(e) and (f), whereas $\lambda(\widetilde{\boldsymbol{\varphi}})$ always increases with T, reaching the values of 18.7314, 59.4102, and 124.8558 for $T = 10, 15$ and 20 hours, respectively, $\lambda(\widetilde{\boldsymbol{v}}_R)$ is not always increasing, reaching the values of 59.0798, 411.9909 and 55.5850 for $T = 10, 15$ and 20 hours, respectively. This can explain why the performance is not improving from $T = 15$ hours to $T = 20$ hours for all the detectors under consideration.

Remark 5.6.1. *As anticipated in Subsection 5.4.1, it is possible to optimally select the thresholds τ' and τ'', in order to maximize the probability of correct decision P_{C}.*

Indeed, in the specific case concerning the varying deviation of the spoofed trajectory, since $\lambda(\widetilde{\boldsymbol{v}}_R) \approx 379$, the decision about the abnormal deviation (5.23) can be made almost perfectly (hence we have selected a false alarm probability $P_{\text{FA}} = 10^{-10}$ corresponding to $1 - P_{\text{D}} \approx 10^{-10}$). Consequently, the C-GLRT performance can be optimized assuming that the correct decision

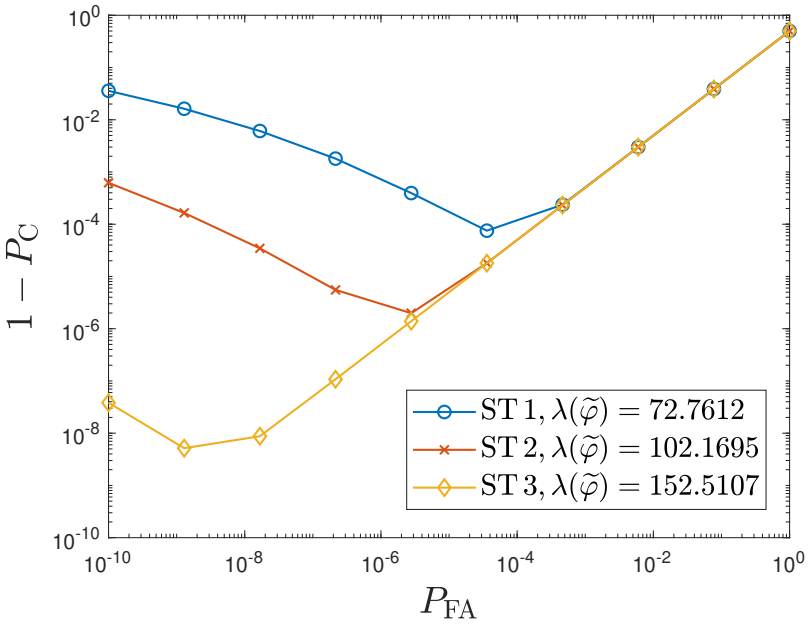


Figure 5.7: Approximate error decision probability (5.28) assuming a perfect detection of the anomalous deviation from the nominal route plotted for each of the three spoofed trajectories considered in the first analysis of Subsection 5.6.1 and in Fig. 5.6(a)-(c).

event is made only based on the spoofing decision in (5.20) *or, in other words, the average probability of correct decision is only function of* τ'*, see equation* (5.28).

In Fig. 5.7 we reported the error probability $1 - P_{\mathrm{C}}$ (5.28) *versus* P_{FA} *for each value of* $\lambda(\widetilde{\boldsymbol{\varphi}}) = 72, 102, 152$ *corresponding to the three spoofed trajectories. It is interesting to note that for larger values of* $P_{\mathrm{FA}}(> 10^{-4})$ *there are no differences among the three spoofed trajectories in Fig. 5.4, in other words* $1 - P_{\mathrm{C}} \approx P_{\mathrm{FA}}$ *since* $P_{\mathrm{D}} \approx 1$*. However, decreasing* P_{FA} *we observe a minimum value, or, in other words, an optimal setting for* P_{FA}*, which would lead to the minimum error probability, or the maximum probability of correct decision. With regard to the C-GLRT, we get* $P_{\mathrm{C}} = 0.99984$ *for the first spoofed trajectory, while for the others the empirical probability is practically equal to one since the theoretical error reported in Fig. 5.6 would be below the inverse of the Monte Carlo runs* (10^4)*, specifically* $2 \cdot 10^{-6}$ *and* $5 \cdot 10^{-9}$

for the second and the third spoofed trajectory, respectively.

5.7 Conclusion

In this chapter it has been documented a means to detect whether a vessel is reporting spoofed AIS data and deviations from the planned route within a multiple hypothesis testing formulation. Two different anomaly detection strategies are developed based on the changes in the OU process long-run mean velocity parameter and supported by the information from reliable radar measurements. In particular, the first strategy aims at determining whether the trajectory reported by the vessel via the AIS is trustworthy or not by means of a GLRT approach, and, similarly, a further GLRT is exploited in order to detect a possible deviation from the nominal route relying on the velocity parameter. Eventually, the resulting decisions from the two GLRTs are then sub-optimally combined in a C-GLRT.

The second strategy is designed to tackle jointly both the spoofing and the deviation events from the nominal route, resorting to the Model Order Selection rules, such as the direct KL approach (also known as NN rule), the AIC, and the BIC.

The detection strategies have been tested within an operational scenario concerning some varying parameters, such as the spoofed trajectory, the number and position of both AIS and radar contacts, and the time window.

It is observed that the NN rule is always outperformed by the others, while there is not a rule among AIC, BIC and C-GLRT that is always better than the others. However, all the detection strategies are able to effectively achieve good performance in terms of probability of correct decision in most of the scenarios analyzed.

Chapter 6

Optimal opponent stealth trajectory planning

In this chapter we play the opponent's side, and describe the least-detectable trajectory that an elusive vessel can follow. The opponent's route planning problem is formalized as a non-convex optimization problem, and physical and practical requirements are accounted for by enforcing several constraints at the optimization design stage. To handle the resulting non-convex optimization problem, we propose a globally-optimal and computationally-efficient technique, called Non-Convex Optimized Stealth Trajectory (N-COST) algorithm, which amounts to solve a number of convex problems. The effectiveness of the proposed approach is demonstrated through some case studies and a real-world example. [1]

6.1 Introduction

With reference to the detection strategies proposed in Chapters 4 and 5, this chapter moves from the point of view of the detector to that of a malicious and elusive vessel. In the specific, we propose a new methodology to design the *optimal* stealth trajectory of a vessel that intends to deviate from

[1] This chapter has been published as "Optimal Opponent Stealth Trajectory Planning Based on an Efficient Optimization Technique" in *IEEE Transactions on Signal Processing*, vol. 69, pp. 270-283, 2020 [24].

the nominal traffic route to accomplish some specific tasks, e.g., to reach a location point where it can rendezvous with another vessel. To hide this behavior, the vessel turns its AIS transponder OFF for the time required to arrive at the specific point, stops and returns to the expected route. In addition, we assume that, during the period of the AIS disablement, the vessel can only be observed (and that, very seldom) by space-borne sensors, such as SAR, whose sequence of sensor measurement time instants is known to the vessel.

The proposed route planning problem aims at defining the optimal trajectory of the opponent vessel and is formalized as a constrained optimization problem, where physical and practical requirements are accounted for at the design stage. The resulting non-convex optimization problem is handled by means of an efficient algorithm that determines the sequence of positions and velocities that the opponent vessel has to follow to result as stealthy as possible to the surveillance system. The effectiveness of the proposed opponent route planning strategy is then tested against the anomaly detector built in Chapter 4 with applications to some relevant case studies. The ultimate goal of this chapter is to highlight the limitations of the anomaly detection strategy presented in Chapter 4 by identifying the worst-case scenario in terms of surveillance capability. This would facilitate possible future work focusing on the available countermeasures (e.g., optimal asset allocation) to improve the detection capability, and perhaps thereby framing a game-theoretic analysis and strategy.

The main contribution of this chapter is twofold, specifically:

- a novel optimal path planning formulation to blind an anomaly detection procedure so as to execute covertly a given task (e.g., a rendezvous with another ship), and

- an optimal and efficient solution for the aforementioned problem formulation, called Non-Convex Optimized Stealth Trajectory (N-COST).

Without the (non-trivial – see Theorem 6.3.4) derivation of the N-COST, the proposed optimization problem would not have been solvable. A *brute force* implementation procedure would not be feasible, since the optimization space would grow exponentially with the length of the velocity and position sequences for any given discretization step of the optimization variables.

6.2 Maritime anomalous scenario and deviations detector's standpoint

We take into consideration the situation already depicted in Section 4.2, where a vessel is following a nominal route with a nominal velocity, and, at some point, the AIS data streaming is interrupted for a period of duration T. Excluding the event of AIS data spoofing in this chapter, as in Section 4.2, and considering the set of the time instants $t_0 < \ldots < t_n < \ldots < t_N$, the detection system can envision two hypotheses for the behavior of the ship during this silent period:

1. Hypothesis H_0: the AIS interruption was accidental or an obscuration, and the vessel keeps on following the planned route at the nominal velocity. The detector assumes that the vessel moves according to a piecewise OU model, with a sequence of N long-run mean velocity values, stored in a matrix $\boldsymbol{V}_0 = [\boldsymbol{v}_{0,1}, \ldots, \boldsymbol{v}_{0,n}, \ldots, \boldsymbol{v}_{0,N}] \in \mathbb{R}^{2,N}$, which identifies the nominal N-section path;

2. Hypothesis H_1: the vessel deviates from the nominal route, and is intentionally keeping its AIS transponder switched OFF to hide the deviation. The detector assumes that the vessel moves according to a piecewise OU model, with a sequence of unknown N long-run mean velocity values, stored in a matrix $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \ldots, \boldsymbol{v}_N] \in \mathbb{R}^{2,N}$, which identifies an alternative N-section path.

The nominal vessel's positions along the N-section path are represented by the random variables $\boldsymbol{x}_n$, $n = 0, \ldots, N$. The time intervals are assumed to be all equal, i.e., $\Delta t \triangleq t_n - t_{n-1}$, $\forall n = 1, \ldots, N$, and sufficiently large with respect to the OU parameters, in order to guarantee that the process reaches a steady state behavior in each time interval.

The detector disposes of $K+1$ sporadic measurements (AIS, radar, SAR, etc.) during the period between t_0 and $t_0 + T$, where the k-th measurement is given by (4.13). Hence, as in Chapter 4, the anomaly detection problem is abstracted to the Gaussian composite hypothesis testing problem (4.23). Assuming known the number of sections N and the sequence of the time intervals $\mathcal{D}_n$, and $\mathbf{H}$ a full rank model matrix given in (4.22), the detector considered in this chapter has the form (4.30), and reported below for

simplicity:

$$\left(\widehat{\boldsymbol{v}} - \mathrm{vec}(\boldsymbol{V}_0)\right)^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H} \left(\widehat{\boldsymbol{v}} - \mathrm{vec}(\boldsymbol{V}_0)\right) \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{6.1}$$

where $\widehat{\boldsymbol{v}}$ is the MLE of $\boldsymbol{v} = \mathrm{vec}(\boldsymbol{V})$, given by (4.31). The detection performance is thus given by (4.27) with $d = 2N$ degrees of freedom, corresponding to the size of the known parameter $\boldsymbol{v}_0 = \mathrm{vec}(\boldsymbol{V}_0)$, and non-centrality parameter

$$\lambda\left(\boldsymbol{V}\right) = \mathrm{vec}\left(\boldsymbol{V} - \boldsymbol{V}_0\right)^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H}\, \mathrm{vec}\left(\boldsymbol{V} - \boldsymbol{V}_0\right). \tag{6.2}$$

6.3 Opponent stealth route planning: problem formulation & solution

The goal of the opponent is to minimize the detection capability, quantified by the anomaly distance $\lambda\left(\boldsymbol{V}\right)$ (6.2), as it will be discussed in Subsection 6.3.2, while accomplishing a specific mission. In this section, we formalize a route planning strategy of the opponent vessel to optimally determine its trajectory; we also propose an efficient technique to handle the resulting non-convex optimization problem.

As a first step towards a meaningful problem definition, let us establish the main tasks and objectives to account for at the design stage:

- the opponent vessel has to meet and fulfill its mission, i.e., to reach a specific point, indicated with $\boldsymbol{p}_c = [p_{c_x},\ p_{c_y}]^{\mathrm{T}}$, unknown to the detector, where, for instance, a possible illicit ship-to-ship transfer would occur;
- the kinematics of the opponent vessel have to satisfy physical constraints such as the maximum speed and acceleration, the presence of land or other forbidden navigation regions, etc.; and
- the anomalous trajectory has to be as hidden as possible, namely the detection capability has to be minimized.

As it will be detailed in the following sections, the ability to design a stealth trajectory improves as T increases. In the proposed approach, T is a design parameter, which can be easily included in the optimization formulation at the expense of increasing the computational cost. However, the value

of T cannot be chosen arbitrarily large, since, if it were significantly greater than typical AIS data gaps in the area, other anomaly detection systems and modalities would make our OU-based analysis moot.

In principle, we could optimize only the stack of long-run mean velocity parameters $\boldsymbol{V}$ of the underlying OU process under H_1. However, by doing so, the positions would not be controlled. Consequently, it would be regulated by an IOU process, and the position error would increase linearly over time as $n\Delta t$, hence there would be no assurance of successful rendezvous. To avoid this, we propose to control both the velocity and position sequences, as it is common for ships equipped with navigation control systems. Thus, the available degrees of freedom for the route planner are given by:

- the N long-run mean velocity vectors stacked in the $\boldsymbol{V}$ matrix, where the n-th vector is the OU mean velocity during the time interval $[(n-1)\Delta t,\ n\Delta t]$;

- the N vessel's positions stacked in the matrix $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n, \ldots \boldsymbol{p}_N] \in \mathbb{R}^{2,N}$, where the n-th vector is the location, at the time instant $n\Delta t$, while the initial position $\boldsymbol{p}_0$ is assumed known and fixed.

Remark 6.3.1. *The optimal control of both velocity and position returns a deterministic trajectory as output, forced to evolve according to the realization of an OU process.*

For its part, the detector (6.1) assumes the vessel is moving according to an OU process and attempts to rebuild the proper trajectory under hypothesis H_1 (estimating the unknown velocity sequence), among all the possible trajectories, based on the available observations. Then, the use of detector (6.1), when the optimizer is controlling both position and velocity, would be based on the assumption that, under H_1, the observed trajectory is stochastic, instead of deterministic. It is worth mentioning that this approximation is also valid in light of real-world data analysis, in which the OU model has been corroborated by a great deal of empirical data [154].

6.3.1 Route constraints

C_1) *Speed and acceleration constraint*: Neither the OU long-run mean velocity parameter, nor the inter-segment acceleration, can be arbitrarily large;

consequently, the following constraints are included

$$\|\boldsymbol{v}_h\| \leq v_{\max}, \qquad h = 1, \ldots, N, \tag{6.3}$$

$$\|\boldsymbol{v}_{h'} - \boldsymbol{v}_{h'-1}\| \leq \Delta v_{\max}, \qquad h' = 2, \ldots, N, \tag{6.4}$$

where $v_{\max}$ is the maximum mean speed allowed to the ship and $\Delta v_{\max} = a_{\max}\Delta t$ the maximum absolute acceleration in successive periods, with $a_{\max}$ denoting the acceleration.

C_2) *OU model compliance constraint*: This forces the trajectory to be *compliant* with the underlying OU statistical model. To this end, at the planning stage, it is required that

$$\left\|\boldsymbol{p}_h - \left(\boldsymbol{v}_h\Delta t + \boldsymbol{p}_{h-1}\right)\right\|^2_{\mathbf{C}_p^{-1}} \leq \epsilon, \; h = 1, \ldots, N, \tag{6.5}$$

where $\left(\boldsymbol{v}_h\Delta t + \boldsymbol{p}_{h-1}\right)$ and $\mathbf{C}_p$ are the expected value and the position covariance matrix of $\boldsymbol{x}_h$, given $\boldsymbol{x}_{h-1}$ (where $\boldsymbol{x}_h$ is defined in Section 6.2), respectively, neglecting the transitory terms.[2] The parameter ϵ rules the proximity between the expected and the actual trajectory, which can be formally described by the set:

$$\mathcal{X}_{h,\epsilon} = \left\{\boldsymbol{x} \in \mathbb{R}^2 : \left\|\boldsymbol{x} - (\boldsymbol{v}_h\Delta t + \boldsymbol{p}_{h-1})\right\|^2_{\mathbf{C}_p^{-1}} \leq \epsilon\right\}.$$

In other words, (6.5) forces the trajectory points to lie within the IOU model's uncertainty region at given percentage α_ϵ, as:

$$\mathrm{P}\left(\boldsymbol{x}_h \in \mathcal{X}_{h,\epsilon} | \boldsymbol{x}_{h-1} = \boldsymbol{p}_{h-1}\right) = \alpha_\epsilon.$$

In this way, the optimized trajectory is statistically compliant with a trajectory generated from an OU process. This compliance condition, imposed to the optimized trajectory, is somehow close to the *typical* property of the sequences generated by random processes [63]. Thanks to this condition, the optimized trajectory will be statistically *typical*, or *compliant*, with OU-generated trajectories.

C_3) *Land constraint*: As an additional requirement, the sought route is also forced to reside within a specific portion of the sea surface, to account for the possible presence of land in the region. More precisely,

$$\boldsymbol{p}_h \in \mathcal{S}_c, \quad h = 1, \ldots, N, \tag{6.6}$$

[2]Note that this condition can be easily generalized to the case of a different stochastic process using a different covariance matrix.

where $\mathcal{S}_c \subseteq \mathbb{R}^2$ is a subset of the plane, that is assumed convex for mathematical tractability. For instance, $\mathcal{S}_c$ can be described by means of polyhedra, ellipsoids, as well as their intersections with appropriate regions, to account for the shape of the coast.

C_4) *Rendezvous constraint*: As already highlighted, the vessel's main goal is to reach and transact some business at the specific point $\boldsymbol{p}_c$. From a mathematical point of view, this can be modeled with a constraint on the vessel's path to be close to the point $\boldsymbol{p}_c$ for some, say L, time instants. Otherwise stated, it is required that there exists a time instant $h\Delta t$, $1 \leq h \leq N - L + 1$, such that

$$\max_{k=h,\ldots,h+L-1} (\|\boldsymbol{p}_k - \boldsymbol{p}_c\|) \leq \delta, \tag{6.7}$$

with δ defining the stopping area size around $\boldsymbol{p}_c$ where the ship will lie for a time of at least $L\Delta t$, assumed known.

C_5) *Starting point condition*: Finally, let us observe that the AIS transmitter is turned OFF at time instant $t_0 = 0$, thus the position of the vessel at this instant, i.e., $\boldsymbol{p}_0$, does not represent a degree of freedom, but it is constrained to be at the communicated location. Without loss of generality, in the following, it is assumed that such position coincides with the origin of the reference system, i.e., $\boldsymbol{p}_0 = \boldsymbol{0}$.

6.3.2 Figure of merit for surveillance system covertness

In order to control the surveillance capabilities of the sensing system, it is worth observing that the detection performance depends on the available measurements. Furthermore, as also seen in Chapters 4 and 5, the detection probability in (4.27) is mainly influenced by the non-centrality parameter in (6.2); that is, for a given probability of false alarm and detector structure (i.e., set of measurements), the higher the non-centrality parameter is, the better the detection performance results.

The opponent vessel is not necessarily aware of the time instants when its kinematics will be monitored by the surveillance assets. However, the overall set of possible acquisition instants from which these times could be chosen can be assumed known to the opponent vessel. From the perspective of the opponent vessel, one possible strategy could be to minimize the worst-case risk of detection. Hence, we consider $\lambda(\boldsymbol{V})$ in (6.2) as performance metric. To shed light on such a choice and further corroborate (6.2)

as a meaningful/useful figure of merit, let us remind that $\lambda(\boldsymbol{V})$ is twice the KL-divergence between the actual distributions under the H_0 and H_1 hypotheses. As a consequence, since the KL-divergence controls the performance limits for any given hypothesis testing problem [63], the non-centrality parameter represents a viable means to predict the surveillance capability of the sensing system. It is also worth noting that the detection performance of the clairvoyant receiver, obtained assuming known the ship velocity profile under the two hypotheses, is ruled by $\lambda(\boldsymbol{V})$, too. As a result, the upper bound to the surveillance system performance can be expressed in terms of $\lambda(\boldsymbol{V})$. In conclusion, $\lambda(\boldsymbol{V})$ provides a meaningful performance measure of the detectability of the anomaly.

The highest non-centrality parameter is obtained when the surveillance system uses all the possible measurements. This is a direct consequence of information-theoretic considerations, and the chain rule for KL-divergence [63], which implies the monotonicity of $\lambda(\boldsymbol{V})$ (with respect to the inclusion of the measurement). Hence, hereafter $\lambda(\boldsymbol{V})$ in (6.2) directly refers to the non-centrality parameter associated to the maximum number of measurements.

Remark 6.3.2. *It is assumed that the opponent vessel has knowledge of all the possible measurement time instants. This assumption is often realistic, since some satellite orbits are publicly known. On the other hand, when such surveillance comes from patrolling assets, it cannot be reasonably assumed that the malicious vessel knows their trajectories in advance. However, their perfect knowledge can still be assumed to consider the worst-case condition for the anomaly detector. Summarizing, the opponent optimizes its trajectory based on the worst set of observations in terms of $\lambda(\boldsymbol{V})$, but with a perfect knowledge of this figure of merit.*

To summarize, the idea is that the stealthy vessel has turned its AIS OFF for a time T in order to linger at a rendezvous location $\boldsymbol{p}_c$. To be conservative, the vessel assumes all times are used by the detector, although the latter may (or may not) observe the positions of the vessel at all possible available times of measurements acquirement. The detector does not check the absolute positions of the measurements versus the expected ones, since to do so would make little sense: the only information the detector has is in terms of the nominal OU operation of the vessel, meaning $\boldsymbol{V}_0$. Hence, the detector (6.1) checks the best MLE of $\boldsymbol{V}$ against $\boldsymbol{V}_0$ and, essentially, asks: does this observed OU process make sense? The performance of such detection is encapsulated in the figure of merit $\lambda(\boldsymbol{V})$, and the stealthy vessel's

goal is to achieve $\boldsymbol{p}_c$ while minimizing $\lambda(\boldsymbol{V})$.

6.3.3 Route design problem and solution technique

According to guidelines defined earlier and constraints C_1–C_5 in Subsection 6.3.1, the route planning problem, which computes the ship positions $\boldsymbol{P}$, as well as the OU long-run mean velocity values $\boldsymbol{V}$, can be formulated as the following constrained optimization problem

$$\mathcal{P}\begin{cases} \min\limits_{\boldsymbol{V},\boldsymbol{P}} & \lambda(\boldsymbol{V}) \\ \text{s.t.} & \|\boldsymbol{v}_h\| \leq v_{\max} \\ & \|\boldsymbol{v}_{h'} - \boldsymbol{v}_{h'-1}\| \leq \Delta v_{\max} \\ & \|\boldsymbol{p}_h - \left(\boldsymbol{v}_h \Delta t + \boldsymbol{p}_{h-1}\right)\|^2_{\mathbf{C}_p^{-1}} \leq \epsilon \\ & \boldsymbol{p}_h \in \mathcal{S}_c \\ & h = 1, \ldots, N, \quad h' = 2, \ldots, N \\ & \min\limits_{l=1,\ldots,N-L+1} \max\limits_{k=l,\ldots,l+L-1} (\|\boldsymbol{p}_k - \boldsymbol{p}_c\|) \leq \delta \\ & \boldsymbol{p}_0 = \mathbf{0} \\ & \boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \ldots, \boldsymbol{v}_N] \in \mathbb{R}^{2,N} \\ & \boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n, \ldots, \boldsymbol{p}_N] \in \mathbb{R}^{2,N} \end{cases} \tag{6.8}$$

where: the first two constraints account for the vessel kinematic feasibility (C_1); the third constraint guarantees the compliance of the trajectory with the underlying statistical model (C_2); the fourth constraint imposes land avoidance (C_3); the fifth constraint ensures the existence of a time interval of at least L consecutive time samples, to allow the illicit activity at the rendezvous (C_4); the last requirement defines the starting point of the vessel trajectory (C_5).

Remark 6.3.3. *Different design problems can be formulated and solved varying both Δt and N. If we fix $T = N\Delta t$, intuitively, the decrease of Δt and the increase of N should lead to better achievable performance. Indeed, the optimal solution to (6.8) is still available by doubling N but a better solution can be found in this case. Analysis is given in Appendix E.1.*

In order to proceed with the solution of $\mathcal{P}$, let us observe that the problem $\mathcal{P}$ is non-convex, due to the C_4 constraint. However, the following theorem paves the way for an efficient globally-optimal solution technique for the formulated route planning problem.

Let $\mathcal{P}_l$, $l \in \mathcal{A} = \{1, \ldots, N - L + 1\}$, be the optimization problem

$$\mathcal{P}_l \begin{cases} \min\limits_{\boldsymbol{V},\boldsymbol{P}} & \lambda(\boldsymbol{V}) \\ \text{s.t.} & \|\boldsymbol{v}_h\| \leq v_{\max} \\ & \|\boldsymbol{v}_{h'} - \boldsymbol{v}_{h'-1}\| \leq \Delta v_{\max} \\ & \|\boldsymbol{p}_1 - \boldsymbol{v}_1 \Delta t\|_{\mathbf{C}_p^{-1}} \leq \sqrt{\epsilon} \\ & \left\|\boldsymbol{p}_{h'} - (\boldsymbol{v}_{h'} \Delta t + \boldsymbol{p}_{h'-1})\right\|_{\mathbf{C}_p^{-1}} \leq \sqrt{\epsilon} \\ & \boldsymbol{p}_h \in \mathcal{S}_c, \ h = 1, \ldots, N \\ & h = 1, \ldots, N, \quad h' = 2, \ldots, N \\ & \max\limits_{k=l,\ldots,l+L-1} (\|\boldsymbol{p}_k - \boldsymbol{p}_c\|) \leq \delta \\ & \boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \ldots, \boldsymbol{v}_N] \in \mathbb{R}^{2,N} \\ & \boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n, \ldots, \boldsymbol{p}_N] \in \mathbb{R}^{2,N} \end{cases} \tag{6.9}$$

where the only difference with problem $\mathcal{P}$ is in the rendezvous constraint: in $\mathcal{P}_l$ the rendezvous time index l is fixed; then, for any $l \in \mathcal{A}$, $\mathcal{P}_l$ is a convex optimization problem.

Theorem 6.3.4. *Let us consider the problems $\mathcal{P}$ and $\mathcal{P}_l$, $l \in \mathcal{A}$.*

- *The problem $\mathcal{P}$ is feasible if and only if at least one among the problems $\mathcal{P}_l$, $l \in \mathcal{A}$, is feasible.*
- *Assume the existence of some $l \in \mathcal{A}$ such that $\mathcal{P}_l$ is feasible, and $(\boldsymbol{V}^l, \boldsymbol{P}^l)$ is the optimal solution to $\mathcal{P}_l$.*

Then, an optimal solution to $\mathcal{P}$ is $(\boldsymbol{V}^\star, \boldsymbol{P}^\star)$, with

$$\boldsymbol{V}^\star = \boldsymbol{V}^{l^\star}, \qquad \boldsymbol{P}^\star = \boldsymbol{P}^{l^\star}, \tag{6.10}$$

where, denoting with $v(\mathcal{P}_l)$ the optimal value of Problem $\mathcal{P}_l$,

$$l^\star = \arg\min_{l \in \mathcal{A}} v(\mathcal{P}_l).$$

Proof of Theorem 6.3.4 is provided in Appendix E.2.

As shown in Theorem 6.3.4, $\mathcal{P}_l$, $l \in \mathcal{A}$, is a convex optimization problem. In particular, resorting to the epigraph representation of $\mathcal{P}_l$, as well as the fact that the square-root function is monotonically increasing, $\mathcal{P}_l$, $l \in \mathcal{A}$, is equivalent to the following Second-Order Cone Program (SOCP) [39], provided that $\mathcal{S}_c$ can be expressed in terms of second-order constraint, as it

is the case of polyhedra and ellipsoids:

$$\bar{\mathcal{P}}_l \begin{cases} \min\limits_{\boldsymbol{V},\boldsymbol{P},t} & t \\ \text{s.t.} & \sqrt{\lambda(\boldsymbol{V})} \leq t \\ & \|\boldsymbol{v}_h\| \leq v_{\max} \\ & \|\boldsymbol{v}_{h'} - \boldsymbol{v}_{h'-1}\| \leq \Delta v_{\max} \\ & \|\mathbf{C}_p^{-\frac{1}{2}} (\boldsymbol{p}_1 - \boldsymbol{v}_1 \Delta t)\| \leq \sqrt{\epsilon} \\ & \|\mathbf{C}_p^{-\frac{1}{2}} \left(\boldsymbol{p}_{h'} - (\boldsymbol{v}_{h'} \Delta t + \boldsymbol{p}_{h'-1})\right)\| \leq \sqrt{\epsilon} \\ & \boldsymbol{p}_h \in \mathcal{S}_c \\ & h = 1, \ldots, N, \quad h' = 2, \ldots, N \\ & \|\boldsymbol{p}_k - \boldsymbol{p}_c\| \leq \delta, \quad k = l, \ldots, l + L - 1 \\ & \boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \ldots, \boldsymbol{v}_N] \in \mathbb{R}^{2,N} \\ & \boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n, \ldots, \boldsymbol{p}_N] \in \mathbb{R}^{2,N} \\ & t \in \mathbb{R} \end{cases} \tag{6.11}$$

where $\sqrt{\lambda(\boldsymbol{V})} = \left\| \left(\mathbf{H}_M^{\mathrm{T}} \mathbf{C}_M^{-1} \mathbf{H}_M\right)^{\frac{1}{2}} \mathrm{vec}\left(\boldsymbol{V} - \boldsymbol{V}_0\right) \right\|$.

The optimization algorithm, named N-COST, follows directly by Theorem 6.3.4, and consists is solving all the feasible problems $\mathcal{P}_l$, $l \in \mathcal{A}$. The optimal solution among all of them will also be the optimal solution of $\mathcal{P}$, as summarized in Algorithm 1. The computational complexity connected with the implementation of the N-COST algorithm depends on the number of outer iterations $N-L+1$, as well as on the complexity of each iteration. Precisely, the complexity of each cycle corresponds to that required to solve (for our specific application) a SOCP, which is $O(N^{3.5}\log(\zeta))$ (see [33]), where ζ is a prescribed accuracy. Thus, the overall complexity is $O(N^{4.5}\log(\zeta))$. A brute force implementation procedure to solve the problem $\mathcal{P}$ would not be feasible, given that the optimization space would need to grow exponentially with the length of the optimization variables (position and velocity) for any given discretization step of them.

Summarizing N-COST provides the optimal solution in virtue of Theorem 6.3.4 and is efficient from a computational perspective.

Remark 6.3.5. *When implementing the N-COST algorithm, the resolution of the $N-L+1$ problems $\mathcal{P}_l$ could lead to a number n_l of overlapping optimal solutions, with $1 < n_l \leq N - L + 1$. In such a case it does not make any difference which solution is chosen among all the possible optimal ones.*

Algorithm 1 N-COST Algorithm

1: **Input.** $\mathbf{H}_M$, $\mathbf{C}_M$, $\boldsymbol{V}_0$, $\mathbf{C}_p$, $\boldsymbol{p}_c$, L, $v_{\max}$, $\Delta v_{\max}$, ϵ, $\delta, \mathcal{S}_c$.
2: **Output.** The optimal trajectory $\boldsymbol{P}^\star$ as well as the optimal mean velocity profile $\boldsymbol{V}^\star$.
3: **Initialization.** Set $l = 0$, $v = +\infty$, $\boldsymbol{V}^\star = \mathbf{0}$, and $\boldsymbol{P}^\star = \mathbf{0}$.
4: **repeat**
5: $l = l + 1$.
6: Solve Problem $\mathcal{P}_l$. If $v(\mathcal{P}_l) < v$, then $v = v(\mathcal{P}_l)$, $\bar{\boldsymbol{V}}^\star = \boldsymbol{V}^l$ and $\bar{\boldsymbol{P}}^\star = \boldsymbol{P}^l$. Otherwise go to step 5.
7: **until** $l \leq N - L + 1$
8: **Output.** If $v < +\infty$, $\boldsymbol{V}^\star = \bar{\boldsymbol{V}}^\star$ and $\boldsymbol{P}^\star = \bar{\boldsymbol{P}}^\star$, otherwise the design problem is not feasible.

6.4 N-COST algorithm performance analysis

In order to assess the effectiveness of the N-COST algorithm, in this section we investigate some interesting case studies concerning both synthetic and real-world scenarios.

To solve the $N - L + 1$ problems $\mathcal{P}_l$ in the N-COST algorithm, we used *CVX*, a package for specifying and solving convex programs [100, 101].

6.4.1 Analysis of a synthetic scenario

The configuration set-up for the synthetic scenario considers a vessel navigating under nominal conditions along a straight route with constant velocity $\boldsymbol{v}_0 = [8,\ 0]^{\mathrm{T}}$ m/s, which, at some point, turns its AIS device off for an overall time of $T = 15$ hours. Under hypothesis H_1, the vessel is supposed to follow an N-section path with $N = 20$, and equal time intervals $\Delta t = 0.75$ hours. Along its stealth trajectory, the vessel must reach the rendezvous point $\boldsymbol{p}_c = [150,\ 50]^{\mathrm{T}}$ km, where it is confined for a period of at least $L\Delta t = 0.75$ hours, with $L = 1$. The stopping area size is defined by $\delta = 100$ m. The other constraint parameters are: $v_{\max} = 10$ m/s, $\Delta v_{\max} = 1$ m/s, and $\epsilon = 0.1$.

Finally, the reversion rate associated to the underlying OU dynamic model is set as $\gamma = 0.9 \cdot 10^{-2}$ and noise level $\sigma^2 = 10^{-2}$, for both coordinates. The noise covariance matrix is set as $\mathbf{C}_{\boldsymbol{n}_m} = \mathrm{diag}(50^2, 50^2, 1, 1)$, $\forall m = 1, \ldots, M$, while $\mathbf{C}_{\boldsymbol{n}_0} = \mathrm{diag}(10^2, 10^2, 1, 1)$.

First analysis

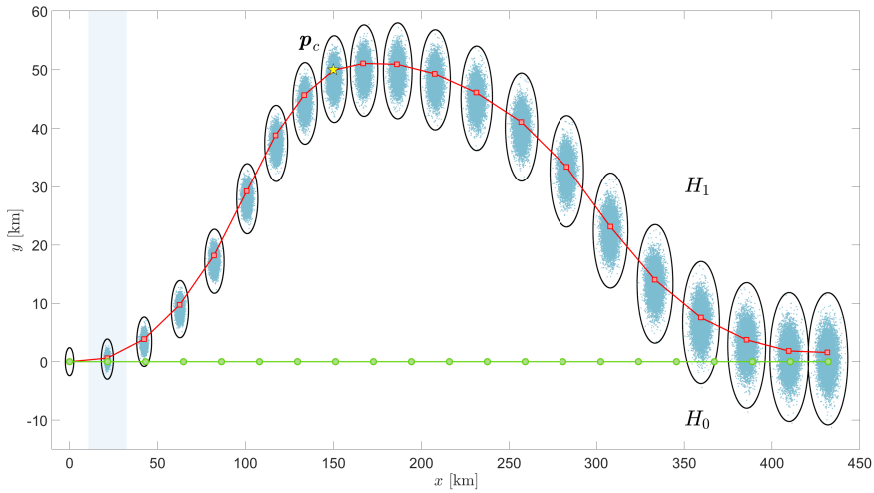
As first analysis, we study the possible output trajectory from the optimizer. The optimized trajectory under hypothesis H_1 is drawn in Fig. 6.1 as against the nominal trajectory (whose positions are marked with green dots) under hypothesis H_0. As already mentioned in 6.3,

- the optimal control of the velocity sequence returns a *stochastic* trajectory forced to evolve according to an OU process with long-run mean velocities determined by the N-COST algorithm. The OU process positions are depicted by small cerulean points in Fig. 6.1.

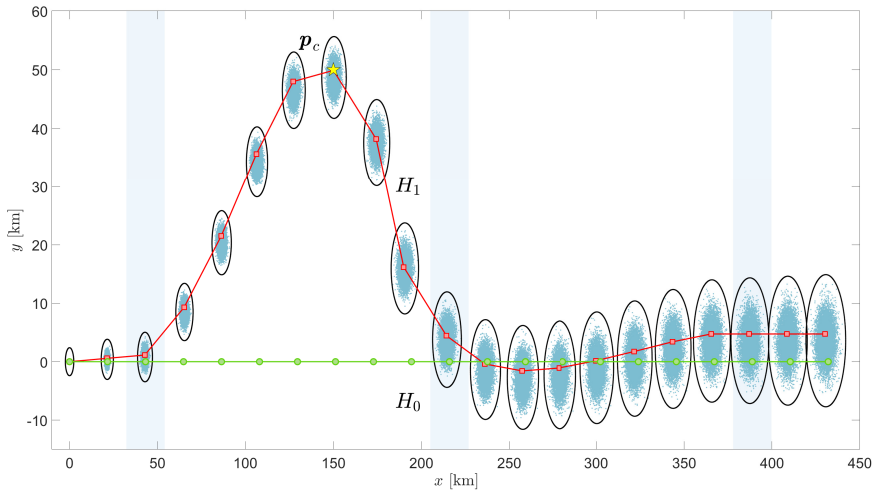
- the optimal control of both the velocity and position sequences returns a *deterministic* trajectory forced to evolve according to any realization of an OU process. The deterministic positions resulting from the application of N-COST algorithm are represented by red square markers in Fig. 6.1.

The comparison between the deterministic positions and the OU process positions shows that the N-COST trajectory is compatible with the OU statistical evolution, always being in the region of uncertainty indicated by the related ellipses (black solid lines) at any of the time intervals. Specifically, in the first synthetic scenario depicted in Fig. 6.1(a), in addition to the two AIS contacts available, there is a single radar contact located in $p_r = 1$, highlighted by a light blue patch. Instead, Fig. 6.1(b) shows the N-COST trajectory (red square markers) forced to evolve according to the presence of three in-between radar contacts located in $p_r \in \{2, 10, 18\}$ (light blue patches), and achieved by relaxing the constraint on the maximum absolute acceleration with $\Delta v_{\max} = 5$ m/s.

The performance of detector (6.1) is reported in Fig. 6.2 in terms of false alarm probability P_{FA} versus detection probability P_{D}. Three different simulations are considered. In the first one (black solid line), the trajectory under H_1 is assumed to evolve according to an OU process with optimal long-run mean velocity sequence provided by the N-COST algorithm, and the related ROC curve is given by (4.27). In the second simulation (red dashed line), the trajectory under H_1 is exactly the deterministic output of the N-COST algorithm, and the only randomness is in the measurements. The ROC curve is determined using the detector (6.1) with 10^4 Monte Carlo runs. The third simulation concerns a heuristic and sub-optimal deviation

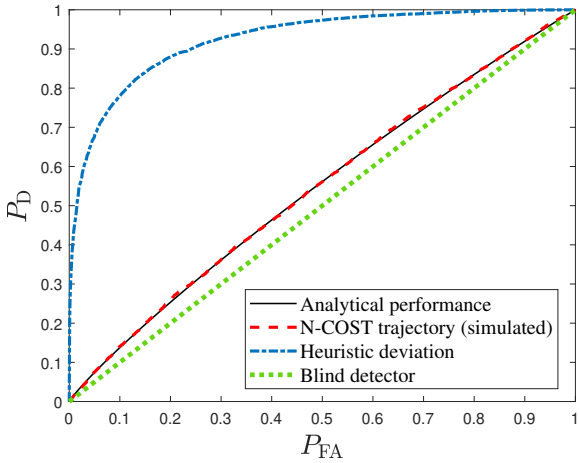


(a) Single in-between contact.

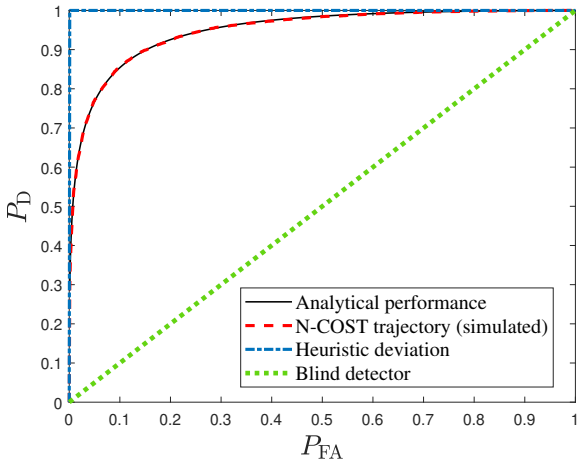


(b) Multiple in-between contacts.

Figure 6.1: Comparison of the optimized positions (red squares) versus the positions drawn from the OU process evolution (cerulean points) and the positions of the nominal trajectory (green dots). The ellipses (black solid lines) represent the 95%-confidence covariance of the OU/IOU process given the initial point of the trajectory. The point $\boldsymbol{p}_c$ near which the vessel lingers is marked with a yellow star. In the subplot (a) a single radar contact, indicated by light blue shade, located in $p_r = 1$ is taken into account, while in (b) there are multiple radar contacts located in $p_r \in \{2, 10, 18\}$.



(a) Single in-between contact.



(b) Multiple in-between contacts.

Figure 6.2: ROC curves describing the performance of detector (6.1) in terms of false alarm probability P_{FA} versus detection probability P_{D} when the optimal deviation, occurring under H_1, is an OU process (black solid line) and a deterministic output of the N-COST algorithm (red dashed line). The first ROC curve is provided by (4.27), while the second is simulated with 10^4 Monte Carlo runs. The blue dash-dotted ROC curve describes the sub-optimal deviation, while the blind detector ROC $P_{\mathrm{D}} = P_{\mathrm{FA}}$ is indicated by the green dotted line. The subplot (a) refers to the single radar contact located at the beginning of the N-path, while the subplot (b) highlights the difference in terms of performance achieved by considering multiple in-between radar contacts along the N-path.

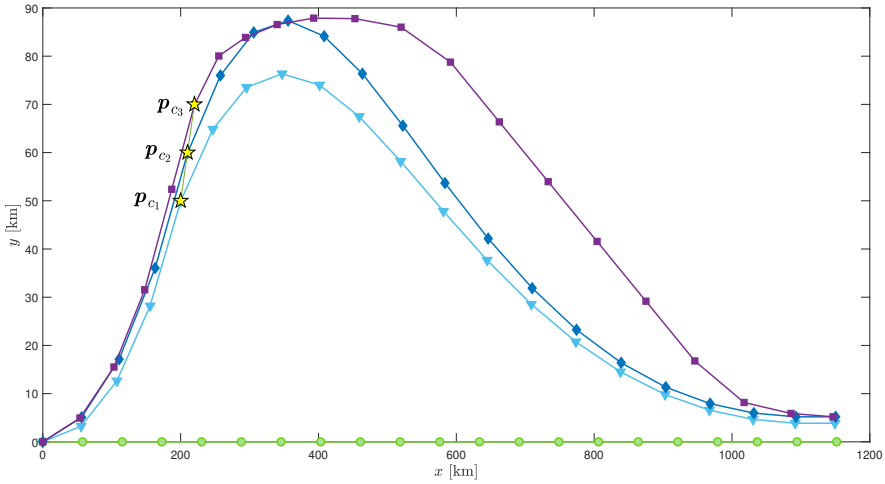


Figure 6.3: Three different optimized stealth trajectories, reaching the rendezvous points (denoted with yellow star markers) $\boldsymbol{p}_{c_1} = [200,\ 50]^{\mathrm{T}}$ km, $\boldsymbol{p}_{c_2} = [210,\ 60]^{\mathrm{T}}$ km, and $\boldsymbol{p}_{c_3} = [220,\ 70]^{\mathrm{T}}$ km, respectively.

represented by a first path section with velocity $\boldsymbol{v}_0$ and then a triangle trajectory (with vertices $\boldsymbol{p}_0$, $\boldsymbol{p}_c$ and $\boldsymbol{v}_0 T$) executed by a vessel that goes to $\boldsymbol{p}_c$, stops there and then comes back to the nominal route. In Fig. 6.2(a) it is possible to observe that both the optimized scenarios exhibit similar (poor) detection performance; and indeed, the ROC is close to that of the blind detector $P_{\mathrm{D}} = P_{\mathrm{FA}}$ (green dotted line), and this is not surprising, as the optimized trajectory was designed precisely to overcome the detector and the single radar contact provides a very low contribution to the detector since it is located where the vessel is still very close to the nominal position. While the detection performance is quite limited for the optimized deviation, the sub-optimal deviation is easily detectable (blue dash-dotted ROC curve) being P_{D} around 80% when P_{FA} is around 10%. A different situation is shown in Fig. 6.2(b), where the three radar contacts affect the performance related to both the optimized scenarios (black solid and red dashed lines) by making the corresponding trajectories much more detectable, but still far less detectable than with a stealth-agnostic deviation.

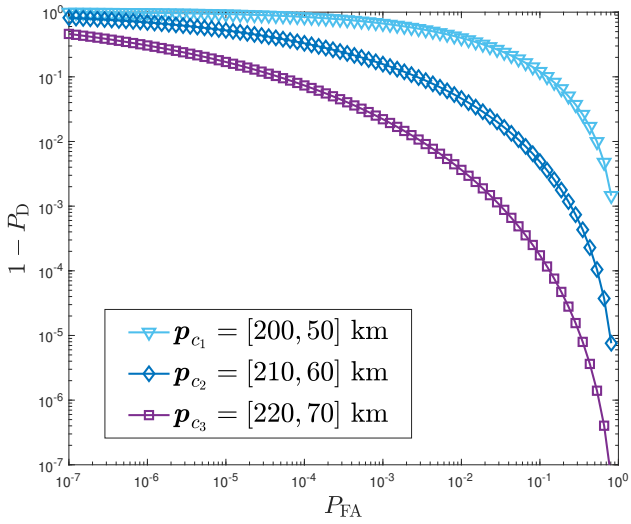


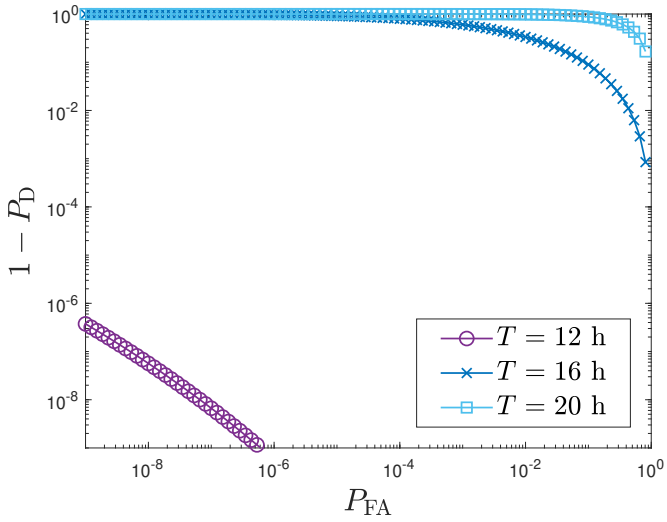
Figure 6.4: Performance of detector (6.1) in terms of missed detection probability $1 - P_{\mathrm{D}}$ versus false alarm probability P_{FA} for three different locations of the rendezvous point.

Second analysis

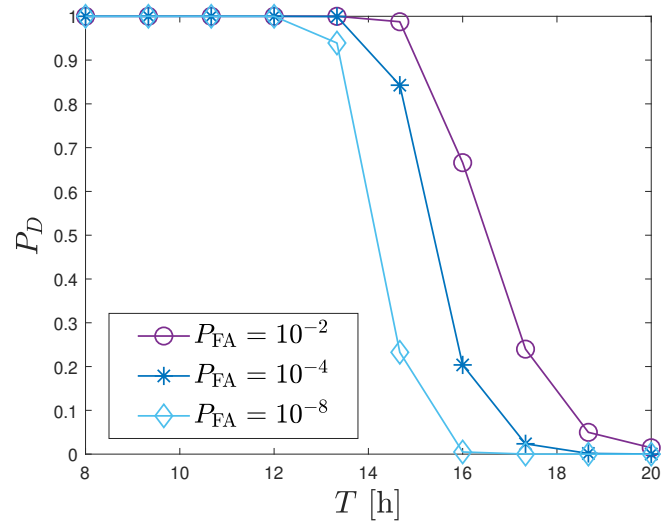
The performance of detector (6.1) is then analyzed at different locations of the rendezvous point, indicated in Fig. 6.3 with $\boldsymbol{p}_{c_1} = [200,\ 50]^{\mathrm{T}}$ km, $\boldsymbol{p}_{c_2} = [210,\ 60]^{\mathrm{T}}$ km, and $\boldsymbol{p}_{c_3} = [220,\ 70]^{\mathrm{T}}$ km. Figure 6.4 shows a clear performance improvement as the rendezvous point moves away from the nominal trajectory.

Third analysis

The proposed N-COST algorithm is also tested against varying time windows when the AIS transponder is shut down. Specifically, T is supposed to assume values within the set $\{12, 16, 20\}$ hours. The performance of detector (6.1) is shown in Fig. 6.5(a), and, as expected, the detection probability decreases with time, underlining that the longer the AIS transponder is inactive, the harder it is to detect possible anomalous deviation from the nominal route. Figure 6.5(b) depicts again the performance of detector (6.1), but in terms of detection probability versus the period T for three different fixed values of the false alarm probabilities, i.e., $P_{\mathrm{FA}} \in \left\{10^{-2}, 10^{-4}, 10^{-8}\right\}$; in



(a) Missed detection probability vs. false alarm probability.



(b) Detection probability vs. time.

Figure 6.5: Panel (a): Performance of detector (6.1) in terms of missed detection probability $1 - P_{\mathrm{D}}$ versus false alarm probability P_{FA} for three different periods of the AIS shutdown, i.e., $T \in \{12, 16, 20\}$ hours. Panel (b): Performance of detector (6.1) in terms of detection probability P_{D} versus T for three different values of the false alarm probability, i.e., $P_{\mathrm{FA}} \in \{10^{-2}, 10^{-4}, 10^{-8}\}$.

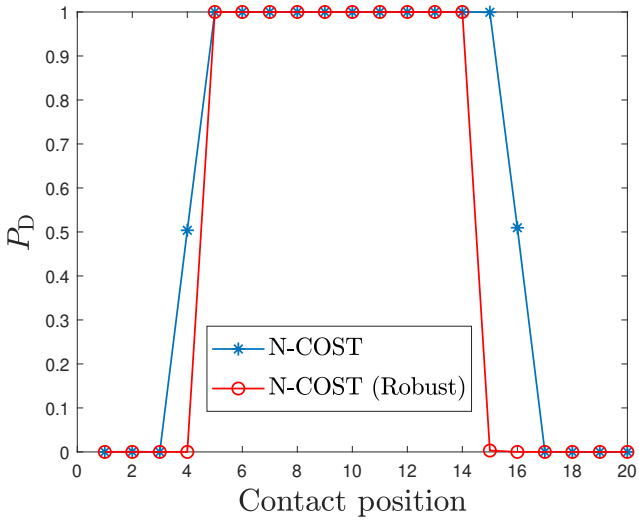


Figure 6.6: Comparison between the robust N-COST and original N-COST algorithms with $P_{FA} = 10^{-8}$ with a single contact observing the vessel.

such a case, the common trend shows a decrease to zero after a transitional period, underlining the increasing difficulty in performing the anomaly detection with the increasing time of the AIS device shutdown. We can observe a clear *phase transition* characterized by a critical time, dependent on the choice of the P_{FA}, before which $P_D \approx 1$ and the vessel is detectable, while after this time $P_D \approx 0$. To summarize, right after the critical time the vessel is able to accomplish its mission remaining hidden, and before the critical time the vessel is detectable while accomplishing its mission. Basically, the analysis shown in this third synthetic scenario draws the guideline in choosing the most suitable period T of the AIS device shutdown in such a way that the vessel is hardly detectable.

Fourth analysis - robust design

We explored the possibility to design the optimization strategy when the measurement time slot is not perfectly known. Specifically, we assume to optimize the trajectory under a worst condition scenario, where a measurement is supposed to observe the vessel not only in a specific time but also in

its neighboring slots.[3] The effectiveness of such robust design is tested in the same conditions of the first synthetic scenario (see Fig. 6.1(a)). Specifically, measurements are assumed to be in time slots $2-4$ and $16-18$ to derive the robust N-COST. This robust solution is then tested against the original N-COST (reported in Fig. 6.1(a)) optimized assuming a single contact only in the second time slot.

In Fig. 6.6 we compare in terms of P_{D} (see (4.27) with $P_{\mathrm{FA}} = 10^{-8}$) the robust and original N-COST varying the position of a single contact. Clearly, both solutions do not perform well in the centering region, where no measurements should have been taken, while the vessel is moving away from the nominal trajectory towards the rendezvous point. However, the robust N-COST is not detectable from time slot 1 to 4 and from 15 to 20 while the original N-COST is sensibly more detectable in time slots 4, 15 and 16. In conclusion, under the same physical constraints and for the same mission to accomplish, a noticeable improvement of performance is observed.

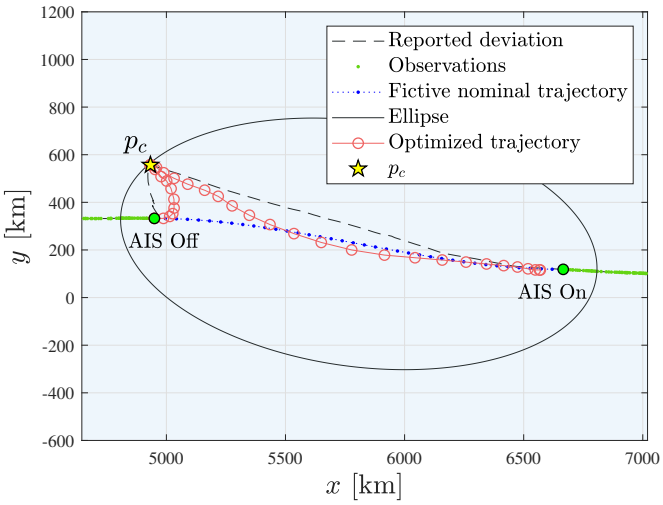
6.4.2 Analysis of real-world vessel traffic data

We applied the proposed optimization strategy to the real-world AIS track of the vessel navigating in the waters of the Pacific Ocean and approaching the Galápagos Marine Reserve [17], previously introduced in Subsection 1.3.5. The information provided by the track reveals a deviation from the nominal route, depicted in Fig. 6.7(a) by the black dashed line (ground truth), in order to rendezvous with other ships in the specified point $\boldsymbol{p}_c$ depicted by a yellow star marker.

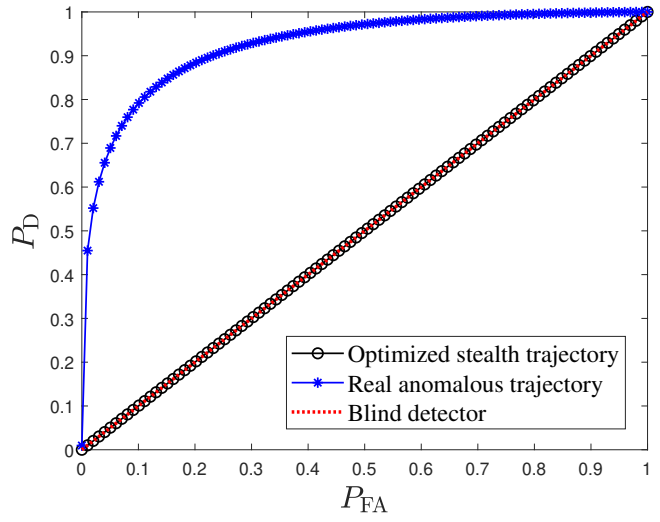
To test the N-COST algorithm, we artificially introduced a gap in AIS data of about $T = 145.5$ hours to simulate the AIS device shutdown during the actual deviation. Additionally, we assumed a fictitious nominal trajectory (blue dotted line) when the data gap occurs, which is characterized by a piecewise nominal velocity sequence $\boldsymbol{V}_0 \in \mathbb{R}^{2,N}$, with $N = 37$, and whose mean value is $\boldsymbol{v}_0 = [3.36,\ -0.42]^{\mathrm{T}}$ m/s. The other OU parameters are also estimated on this fictitious nominal trajectory, whose values are given by $\boldsymbol{\gamma} = \left[1.63 \times 10^{-5},\ 5.95 \times 10^{-4}\right]$ and $\boldsymbol{\sigma} = \left[1.1 \times 10^{-2},\ 1.58 \times 10^{-2}\right]$.

As regards the optimization procedure, the parameters related to the constraints are set as: $v_{\max} = 10$ m/s, $\Delta v_{\max} = 1$, $\epsilon = 1$, and $\delta = 1$ km. Moreover, we imposed an ellipse constraint (black solid line) to contain $\boldsymbol{p}_c$

[3] Clearly, other robust design strategies are possible and will be investigated in future works.



(a)



(b)

Figure 6.7: Panel (a): Real-world AIS track (black dashed line) versus optimized trajectory (red circle markers line). Panel (b): Performance of the anomaly detector (6.1) shows the optimized trajectory being less detectable then the real one.

and the resulting optimized trajectory, and it has been designed by aligning the focal points with the positions where the AIS device is turned OFF and where is then reactivated, respectively.

As can be seen in Fig. 6.7(a), the optimized trajectory is effectively devised (red circle markers line) so that the vessel can arrive to the meeting point, loiter there for the desired time ($L\Delta t \simeq 19.6$ hours), and finally restore the nominal condition before the AIS device reactivation.

The performance of detector (6.1) is provided in terms of P_{FA} versus P_{D} in Fig. 6.7(b), showing the ROC curves related to the real trajectory (blue asterisks markers curve) and the optimized one (black circle markers curve). Remarkably, the synthesized stealth trajectory exhibits a behavior similar to the real anomalous route, but it turns out to be stealthy as the ROC is almost equal to the blind detector (red dotted line), revealing the effectiveness of the new proposed route planning tool.

6.5 Conclusion

In this chapter we propose a computationally efficient technique, called N-COST algorithm, to solve the route planning problem with the goal to make a vessel's trajectory as stealthy as possible to an anomaly detector, so as to hide a deviation from a nominal traffic route to accomplish a specific mission.

The proposed approach assumes the velocity of the vessel evolves according to an OU mean-reverting stochastic process, while proper kinematic and practical constraints are taken into account. The optimization problem minimizes the KL divergence between the statistical hypotheses of the nominal and the anomalous trajectories. Interesting case studies concerning both synthetic and real-world scenarios are reported to prove the effectiveness of the N-COST algorithm.

Chapter 7

Conclusion

This chapter summarizes the contribution of the thesis and discusses avenues for future research.

7.1 Summary of contribution

The proposed research aims at contributing to advances in the maritime anomaly detection methodologies in order to improve the ability to reveal, understand, anticipate and prevent illegitimate activities at sea. The work has been developed in the statistical framework by exploiting *i*) a prior information from a maritime traffic graph (MTG) that can be derived from a route atlas or from historical data, *ii*) the Ornstein-Uhlenbeck (OU) mean reverting stochastic process to model the vessel's dynamics, and *iii*) the complete or incomplete observation of the available data.

The first part of the research work has proposed a novel maritime anomaly detector to reveal path deviations of ships during an intentional Automatic Identification System (AIS) device disablement, in order to possibly perform activities that would not be normally allowed. The problem has been treated within the statistical hypothesis testing framework, based on the Generalized Likelihood Ratio Test for Gaussian linear model, designed to identify changes in the OU long-run mean velocity parameter of the vessel. The proposed anomaly detection strategy has been successfully applied in two real-world scenarios: the illicit fishing activities around the Galápagos Marine Reserve in 2017, and the grounding of the container vessel *Ever Given* in the Suez Canal in 2021.

The second part of the research work proceeded considering that the information contained in AIS messages can be entered manually by the crew and maliciously altered, reporting falsified dynamic information and ruining the reliability of AIS data. Therefore, the problem addressed consists in determining whether the AIS data received from a vessel are trustworthy or not, and whether a deviation from the expected route occurred or not, with the support of additional reliable information provided by surveillance systems. The proposed solution involves two detection strategies both designed in a hypothesis testing framework and based on the changes of the OU process velocity parameter. The first strategy splits the joint problem into two binary hypothesis tests via the Generalized Likelihood Ratio Test (GLRT). Specifically, the first GLRT decides if the AIS data are truthful or not, while the second decides if a deviation occurred or not, and finally the two decisions are combined through a decision rule named Coupled GLRT (C-GLRT). The second strategy is designed to tackle a multiple hypothesis test, relying on the Model Order Selection methodology, which applies an appropriate penalty term to the maximized log-likelihood. Specifically, the direct Kullback-Leibler approach, also known as No-Name rule, the Akaike's Information Criteria, and the Bayesian Information Criteria have been considered among all the different criteria to choose such a penalty term. The effectiveness of the two detection strategies has been assessed through the investigation of some case studies concerning a maritime operational scenario. The performance of both strategies has been analyzed by varying some parameters, such as the falsified trajectory, the configurations (number and position) of the AIS contacts and the additional reliable contacts provided by the surveillance system, and the time window.

Finally, the third and last part of the research work moved from the anomaly detector to the opponent vessel's point of view, and proposed an optimization methodology to make the anomalous trajectory of the malicious vessel as stealth as possible to the anomaly detector build in the first part of this work. Assuming that the vessel can be only observed by satellite sensors, and that it has knowledge of the surveillance system acquisition instants, the route planning is formalized as a min-max problem by capitalizing the Kullback-Leibler divergence between the statistical hypotheses of the nominal and the anomalous trajectories as key performance measure. A computationally efficient technique, called Non-Convex Optimized Stealth Trajectory N-COST algorithm, is proposed to handle the resulting con-

strained optimization problem, where physical requirements are accounted for at the design stage. Such requirements concern the mean velocity dynamic, the compliance of the trajectory with the underlying OU statistical model, the existence of a time instant allowing a possible rendezvous, and sea coast limitations. Interesting case studies, concerning both synthetic and real-world scenarios, are also investigated to prove the effectiveness of the proposed planning strategy. Leveraging the proposed optimization tool, it is possible to compute the worst-case conditions in terms of anomaly detection capability and consequently to determine the optimal surveillance system acquisition instants in the region of interest.

7.2 Directions for future work

The findings presented throughout this work show excellent performance, but it is essential to mention also the limitations related to the detection procedure:

- An extension of the problem should encompass a multi-target scenario and/or the presence of clutter, prompting a data association issue; this offers the opportunity to delve into the application of the belief propagation methodology able to provide a highly effective, efficient, and scalable solution to the probabilistic data association problem [91, 152].

- The assumptions on the perfect knowledge of the nominal velocity could be questioned. The nominal velocity could actually not be perfectly known in advance, and, thus, should be described by a proper stochastic model that takes into account also the associated uncertainty. For instance, if the MTG exhibits more than one edge in a certain area of interest, then the nominal velocity can be modeled as a mixture whose number of components depends on the number of the edges in the graph.

- Due to environmental and/or financial changes, a variation of the velocity and/or the planned routes should be taken into consideration. For instance, what was standard in 2019 is not anymore after the outbreak of the Coronavirus disease (COVID-19), since an unprecedented drop in maritime mobility across all categories of commercial shipping

> and a generally reduced activity has been observed when the most severe restrictions were in force [156].

Future areas of research could build up on this research work by looking at improving the anomaly detection capabilities by exploring the aforementioned limitations and coming up with innovative solutions.

Moreover, regardless the strategy proposed in Chapter 5, further works will be in the direction of mathematically establishing the observability of the AIS data spoofing and deviations anomalies from the set of available data, and more advanced anomalous behaviors could be considered to test the (C-GLRT) approach. Another important aspect that can be investigated is related to the use of the proposed approach in the context of the air traffic control. Specifically, the behavior of stealth anomalous deviations of private aircraft could be taken into consideration.

Although in Chapter 4 it has been demonstrated that even with few contacts (e.g., space-borne acquisitions) the anomaly detection procedure can work very well when the trajectory is not optimized against the detector, a possible future development could overcome the lack of an optimal surveillance asset. Indeed, countermeasures to the N-COST algorithm could be investigated to benefit the maximization of the anomaly detection capability in the worst condition, and this could include the optimal satellite scheduling acquisition combined with the optimal path planning of the surveillance asset.

Appendix A

Automatic Identification System

This appendix is related to the Automatic Identification System (AIS), previously presented in Chapter 1.

A.1 AIS functioning

AIS integrates a standardized VHF transceiver to automatically broadcast information, such as a vessel's position, speed, and navigational status, at regular intervals. The information originates from the vessel's navigational sensors, typically its GPS receiver together with a gyrocompass or a rate of turn indicator. Permanent information, such as the vessel name and VHF call sign, are programmed when installing the equipment and is also transmitted regularly. More precisely, the acquired GPS coordinates are broadcast via the dedicated Marine VHF bandwidth (through two radio channels operating at 161.975 MHz and 162.025 MHz and based on a time division multiple access (TDMA) system, which permits each device to communicate during a given time period, or time slot) [114].

In order to transmit and receive AIS signals, some dedicated devices have been put in place since the introduction of the system: class-A transponders, for large ships, and class-B transponders, for smaller vessels. The IMO [111] imposes that every ship of more than 300 GT, all passenger ships and all fishing vessels with a length above 15 meters be equipped with class-A transponders. Conversely, class-B transponders are designed to bring the benefits of AIS on smaller vessels; indeed, they are smaller and less expensive than class-A type transceivers. As such, they can be installed on small ships such as

recreational vessels that want to have the benefits of having the AIS even if, for their size, they are not required to fit a transponder onboard. A class-A AIS transponder allows vessels to automatically broadcast and receive static, dynamic, and voyage-related information as listed in Table A.1 [114]: vessels while underway transmit information updates every 2 to 10 seconds, while vessels at anchor transmit information every 3 minutes. Static permanent information are broadcast every 6 minutes. Whereas, position updates for

Table A.1: Information broadcast from a class-A AIS transponder.

Every 2 to 10 seconds while underway and every 3 minutes while at anchor	
MMSI	unique referenceable nine digit identification number
Navigation status	"at anchor" "under way using engines", "not under command"
Rate of turn	right or left, from 0 to 720 degrees per minute
Speed over ground	0.1-knot (0.19 km/h) resolution from 0 to 102 knots (189 km/h)
Positional resolution	Longitude and Latitude
Course over ground	relative to true north to 0.1°
True heading	0 to 359°
True bearing at own position	0 to 359°
Time stamp	the seconds field of the UTC time when these data were generated
Every 6 minutes	
IMO ship identification number	a seven digit number that remains unchanged upon transfer of the ship's registration to another country
Radio call sign	international radio call sign, up to 7 characters, assigned to the vessel by its country of registry
Name	20 characters to represent the name of the vessel
Type of ship	"cargo", "tanker", "passenger", "fishing"
Dimension of ship	to nearest meter
Location of positioning system's antenna on board	in meters aft of bow and meters port or starboard
Type of positioning system	various options such as GPS or differential GPS
Draught of ship	0.1 meter to 25.5 meters
Destination	maximum 20 characters
ETA at destination	month, day, hour, and minute in UTC

class-B transponders are broadcast less often than class-A transponders.

Furthermore, two types of transmission are available: terrestrial and by satellite. At first, the system was only terrestrial, with transmission occurring from one vessel to another, or between a vessel and a land based systems, such as VTS systems, in a range of distance which is limited by the curvature of the Earth. However, the development of satellites enabled

to receive messages even far from the coast line, by uploading and storing the received messages then downloading information as soon as a coast line and a shore station is reached. Therefore, when out of range of terrestrial networks, vessels fitted with AIS device can be tracked through a growing number of satellites that are fitted with special AIS receivers, known as S-AIS. Using satellites with existing AIS technology gives a comprehensive picture of vessel activity, even in high-traffic and remote areas. Hence, S-AIS greatly extends the range of traditional AIS: signals are sent and received from many kilometers above land and sea, so that the barrier of the horizon does not limit these signals. The two types of transmissions are conceivably represented in Fig. 1.1 shown in Chapter 1.

Appendix B

Ornstein-Uhlenbeck stochastic process

This appendix is related to the OU model, previously presented in Chapter 3. Here we present the solution of the SDE (3.7) which provides for the target state prediction and the related variance (see Subsection 3.2.3).

B.1 Coupled OU process

The solution of the SDE (3.7) can be found using Itô calculus [93, 163], and the first two moments are as follows

$$\boldsymbol{s}\left(t|t_0\right) = \mathrm{e}^{\mathbf{A}(t-t_0)}\boldsymbol{s}(t_0) + \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}(t-s)}\mathbf{D}\,\boldsymbol{v}\,\mathrm{d}s, \tag{B.1}$$

and

$$\mathrm{Cov}\left[\boldsymbol{s}(t)|\boldsymbol{s}(t_0)\right] = \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}(t-s)}\mathbf{B}\mathbf{B}^{\mathrm{T}}\left(\mathrm{e}^{\mathbf{A}(t-s)}\right)^{\mathrm{T}}\mathrm{d}s. \tag{B.2}$$

Assuming that $\boldsymbol{\Theta}$ is diagonalizable, then $\boldsymbol{\Theta} = \mathbf{R}\boldsymbol{\Gamma}\mathbf{R}^{-1}$ represents its eigen-decomposition, and the following relations hold

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -\boldsymbol{\Theta} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -\mathbf{R}\boldsymbol{\Gamma}\mathbf{R}^{-1} \end{bmatrix} = \widetilde{\mathbf{R}}\widetilde{\mathbf{A}}\widetilde{\mathbf{R}}^{-1}, \tag{B.3}$$

where

$$\widetilde{\mathbf{R}} \triangleq \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}, \quad \text{and} \quad \widetilde{\mathbf{A}} \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -\boldsymbol{\Gamma} \end{bmatrix}. \tag{B.4}$$

We can now exploit the power series of the matrix exponential to obtain a convenient expression for $\mathrm{e}^{\mathbf{A}t}$:

$$\begin{aligned}
\mathrm{e}^{\mathbf{A}t} &= \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k t^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \prod_{i=0}^{k} \left(\widetilde{\mathbf{R}} \widetilde{\mathbf{A}} \widetilde{\mathbf{R}}^{-1} \right) t^k \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \widetilde{\mathbf{R}} \widetilde{\mathbf{A}}^k t^k \widetilde{\mathbf{R}}^{-1} \\
&= \widetilde{\mathbf{R}} \left(\sum_{k=0}^{\infty} \frac{1}{k!} \widetilde{\mathbf{A}}^k t^k \right) \widetilde{\mathbf{R}}^{-1} \\
&= \widetilde{\mathbf{R}}\, \mathrm{e}^{\widetilde{\mathbf{A}}t}\, \widetilde{\mathbf{R}}^{-1},
\end{aligned} \tag{B.5}$$

that can be reworked to highlight the dependence on t and $\boldsymbol{\gamma}$ by defining the state transition matrix

$$\boldsymbol{\Phi}(t, \boldsymbol{\gamma}) \triangleq \mathrm{e}^{\widetilde{\mathbf{A}}t} = \begin{bmatrix} \mathbf{I} & \left(\mathbf{I} - \mathrm{e}^{-\boldsymbol{\Gamma} t} \right) \boldsymbol{\Gamma}^{-1} \\ \mathbf{0} & \mathrm{e}^{-\boldsymbol{\Gamma} t} \end{bmatrix}. \tag{B.6}$$

On the other hand, the integral in (B.1) can be written in terms of $\boldsymbol{\Phi}(t, \boldsymbol{\gamma})$ as

$$\boldsymbol{\Psi}(t, \boldsymbol{\gamma}) \triangleq \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}(t-s)} \mathbf{D}\, ds = \int_{t_0}^{t} \boldsymbol{\Phi}(t-s, \boldsymbol{\gamma}) \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Gamma} \end{bmatrix} \mathrm{d}s, \tag{B.7}$$

and allows a closed-form solution

$$\boldsymbol{\Psi}(t, \boldsymbol{\gamma}) = \begin{bmatrix} t\mathbf{I} - \left(\mathbf{I} - \mathrm{e}^{-\boldsymbol{\Gamma} t} \right) \boldsymbol{\Gamma}^{-1} \\ \mathbf{I} - \mathrm{e}^{-\boldsymbol{\Gamma} t} \end{bmatrix}, \tag{B.8}$$

which is identified as the control input function.

Combining together (B.1), (B.6) and (B.8) lead us to the first moment of the SDE solution, which has the form

$$\boldsymbol{s}\,(t|t_0) = \widetilde{\mathbf{R}} \boldsymbol{\Phi}(t - t_0, \boldsymbol{\gamma}) \widetilde{\mathbf{R}}^{-1} \boldsymbol{s}(t_0) + \widetilde{\mathbf{R}} \boldsymbol{\Psi}(t - t_0, \boldsymbol{\gamma}) \widetilde{\mathbf{R}}^{-1} \boldsymbol{v}. \tag{B.9}$$

We can proceed analogously with the second order solution using (B.2) and (B.6)

$$\mathrm{Cov}\left[\boldsymbol{s}(t) | \boldsymbol{s}(t_0) \right] = \widetilde{\mathbf{R}} \left(\int_{t_0}^{t} \boldsymbol{\Phi}(t-s, \boldsymbol{\gamma}) \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widetilde{\boldsymbol{\Sigma}} \end{bmatrix} \boldsymbol{\Phi}^{\mathrm{T}}(t-s, \boldsymbol{\gamma})\, \mathrm{d}s \right) \widetilde{\mathbf{R}}^{-1}, \tag{B.10}$$

where $\widetilde{\boldsymbol{\Sigma}}$ is the noise covariance in the *transformed* domain, wose entries are defined as follows

$$\widetilde{\boldsymbol{\Sigma}} \triangleq \mathbf{R}^{-1}\boldsymbol{\Sigma}\left(\mathbf{R}^{-1}\boldsymbol{\Sigma}\right)^{\mathrm{T}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}. \tag{B.11}$$

Again, the problem can be solved in algebraically closed-form, and for the second moment of the SDE solution, we obtain

$$\mathrm{Cov}\left[\boldsymbol{s}(t)|\boldsymbol{s}(t_0)\right] = \widetilde{\mathbf{R}}\,\boldsymbol{\Omega}(t-t_0)\widetilde{\mathbf{R}}^{-1}, \tag{B.12}$$

where $\boldsymbol{\Omega}(t) = \boldsymbol{\Omega}_1 \circ \boldsymbol{\Omega}_2(t)$, with the $\circ$ operator denoting the Hadamard product. Matrix $\boldsymbol{\Omega}_1$ has the following form

$$\boldsymbol{\Omega}_1 = \begin{bmatrix} \frac{\sigma_x^2}{\gamma_x^3} & \frac{\sigma_{xy}}{\gamma_x\gamma_y} & \frac{\sigma_x^2}{2\gamma_x^2} & \frac{2\sigma_{xy}}{\gamma_x} \\ \frac{\sigma_{xy}}{\gamma_x\gamma_y} & \frac{\sigma_y^2}{\gamma_y^3} & \frac{2\sigma_{xy}}{\gamma_y} & \frac{\sigma_y^2}{2\gamma_y^2} \\ \frac{\sigma_x^2}{2\gamma_x^2} & \frac{2\sigma_{xy}}{\gamma_y} & \frac{\sigma_x^2}{\gamma_x} & \frac{2\sigma_{xy}}{\gamma_x+\gamma_y} \\ \frac{2\sigma_{xy}}{\gamma_x} & \frac{\sigma_y^2}{2\gamma_y^2} & \frac{2\sigma_{xy}}{\gamma_x+\gamma_y} & \frac{\sigma_y^2}{\gamma_y} \end{bmatrix}, \tag{B.13}$$

whereas $\boldsymbol{\Omega}_2(t)$ is defined as

$$\boldsymbol{\Omega}_2(t) = \begin{bmatrix} \boldsymbol{\Xi}_1(t) & \boldsymbol{\Xi}_2(t) \end{bmatrix}, \tag{B.14}$$

with

$$\boldsymbol{\Xi}_1(t) = \begin{bmatrix} f(t\,\gamma_x) & h(t,\boldsymbol{\gamma}) \\ h(t,\boldsymbol{\gamma}) & f(t\,\gamma_y) \\ k(t\,\gamma_x) & \frac{g\left(\gamma_x\frac{t}{2}\right)}{\gamma_x} - \frac{g\left((\gamma_x+\gamma_y)\frac{t}{2}\right)}{\gamma_x+\gamma_y} \\ \frac{g\left(\gamma_y\frac{t}{2}\right)}{\gamma_y} - \frac{g\left((\gamma_x+\gamma_y)\frac{t}{2}\right)}{\gamma_x+\gamma_y} & k(t\,\gamma_y) \end{bmatrix},$$

$$\boldsymbol{\Xi}_2(t) = \begin{bmatrix} k(t\,\gamma_x) & \frac{g\left(\gamma_y\frac{t}{2}\right)}{\gamma_y} - \frac{g\left((\gamma_x+\gamma_y)\frac{t}{2}\right)}{\gamma_x+\gamma_y} \\ \frac{g\left(\gamma_x\frac{t}{2}\right)}{\gamma_x} - \frac{g\left((\gamma_x+\gamma_y)\frac{t}{2}\right)}{\gamma_x+\gamma_y} & k(t\,\gamma_y) \\ g(t\,\gamma_x) & g\left((\gamma_x+\gamma_y)\frac{t}{2}\right) \\ g\left((\gamma_x+\gamma_y)\frac{t}{2}\right) & g(t\,\gamma_y) \end{bmatrix},$$

by using the following ancillary functions:

$$f(t) \triangleq \frac{1}{2}\left(2t + 4\mathrm{e}^{-t} - \mathrm{e}^{-2t} - 3\right), \tag{B.15}$$

$$g(t) \triangleq \frac{1}{2}\left(1 - \mathrm{e}^{-2t}\right), \tag{B.16}$$

which are the prediction position and velocity error normalized variance, respectively, and

$$h(t,\boldsymbol{\gamma}) \triangleq t - \frac{1-\mathrm{e}^{-t\gamma_x}}{\gamma_x} - \frac{1-\mathrm{e}^{-t\gamma_y}}{\gamma_y} + \frac{1-\mathrm{e}^{-t(\gamma_x+\gamma_y)}}{\gamma_x+\gamma_y}, \tag{B.17}$$

$$k(t) \triangleq \mathrm{e}^{-2t}\left(1-\mathrm{e}^{t}\right)^2. \tag{B.18}$$

Appendix C

Detection of maritime anomalous deviations

This appendix is related to the detection of maritime anomalous deviations, previously presented in Chapter 4. Here we outline the synthetic characterization of the overall process noise (see Subsection 4.3.1); we also calculate the off-diagonal terms in the covariance matrix of the data vector $\boldsymbol{y}$ in the multi-contacts case (see Subsection 4.3.2), and we derive the MLE of the unknown parameter $\boldsymbol{\theta}$ (see Section 4.4). Finally, we present an insight on the detection performance determining the scaling law of $\lambda(\boldsymbol{\theta})$ introduced in Section 4.4.

C.1 Synthetic characterization of $\boldsymbol{\omega}(T)$

We easily show that $\boldsymbol{\omega}(T) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(T))$ returning the synthetic characterization of $\boldsymbol{\omega}(T)$ (4.5) as follows

$$\mathbb{E}[\boldsymbol{\omega}(T)] = \sum_{n=1}^{N-1} \left[\prod_{i=n+1}^{N} \boldsymbol{\Phi}(\Delta_i) \right] \mathbb{E}[\boldsymbol{\omega}_n] + \mathbb{E}[\boldsymbol{\omega}_N] = \mathbf{0}. \tag{C.1}$$

$$\begin{aligned} \mathrm{Cov}[\boldsymbol{\omega}(T)] &= \mathbb{E}\left[\boldsymbol{\omega}(T)\boldsymbol{\omega}(T)^{\mathrm{T}}\right] \\ &\overset{i)}{=} \sum_{n=1}^{N-1} \left[\prod_{i=n+1}^{N} \boldsymbol{\Phi}(\Delta_i) \right] \mathbf{C}(\Delta_n) \left[\prod_{i=n+1}^{N} \boldsymbol{\Phi}(\Delta_i)^{\mathrm{T}} \right] + \mathbf{C}(\Delta_N) \\ &\overset{ii)}{=} \mathbf{C}\left(\sum_{n=1}^{N} \Delta_n \right) = \mathbf{C}(T). \end{aligned} \tag{C.2}$$

Since the $\boldsymbol{\omega}_n$ are independent zero-mean Gaussian random vectors, the terms involving the mean of the mixed products, $\mathbb{E}[\boldsymbol{\omega}_n \boldsymbol{\omega}_m]$, are zero, and equality $i)$ is valid.

The equality $ii)$ is proved by making the following considerations. For $N = 2$, by using the definitions of $\mathbf{C}(\Delta_n)$ in (B.12) and $\boldsymbol{\Phi}(\Delta_n)$ in (B.6), we can show that

$$\boldsymbol{\Phi}(\Delta_2)\mathbf{C}(\Delta_1)\boldsymbol{\Phi}(\Delta_2)^{\mathrm{T}} + \mathbf{C}(\Delta_2) = \mathbf{C}(\Delta_1 + \Delta_2). \tag{C.3}$$

It can be easily shown that for $N = 3$, exploiting the previous expression, we can get

$$\begin{aligned} &\boldsymbol{\Phi}(\Delta_2 + \Delta_3)\mathbf{C}(\Delta_1)\boldsymbol{\Phi}(\Delta_2 + \Delta_3)^{\mathrm{T}} + \boldsymbol{\Phi}(\Delta_3)\mathbf{C}(\Delta_2)\boldsymbol{\Phi}(\Delta_3)^{\mathrm{T}} + \mathbf{C}(\Delta_3) \\ &\quad = \boldsymbol{\Phi}(\Delta_3)\left[\boldsymbol{\Phi}(\Delta_2)\mathbf{C}(\Delta_1)\boldsymbol{\Phi}(\Delta_2)^{\mathrm{T}} + \mathbf{C}(\Delta_2)\right]\boldsymbol{\Phi}(\Delta_3)^{\mathrm{T}} + \mathbf{C}(\Delta_3) \\ &\quad = \boldsymbol{\Phi}(\Delta_3)\mathbf{C}(\Delta_1 + \Delta_2)\boldsymbol{\Phi}(\Delta_3)^{\mathrm{T}} + \mathbf{C}(\Delta_3) \\ &\quad = \mathbf{C}(\Delta_1 + \Delta_2 + \Delta_3). \end{aligned} \tag{C.4}$$

This procedure is valid for each N by induction.

C.2 Off-diagonal terms in $\mathbf{C}_{\boldsymbol{y}}$ calculation

The off-diagonal terms in the covariance matrix $\mathbf{C}_{\boldsymbol{y}}$ (4.20) can be derived as follows. Let us suppose for simplicity that p_i and p_j, denoting the time

location of the i-th and the j-th contacts, respectively, are integer quantities. The result is still valid even if p_i and p_j are not integer.

$$\begin{aligned}\mathbf{C}_{ij} =&\mathbb{E}\left[(\boldsymbol{y}_i-\boldsymbol{\mu}_i)(\boldsymbol{y}_j-\boldsymbol{\mu}_j)^{\mathrm{T}}\right]\\ =&\mathbb{E}\left[(\boldsymbol{\omega}(T_i)+\boldsymbol{n}_i+\boldsymbol{\Phi}(T_i)\boldsymbol{n}_0)(\boldsymbol{\omega}(T_j)+\boldsymbol{n}_j+\boldsymbol{\Phi}(T_j)\boldsymbol{n}_0)^{\mathrm{T}}\right]\\ =&\mathbb{E}\left[\boldsymbol{\omega}(T_i)\boldsymbol{\omega}(T_j)^{\mathrm{T}}\right]+\boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_j)^{\mathrm{T}}\\ \overset{i)}{=}&\sum_{n=1}^{p_i}\left[\prod_{l=n+1}^{p_i}\boldsymbol{\Phi}(\Delta_l)\right]\mathbf{C}(\Delta_n)\left[\prod_{l'=n+1}^{p_j}\boldsymbol{\Phi}(\Delta_{l'})^{\mathrm{T}}\right]+\boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_j)^{\mathrm{T}}\\ =&\mathbf{C}(T_i)\boldsymbol{\Phi}(T_j-T_i)^{\mathrm{T}}+\boldsymbol{\Phi}(T_i)\mathbf{C}_{\boldsymbol{n}_0}\boldsymbol{\Phi}(T_j)^{\mathrm{T}},\end{aligned}\tag{C.5}$$

$\forall i,j=1,...,K$. In i) it has been considered that the random variables $\boldsymbol{\omega}(\Delta_n)$ involved in the expressions of $\boldsymbol{\omega}(T_i)$ and $\boldsymbol{\omega}(T_j)$ given in (4.15) , are all independent so that $\mathbb{E}[\boldsymbol{\omega}_n\boldsymbol{\omega}_m^{\mathrm{T}}]=0$ if $n\neq m$. In the same way it is shown that $\mathbf{C}_{ji}=\mathbf{C}_{ij}^{\mathrm{T}}$, $\forall i,j=1,2,\ldots,K$ with $T_i<T_j$.

C.3 MLE of the unknown parameter $\boldsymbol{\theta}$

Incorporating the term independent of the unknown parameter appearing in the GLRT (4.26) in the threshold τ, under the condition in which N and $\mathcal{D}_N$ are unknown, the MLE of $\boldsymbol{\mu}=\boldsymbol{\theta}$ is given by

$$\begin{aligned}\widehat{\boldsymbol{\theta}}&=\arg\max_{\boldsymbol{\theta}}\left\{\ln\left[p_{\boldsymbol{\theta}}(\boldsymbol{y})\right]\right\}\\ &=\arg\min_{\boldsymbol{\theta}}\left\{(\boldsymbol{y}-\boldsymbol{\theta})^{\mathrm{T}}\,\mathbf{C}_{\boldsymbol{y}}^{-1}\,(\boldsymbol{y}-\boldsymbol{\theta})\right\},\end{aligned}\tag{C.6}$$

whose exact expression can be found by setting to zero the derivative with respect to $\boldsymbol{\theta}$:

$$\frac{\partial}{\partial\boldsymbol{\theta}}\left\{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{C}_{\boldsymbol{y}}^{-1}\boldsymbol{\theta}-2\boldsymbol{\theta}^{\mathrm{T}}\mathbf{C}_{\boldsymbol{y}}^{-1}\boldsymbol{y}\right\}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}=0,\tag{C.7}$$

from which it follows that the MLE of the unknown parameter corresponds to data, $\widehat{\boldsymbol{\theta}}=\boldsymbol{y}$.

In the alternative case where N and $\mathcal{D}_N$ are known, the MLE of the vector $\boldsymbol{\theta}=\boldsymbol{v}_{1:N}$ is given by

$$\widehat{\boldsymbol{\theta}}=\arg\min_{\boldsymbol{\theta}}\left\{(\boldsymbol{y}-\mathbf{H}\,\boldsymbol{\theta})^{\mathrm{T}}\,\mathbf{C}_{\boldsymbol{y}}^{-1}\,(\boldsymbol{y}-\mathbf{H}\,\boldsymbol{\theta})\right\},\tag{C.8}$$

whose exact expression can be found by setting to zero the derivative with respect to $\boldsymbol{\theta}$:

$$\frac{\partial}{\partial \boldsymbol{\theta}} \left\{ \boldsymbol{\theta}^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H} \boldsymbol{\theta} - 2\boldsymbol{\theta}^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \boldsymbol{y} \right\} \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = 0, \tag{C.9}$$

from which

$$\widehat{\boldsymbol{\theta}} = \left(\mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1} \boldsymbol{y}. \tag{C.10}$$

In the same way it is shown that the MLE of the unknown parameter $\widetilde{\boldsymbol{\theta}}$, achieved by applying the rank reduction method, is given by

$$\widehat{\widetilde{\boldsymbol{\theta}}} = \left(\widetilde{\mathbf{H}}^{\mathrm{T}} \widetilde{\mathbf{H}} \right)^{-1} \widetilde{\mathbf{H}}^{\mathrm{T}} \widetilde{\mathbf{C}} \boldsymbol{y}. \tag{C.11}$$

C.4 More on the detection performance

Assuming that we know the matrix $\mathbf{H}$ (see scenario 2 in 4.4.2) we expect that the detector (4.35) outperforms the detector (4.28) designed under the assumption that $\mathbf{H}$ is unknown (see scenario 1 in 4.4.1). Indeed, given that the parameter $\lambda(\boldsymbol{\theta})$ is equal for both the scenarios, the number of degrees of freedom d makes the difference. In scenario 1, $d = 4K$, while, for scenario 2, d depends on the structure of $\mathbf{H}$ (or equivalently on the contact time locations): $d = \rho = \mathrm{rank}(\mathbf{H}) \leq \min(4K, 2N)$. Then, scenario 2 is never worse than scenario 1 when $\rho < 4K$ and equivalently when $\rho = 4K$.

In both scenarios, by increasing K, we would like to obtain an improvement of performance. Unfortunately, this is not guaranteed. However, when K is large enough and the structure of $\mathbf{H}$ allows $\lambda(\boldsymbol{\theta})$ to increase with K, the scenario 2 has performance improving with K, since d is bounded by $2N$, implying that $P_{\mathrm{D}} \longrightarrow 1$ for any fixed P_{FA}.

Conversely, in scenario 1, when $\lambda(\boldsymbol{\theta})$ increases with K, the performance is not guaranteed to improve as $d = 4K$ increases as well with K. In the following, we give the scaling law of $\lambda(\boldsymbol{\theta})$ to obtain improving performance of the detection strategy in scenario 1.

Scaling law of $\lambda(\boldsymbol{\theta})$

Let us indicate $d_K = 4K$. The decision statistic (4.28) has a Chi-squared distribution, that is equivalent to the sum of the squares of d_K independent

Normal random variables x_i^2

$$\mathcal{Q}_K = \sum_{i=1}^{d_K} x_i^2, \tag{C.12}$$

with zero mean under H_0, finite mean under H_1, and unit variance:

$$\begin{cases} H_0: & x_i \sim \mathcal{N}(0,1), \quad \mathcal{Q}_K \sim \chi^2_{d_K}, \\ H_1: & x_i \sim \mathcal{N}(q_i,1), \quad \mathcal{Q}_K \sim \chi'^2_{d_K}(\lambda_K), \end{cases} \tag{C.13}$$

where $\lambda_K = \sum_{i=1}^{d_K} q_i^2$. Moreover, we have

$$E[\mathcal{Q}_K] = \begin{cases} d_K & H_0 \\ \lambda_K + d_K & H_1 \end{cases} \tag{C.14}$$

$$Var[\mathcal{Q}_K] = \begin{cases} 2d_K & H_0 \\ 2(d_K + 2\lambda_K) & H_1. \end{cases} \tag{C.15}$$

Exploiting the *central limit theorem* [134], and normalizing the decision statistic $\mathcal{Q}_\mathcal{K}$ under H_0, we obtain the normalized decision statistic

$$\widetilde{\mathcal{Q}}_K \triangleq \frac{\mathcal{Q}_K - d_K}{\sqrt{2d_K}} \underset{H_0}{\overset{H_1}{\gtrless}} \tau, \tag{C.16}$$

which converges to a Gaussian distribution for large K:

$$\begin{cases} H_0: \ \widetilde{\mathcal{Q}}_K \longrightarrow \mathcal{N}(0,1) \\ H_1: \ \widetilde{\mathcal{Q}}_K \longrightarrow \mathcal{N}\left(\widetilde{\lambda}_\infty, \widetilde{\sigma}^2_\infty\right). \end{cases} \tag{C.17}$$

Under hypothesis H_1, $\widetilde{\mathcal{Q}}_K$ can be written as

$$\underbrace{\frac{\mathcal{Q}_K - \lambda_K - d_K}{\sqrt{2(d_K + 2\lambda_K)}}}_{\widetilde{\mathcal{Q}}_K^{(H_1)} \longrightarrow \mathcal{N}(0,1)} \underbrace{\sqrt{\frac{d_K + 2\lambda_K}{d_K}}}_{\widetilde{\sigma}_K \longrightarrow \widetilde{\sigma}_\infty} + \underbrace{\frac{\lambda_K}{\sqrt{2d_K}}}_{\widetilde{\lambda}_K \longrightarrow \widetilde{\lambda}_\infty}, \tag{C.18}$$

where the first limit is a convergence in distribution to a normal random variable while the others are limit of deterministic sequences. Summarizing we have

$$\widetilde{\lambda}_\infty = \lim_{K\to\infty} \widetilde{\lambda}_K = \lim_{K\to\infty} \frac{\lambda_K}{\sqrt{2d_K}}, \tag{C.19}$$

$$\widetilde{\sigma}^2_\infty = \lim_{K\to\infty} \widetilde{\sigma}^2_K = 1 + \lim_{K\to\infty} 2\frac{\lambda_K}{d_K}. \tag{C.20}$$

Exploiting the previous convergence properties, we can analyze the asymptotic detection performance of the decision statistic $\widetilde{\mathcal{Q}}_K$, specifically we have

$$P_{\mathrm{D}}^{\infty} = \lim_{K\to\infty} P_{\mathrm{D},K} = \lim_{K\to\infty} \mathrm{P}\left(\widetilde{\mathcal{Q}}_K > \tau \,|H_1\right) = Q\left(\frac{Q^{-1}\left(P_{\mathrm{FA}}^{\infty}\right) - \widetilde{\lambda}_{\infty}}{\widetilde{\sigma}_{\infty}}\right), \tag{C.21}$$

where P_{FA}^{∞} is the asymptotic false alarm probability, obtained exploiting the convergence in (C.16) and $\tau = Q^{-1}\left(P_{\mathrm{FA}}^{\infty}\right)$.

Let us recall that $d_K = \mathcal{O}(K)$ and assume $\widetilde{\lambda}_K = \mathcal{O}(K^{\frac{n}{2}})$. We distinguish three different cases depending on the parameter $n \geq 0$. If $0 \leq n < 1$, the test cannot distinguish H_0 from H_1 because $\widetilde{\lambda}_{\infty} = 0$ and $\widetilde{\sigma}_{\infty} = 1$, consequently from (C.21) $P_{\mathrm{D}}^{\infty} = P_{\mathrm{FA}}^{\infty}$. If $n = 1$ then $\widetilde{\lambda}_{\infty} < \infty$, $\widetilde{\sigma}_{\infty} = 1$ and the detection probability converges to a value less than 1 provided by (C.21). If $n > 1$ and $\tau < \infty$, then, under H_1, dividing the decision statistic (C.18) by $\widetilde{\lambda}_K$, we obtain

$$\frac{\widetilde{\mathcal{Q}}_K}{\widetilde{\lambda}_K} = \underbrace{\widetilde{\mathcal{Q}}_K^{(H_1)}}_{\longrightarrow\mathcal{N}(0,1)} \underbrace{\widetilde{\sigma}_K/\widetilde{\lambda}_K}_{\longrightarrow 0} + 1, \tag{C.22}$$

implying that $\widetilde{\mathcal{Q}}_K/\widetilde{\lambda}_K$ converges to one in probability under H_1. Then, following (C.21), and, since $\tau/\widetilde{\lambda}_K \longrightarrow 0$, the detection probability converges to one. Summarizing, for any $P_{\mathrm{FA}}^{\infty} < 1$, we have:

$$\begin{cases} P_{\mathrm{D}}^{\infty} = P_{\mathrm{FA}}^{\infty} & \text{if } 0 \leq n < 1, \\ P_{\mathrm{D}}^{\infty} = Q\left(Q^{-1}\left(P_{\mathrm{FA}}^{\infty}\right) - \widetilde{\lambda}_{\infty}\right) & \text{if } n = 1, \\ P_{\mathrm{D}}^{\infty} = 1 & \text{if } n > 1. \end{cases} \tag{C.23}$$

Note that the sequence q_i, for $i = 1, \ldots, d_K$, where in general $q_i \neq q_j$ for $i \neq j$, represents the heterogeneity of the contacts. Each q_i is a measure of the information contained in a component of a single measurement (position or velocity along one of the Cartesian axes). Clearly, if $q_i \ll 1$ this component is adding mostly noise to the decision statistic degrading the detection performance.

Appendix D

Detection of anomalous deviations in the presence of AIS spoofing

This appendix is related to the derivation of the GLRTs and the concerning MLEs previously presented in Chapter 5 (see Subsection 5.4.1 and Section 5.5).

D.1 Derivation of the GLRTs (5.20), (5.23), and (5.40)

Some of the MLE and GLRT properties used in this derivation are provided in [119], and also reported below for the sake of clarity.

Derivation of the GLRT (5.20)

Let us recall the data model after the SVD in (5.18). The derivation of the unrestricted MLE of $\widetilde{\boldsymbol{z}} = \left[\widetilde{\boldsymbol{\varphi}}^{\mathrm{T}}, \widetilde{\boldsymbol{v}}_R^{\mathrm{T}}\right]$, under hypothesis $\{H_2, H_3\}$, is the following

$$p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{\varphi}}}, \widehat{\widetilde{\boldsymbol{v}}}_R\right) = \max_{\widetilde{\boldsymbol{z}}} p\left(\boldsymbol{y}; \widetilde{\boldsymbol{z}}\right) = p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{z}}}\right),$$

and can be obtained from

$$\begin{aligned}\widehat{\widetilde{\boldsymbol{z}}} &= \arg\max_{\widetilde{\boldsymbol{z}}} \left\{\ln\left[p\left(\boldsymbol{y};\widetilde{\boldsymbol{z}}\right)\right]\right\} \\ &= \arg\min_{\widetilde{\boldsymbol{z}}} \left\{\left(\boldsymbol{y}-\widetilde{\mathbf{U}}\,\widetilde{\boldsymbol{z}}\right)^{\mathrm{T}} \mathbf{C}_{\boldsymbol{y}}^{-1}\left(\boldsymbol{y}-\widetilde{\mathbf{U}}\,\widetilde{\boldsymbol{z}}\right)\right\},\end{aligned}$$

whose exact expression can be found by setting to zero the derivative with respect to $\widetilde{\boldsymbol{z}}$, achieving

$$\begin{aligned}\widehat{\widetilde{\boldsymbol{z}}} = \begin{bmatrix}\widehat{\widetilde{\boldsymbol{\varphi}}} \\ \widehat{\widetilde{\boldsymbol{v}}}_R\end{bmatrix} &= \left(\widetilde{\mathbf{U}}^{\mathrm{T}}\mathbf{C}_{\boldsymbol{y}}^{-1}\widetilde{\mathbf{U}}\right)^{-1}\widetilde{\mathbf{U}}^{\mathrm{T}}\mathbf{C}_{\boldsymbol{y}}^{-1}\boldsymbol{y} \\ &\overset{i)}{=} \left(\widetilde{\mathbf{U}}^{\mathrm{T}}\widetilde{\mathbf{C}}^{\mathrm{T}}\widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}\right)^{-1}\widetilde{\mathbf{U}}^{\mathrm{T}}\widetilde{\mathbf{C}}^{\mathrm{T}}\widetilde{\mathbf{C}}\,\boldsymbol{y} \\ &\overset{ii)}{=} \left(\widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{Z}}\right)^{-1}\widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{C}}\,\boldsymbol{y} \sim \mathcal{N}\left(\widetilde{\boldsymbol{z}}, \left(\widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{Z}}\right)^{-1}\right), \qquad \text{(D.1)}\end{aligned}$$

where in i) the Cholesky decomposition has been applied to matrix $\mathbf{C}_{\boldsymbol{y}}^{-1}$, and in ii) $\widetilde{\mathbf{Z}} \triangleq \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}$ has been defined.

Since $\widehat{\widetilde{\boldsymbol{z}}}$ attains the Cramer-Rao Lower Bound (CRLB), it satisfies (see [118])

$$\frac{\partial}{\partial\widetilde{\boldsymbol{z}}}\ln p\left(\boldsymbol{y};\widetilde{\boldsymbol{z}}\right) = \boldsymbol{\mathcal{I}}\left(\widetilde{\boldsymbol{z}}\right)\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right) = \boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right)\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right), \qquad \text{(D.2)}$$

where $\boldsymbol{\mathcal{I}}(\widehat{\widetilde{\boldsymbol{z}}})$ is the Fisher Information Matrix (FIM) related to $\widehat{\widetilde{\boldsymbol{z}}}$ and defined as

$$\boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right) = \widetilde{\mathbf{Z}}^{\mathrm{T}}\widetilde{\mathbf{Z}} = \begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}} & \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{H}} \\ \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{G}} & \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\end{bmatrix}, \qquad \text{(D.3)}$$

with $\widetilde{\mathbf{G}} \triangleq \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_G$, and $\widetilde{\mathbf{H}} \triangleq \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_L$. Then, integrating equation (D.2) with respect to $\widetilde{\boldsymbol{z}}$ produces

$$\ln p\left(\boldsymbol{y};\widetilde{\boldsymbol{z}}\right) = -\frac{1}{2}\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right)\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right) + c\left(\widehat{\widetilde{\boldsymbol{z}}}\right),$$

or, since $c(\widehat{\widetilde{\boldsymbol{z}}}) = \ln p(\boldsymbol{y};\widehat{\widetilde{\boldsymbol{z}}})$,

$$p\left(\boldsymbol{y};\widetilde{\boldsymbol{z}}\right) = p\left(\boldsymbol{y};\widehat{\widetilde{\boldsymbol{z}}}\right)\exp\left[-\frac{1}{2}\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right)\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right)\right]. \qquad \text{(D.4)}$$

In order to find the MLE $\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}$, we need to maximize $p(\boldsymbol{y};\widetilde{\boldsymbol{\varphi}}=\mathbf{0},\widehat{\widetilde{\boldsymbol{v}}}_R)$ over $\widetilde{\boldsymbol{v}}_R$ under hypothesis $\{H_0, H_1\}$, i.e., we need to minimize

$$J\left(\widetilde{\boldsymbol{v}}_R\right) = \left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right)\left(\widehat{\widetilde{\boldsymbol{z}}}-\widetilde{\boldsymbol{z}}\right) \qquad \text{(D.5)}$$

over $\widetilde{\boldsymbol{v}}_R$, with $\widetilde{\boldsymbol{z}} = [\widetilde{\boldsymbol{\varphi}}^{\mathrm{T}} = \mathbf{0}, \widetilde{\boldsymbol{v}}_R^{\mathrm{T}}]^{\mathrm{T}}$. By replacing the components of the FIM (D.3) in equation (D.5), and annulling the gradient of (D.5) with respect to $\widetilde{\boldsymbol{v}}$ in the point $\widetilde{\boldsymbol{v}}_R = \widetilde{\boldsymbol{v}}_R^{(0)}$, we get

$$\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)} = \widehat{\widetilde{\boldsymbol{v}}}_R + \left(\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{G}}\,\widehat{\widetilde{\boldsymbol{\varphi}}},$$

$$J\left(\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right) = \widehat{\widetilde{\boldsymbol{\varphi}}}^{\mathrm{T}}\left[\widetilde{\mathbf{G}}^{\mathrm{T}}\,\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{G}}\right]\widehat{\widetilde{\boldsymbol{\varphi}}}.$$

Consequently, we achieve

$$p\left(\boldsymbol{y}; \widetilde{\boldsymbol{\varphi}} = \mathbf{0}, \widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right) = p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{\varphi}}}\right)\exp\left[-\frac{1}{2}J\left(\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right)\right],$$

and the GLRT is given by

$$L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) = \frac{p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{z}}}\right)}{p\left(\boldsymbol{y}; \widehat{\widetilde{\boldsymbol{z}}}\right)\exp\left[-\frac{1}{2}J\left(\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right)\right]} = \exp\left[\frac{1}{2}J\left(\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right)\right],$$

from which

$$\begin{aligned} 2\ln L_{\widetilde{\boldsymbol{\varphi}}}(\boldsymbol{y}) &= J\left(\widehat{\widetilde{\boldsymbol{v}}}_R^{(0)}\right) \\ &= \widehat{\widetilde{\boldsymbol{\varphi}}}^{\mathrm{T}}\left[\widetilde{\mathbf{G}}^{\mathrm{T}}\,\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\,\widetilde{\mathbf{G}}\right]\widehat{\widetilde{\boldsymbol{\varphi}}} \\ &= \widehat{\widetilde{\boldsymbol{\varphi}}}^{\mathrm{T}}\left(\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{\varphi}}\widetilde{\boldsymbol{\varphi}}}\right)^{-1}\widehat{\widetilde{\boldsymbol{\varphi}}}, \end{aligned}$$

where $\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{\varphi}}\widetilde{\boldsymbol{\varphi}}}$ is the matrix of size $\rho_G \times \rho_G$ corresponding to the upper-left partition of $\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})$, and it is shown to be the covariance matrix of $\widehat{\widetilde{\boldsymbol{\varphi}}}$:

$$\mathbf{C}_{\widehat{\widetilde{\boldsymbol{\varphi}}}} = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{\varphi}}\widetilde{\boldsymbol{\varphi}}}.$$

Derivation of the GLRT (5.23)

Starting from (D.2) and repeating the procedure in terms of $\widetilde{\boldsymbol{v}}_R$ (treating $\widetilde{\boldsymbol{\varphi}}$ as a nuisance parameter), the resulting GLRT is achieved as

$$2\ln L_{\widetilde{\boldsymbol{v}}_R}(\boldsymbol{y}) = \left(\widehat{\widetilde{\boldsymbol{v}}}_R - \widetilde{\boldsymbol{v}}_R^{(0)}\right)^{\mathrm{T}}\left(\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_R\widetilde{\boldsymbol{v}}_R}\right)^{-1}\left(\widehat{\widetilde{\boldsymbol{v}}}_R - \widetilde{\boldsymbol{v}}_R^{(0)}\right), \qquad \text{(D.6)}$$

where $\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_R\widetilde{\boldsymbol{v}}_R}$ is the matrix of size $\rho_H \times \rho_H$ corresponding to the lower-right partition of $\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})$, and it is shown to be the covariance matrix of $\widehat{\widetilde{\boldsymbol{v}}}_R$:

$$\mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_R} = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_R\widetilde{\boldsymbol{v}}_R}. \tag{D.7}$$

Derivation of the GLRT (5.40)

Manipulating the data model in (5.8) to bring matrix $\mathbf{L} \triangleq \mathbf{H} - \mathbf{G}$ up, after the SVD application (see (5.36) and (5.37)) we get

$$\boldsymbol{y} = \widetilde{\mathbf{U}}_G\,\widetilde{\boldsymbol{v}}_A + \widetilde{\mathbf{U}}_L\,\widetilde{\boldsymbol{v}}_R + \boldsymbol{\omega} = \begin{bmatrix}\widetilde{\mathbf{U}}_G & \widetilde{\mathbf{U}}_L\end{bmatrix}\begin{bmatrix}\widetilde{\boldsymbol{v}}_A \\ \widetilde{\boldsymbol{v}}_R\end{bmatrix} + \boldsymbol{\omega} = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{z}} + \boldsymbol{\omega} \tag{D.8}$$

so that, in this case, $\widetilde{\mathbf{U}} \triangleq \begin{bmatrix}\widetilde{\mathbf{U}}_G & \widetilde{\mathbf{U}}_L\end{bmatrix}$, and $\widetilde{\boldsymbol{z}} \triangleq \begin{bmatrix}\widetilde{\boldsymbol{v}}_A \\ \widetilde{\boldsymbol{v}}_R\end{bmatrix}$.

The unrestricted MLE $\widehat{\widetilde{\boldsymbol{z}}}$ attains the CRLB, and, hence, it satisfies the equality (D.2), where the FIM related to $\widehat{\widetilde{\boldsymbol{z}}}$ is defined as

$$\boldsymbol{\mathcal{I}}\left(\widehat{\widetilde{\boldsymbol{z}}}\right) = \begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}} & \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{L}} \\ \widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{G}} & \widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{L}}\end{bmatrix}, \tag{D.9}$$

with $\widetilde{\mathbf{G}} \triangleq \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_G$, and $\widetilde{\mathbf{L}} \triangleq \widetilde{\mathbf{C}}\,\widetilde{\mathbf{U}}_H$ (see the SVD applications (5.15) and (5.37)).

The same procedure applied for the derivation of the GLRTs (5.20) and (5.23) can be also implemented in this case by taking into account (D.9). Consequently, treating $\widetilde{\boldsymbol{v}}_R$ as a nuisance parameter, the resulting GLRT (5.40) has the following expression

$$2\ln L_G(\boldsymbol{y}) = \left(\widehat{\widetilde{\boldsymbol{v}}}_A - \widetilde{\boldsymbol{v}}_0\right)^{\mathrm{T}}\left(\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}\right)^{-1}\left(\widehat{\widetilde{\boldsymbol{v}}}_A - \widetilde{\boldsymbol{v}}_0\right), \tag{D.10}$$

where $\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}$ is the matrix of size $\rho_G \times \rho_G$ corresponding to the upper-left partition of $\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})$, and it corresponds to the covariance matrix of $\widehat{\widetilde{\boldsymbol{v}}}_A$:

$$\mathbf{C}_{\widehat{\widetilde{\boldsymbol{v}}}_A} = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}. \tag{D.11}$$

D.2 MLE expressions of $\widetilde{\varphi}$, $\widetilde{v}_R$ and $\widetilde{v}_A$

MLE of $\widetilde{\varphi}$

Considering the unrestricted MLE (D.1), which can be rewritten as

$$\begin{aligned}\begin{bmatrix}\widetilde{\varphi}\\ \widetilde{v}_R\end{bmatrix} &= \boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\\ \widetilde{\mathbf{H}}^{\mathrm{T}}\end{bmatrix}\widetilde{\mathbf{C}}\,\boldsymbol{y}\\ &= \begin{bmatrix}\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{\varphi}} & \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{v}_R}\\ \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{v}_R\widetilde{\varphi}} & \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{v}_R\widetilde{v}_R}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\\ \widetilde{\mathbf{H}}^{\mathrm{T}}\end{bmatrix}\widetilde{\mathbf{C}}\,\boldsymbol{y},\end{aligned} \tag{D.12}$$

and extracting the upper part, it is easy to get

$$\widehat{\widetilde{\varphi}} = \left(\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{\varphi}}\widetilde{\mathbf{G}}^{\mathrm{T}} + \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{v}_R}\widetilde{\mathbf{H}}^{\mathrm{T}}\right)\widetilde{\mathbf{C}}\,\boldsymbol{y}, \tag{D.13}$$

with

$$\begin{aligned}\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{\varphi}} &= \left[\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{G}}\right]^{-1},\\ \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{v}_R} &= -\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{\varphi}}\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\right)^{-1},\end{aligned}$$

which, replaced in equation (D.13), yield

$$\widehat{\widetilde{\varphi}} = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\varphi}\widetilde{\varphi}}\left[\widetilde{\mathbf{G}}^{\mathrm{T}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^{\mathrm{T}}\right]\widetilde{\mathbf{C}}\,\boldsymbol{y}, \tag{D.14}$$

where $\widetilde{\mathbf{C}}$ is the lower triangular matrix with positive elements on the diagonal derived from the Cholesky decomposition [96] of matrix $\mathbf{C}_{\boldsymbol{y}}^{-1}$.

MLE of $\widetilde{v}_R$

In the same way, it is possible to get the MLE of the velocity parameters vector from the lower part of equation (D.12) as

$$\widehat{\widetilde{v}}_R = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{v}_R\widetilde{v}_R}\left[\widetilde{\mathbf{H}}^{\mathrm{T}} - \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{G}}\left(\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}}\right)^{-1}\widetilde{\mathbf{G}}^{\mathrm{T}}\right]\widetilde{\mathbf{C}}\,\boldsymbol{y}. \tag{D.15}$$

MLE of $\widetilde{\boldsymbol{v}}_A$

In this case, considering the FIM (D.9), the unrestricted MLE can be expressed as

$$\begin{aligned}\begin{bmatrix}\widetilde{\boldsymbol{v}}_A\\ \widetilde{\boldsymbol{v}}_R\end{bmatrix} &= \boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\\ \widetilde{\mathbf{L}}^{\mathrm{T}}\end{bmatrix}\widetilde{\mathbf{C}}\,\boldsymbol{y}\\ &= \begin{bmatrix}\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A} & \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_R}\\ \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_R\widetilde{\boldsymbol{v}}_A} & \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_R\widetilde{\boldsymbol{v}}_R}\end{bmatrix}\begin{bmatrix}\widetilde{\mathbf{G}}^{\mathrm{T}}\\ \widetilde{\mathbf{L}}^{\mathrm{T}}\end{bmatrix}\widetilde{\mathbf{C}}\,\boldsymbol{y},\end{aligned} \tag{D.16}$$

and extracting the upper part, it is easy to get

$$\widehat{\widetilde{\boldsymbol{v}}}_A = \left(\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}\widetilde{\mathbf{G}}^{\mathrm{T}} + \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_R}\widetilde{\mathbf{L}}^{\mathrm{T}}\right)\widetilde{\mathbf{C}}\,\boldsymbol{y}, \tag{D.17}$$

with

$$\begin{aligned}\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A} &= \left[\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{L}}\left(\widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{L}}\right)^{-1}\widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{G}}\right]^{-1},\\ \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_R} &= -\left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}\widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{L}}\left(\widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{L}}\right)^{-1},\end{aligned}$$

which, replaced in equation (D.17), yield

$$\widehat{\widetilde{\boldsymbol{v}}}_A = \left[\boldsymbol{\mathcal{I}}^{-1}(\widetilde{\boldsymbol{z}})\right]_{\widetilde{\boldsymbol{v}}_A\widetilde{\boldsymbol{v}}_A}\left[\widetilde{\mathbf{G}}^{\mathrm{T}} - \widetilde{\mathbf{G}}^{\mathrm{T}}\widetilde{\mathbf{L}}\left(\widetilde{\mathbf{L}}^{\mathrm{T}}\widetilde{\mathbf{L}}\right)^{-1}\widetilde{\mathbf{L}}^{\mathrm{T}}\right]\widetilde{\mathbf{C}}\,\boldsymbol{y}. \tag{D.18}$$

Appendix E

Optimal opponent stealth trajectory planning

This appendix is related to the optimal opponent stealth trajectory planning, previously presented in Chapter 6. Here we demonstrate Remark 6.3.3 and Theorem 6.3.4 (see Subsection 6.3.3).

E.1 Proof of Remark 6.3.3

Let us prove that, given a temporal scale Δt_1, better performance can be obtained including an additional point in each time interval, i.e., doubling the degrees of freedom. To this end, let $(\boldsymbol{V}^\star, \boldsymbol{P}^\star)$ be an optimal solution to the planning problem corresponding to Δt_1, denoted, in the following, as $\mathcal{P}_1$. Without loss of generality, let us assume that the extra time instants $\bar{t}_h$, $h = 1, \ldots, N$, are given by $\bar{t}_h = t_{h-1} + \alpha \Delta t_1$, with $0 < \alpha < 1$, namely $t_h = \bar{t}_h + (1-\alpha)\Delta t_1$. Evidently, the velocity vector matrix $\bar{\boldsymbol{V}} = [\boldsymbol{v}_1^\star, \boldsymbol{v}_1^\star, \boldsymbol{v}_2^\star, \boldsymbol{v}_2^\star, \ldots, \boldsymbol{v}_N^\star, \boldsymbol{v}_N^\star] \in \mathbb{R}^{2,2N}$ is feasible to the design problem associated with $t_0, \bar{t}_1, t_1, \ldots, \bar{t}_N, t_N$, $\mathcal{P}_2$ say, and achieves the optimal value of $\mathcal{P}_1$.

Let us now focus on the definition/construction of a trajectory feasible for $\mathcal{P}_2$ based on $\boldsymbol{P}^\star$, i.e., $\bar{\boldsymbol{P}} = [\bar{\boldsymbol{p}}_1^\star, \boldsymbol{p}_1^\star, \bar{\boldsymbol{p}}_2^\star, \boldsymbol{p}_2^\star, \ldots, \bar{\boldsymbol{p}}_N^\star, \boldsymbol{p}_N^\star] \in \mathbb{R}^{2,2N}$ with $\bar{\boldsymbol{p}}_i$ the vessel position at time $\bar{t}_i$. To this end, let $\bar{\boldsymbol{p}}_i = (1-\alpha)\boldsymbol{p}_i + \alpha \boldsymbol{p}_{i+1}$, with $i = 1, \ldots, K$. By assumption,

$$\left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p(\Delta t_1)} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right) \leq \epsilon \tag{E.1}$$

Hence, denoting by $\Delta t_2 = (1-\alpha)\Delta t_1$ and $\beta = 1-\alpha$, it follows that

$$\begin{aligned}
&\left(\bar{\boldsymbol{p}}_i - (\boldsymbol{v}_i \alpha \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\alpha\Delta t_1)}} \left(\bar{\boldsymbol{p}}_i - (\boldsymbol{v}_i \alpha \Delta t_1 + \boldsymbol{p}_i)\right) \\
&= \alpha^2 \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\alpha\Delta t_1)}} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right) \\
&= \alpha \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\Delta t_1)}} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right) \leq \epsilon\alpha
\end{aligned} \tag{E.2}$$

and

$$\begin{aligned}
&\left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_2 + \bar{\boldsymbol{p}}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\Delta t_2)}} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_2 + \bar{\boldsymbol{p}}_i)\right) \\
&= \beta^2 \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\beta\Delta t_1)}} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right) \\
&= \beta \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right)^{\mathrm{T}} \mathbf{C}^{-1}_{p_{(\Delta t_1)}} \left(\boldsymbol{p}_{i+1} - (\boldsymbol{v}_i \Delta t_1 + \boldsymbol{p}_i)\right) \leq \epsilon\beta
\end{aligned} \tag{E.3}$$

As a consequence, $\bar{\boldsymbol{P}}$ complies with constraint C_2 in $\mathcal{P}_2$. Additionally, due the convexity of $\mathcal{S}_c$, $\{\bar{\boldsymbol{p}}_1^\star, \bar{\boldsymbol{p}}_2^\star, \ldots, \bar{\boldsymbol{p}}_N^\star\} \in \mathcal{S}_c$.

Finally, it follows that

$$\|\bar{\boldsymbol{p}}_i - \boldsymbol{p}_c\| = \left\|(1-\alpha)\boldsymbol{p}_i + \alpha\boldsymbol{p}_{i+1} - \boldsymbol{p}_c\right\| \leq \alpha \|\boldsymbol{p}_i - \boldsymbol{p}_c\| + (1-\alpha) \left\|\boldsymbol{p}_{i+1} - \boldsymbol{p}_c\right\|, \tag{E.4}$$

implying that the constructed trajectory also satisfies the constraint C_3 of $\mathcal{P}_2$ that accounts for the coast shape. Based on the above argumentation, $\left(\bar{\boldsymbol{V}}, \bar{\boldsymbol{P}}\right)$ can be said to be feasible to $\mathcal{P}_2$, and the optimal value of $\mathcal{P}_2$ greater than or equal to that of $\mathcal{P}_1$.

E.2 Proof of Theorem 6.3.4

The convexity of $\mathcal{P}_l$, $l \in \mathcal{A} = 1, \ldots, N-L+1$, easily follows from the convexity of the objective, as well as the constraint functions. In fact, $\lambda(\boldsymbol{V})$ is a quadratic function with a positive semi-definite matrix, thus the objective is a convex function. As to the requirements set, the constraints C_1, C_2 and C_4 involve the norms of linear functions of the optimization variables. Furthermore, the constraint C_4 is the point-wise maximum of convex functions [39]. Finally, constraint C_3 forces the optimization variable to lie in a convex set. Hence, the overall feasible set is convex as the intersection of convex sets.

To proceed further, let us observe that problems $\mathcal{P}$ and $\mathcal{P}_l$, $l \in \mathcal{A}$ share the same objective function. Let us observe that the feasible set of $\mathcal{P}$ is

larger than that of $\mathcal{P}_l$ for all $l \in \mathcal{A}$, being $\forall l \in \mathcal{A}$

$$\min_{h=1,\dots,N-L+1} \max_{k=h,\dots,h+L-1} (\|\boldsymbol{p}_k - \boldsymbol{p}_c\|) \leq \max_{k=l,\dots,l+L-1} (\|\boldsymbol{p}_k - \boldsymbol{p}_c\|) \tag{E.5}$$

Hence, $v(\mathcal{P}) \leq v(\mathcal{P}_l)$, $l \in \mathcal{A}$. Now, if $\mathcal{P}$ is not feasible then $\mathcal{P}_l$, $l \in \mathcal{A}$, is also not feasible and $v(\mathcal{P}) = v(\mathcal{P}_l) = +\infty$, $l \in \mathcal{A}$. On the contrary, let $(\boldsymbol{V}^\star, \boldsymbol{P}^\star)$ be an optimal solution to $\mathcal{P}$. Moreover, let

$$l^\star = \arg\min_{l' \in \mathcal{A}} \max_{k=l',\dots,h+L} (\|\boldsymbol{p}_k^\star - \boldsymbol{p}_c\|)\,. \tag{E.6}$$

Hence, $(\boldsymbol{V}^\star, \boldsymbol{P}^\star)$ is also feasible to $\mathcal{P}_{l^\star}$. As a result, $(\boldsymbol{V}^\star, \boldsymbol{P}^\star)$ is an optimal solution to $\mathcal{P}_{l^\star}$, and $v(\mathcal{P}) = \min_{l' \in \mathcal{A}} v(\mathcal{P}_{l'})$, which concludes the proof.

Appendix F

Publications

This research activity has led to several publications in international journals and conferences. These are summarized below.[1]

International Journals

1. **E. d'Afflisio**, P. Braca, L. M. Millefiori, P. Willett. "Detecting Anomalous Deviations From Standard Maritime Routes Using the Ornstein-Uhlenbeck Process", in *IEEE Transactions on Signal Processing*, vol. 66, no. 24, pp. 6474 – 6487, Dec. 2018. [DOI: 10.1109/TSP.2018.2875887]

2. A. Aubry, P. Braca, **E. d'Afflisio**, A. De Maio, L. M. Millefiori, P. Willett. "Optimal Opponent Stealth Trajectory Planning Based on an Efficient Optimization Technique", in *IEEE Transactions on Signal Processing*, vol. 69, pp. 270 – 283, Dec. 2020. [DOI: 10.1109/TSP.2020.3041925]

3. **E. d'Afflisio**, P. Braca, P. Willett. "Malicious AIS Spoofing and Abnormal Stealth Deviations: A Comprehensive Statistical Framework for Maritime Anomaly Detection", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2093–2108, Jun. 2021.
[DOI: 10.1109/TAES.2021.3083466]

4. N. Forti, **E. d'Afflisio**, P. Braca, L. M. Millefiori, P. Willett, and S. Carniel. "Maritime anomaly detection in a real-world scenario: Ever Given grounding in the Suez Canal", *IEEE Transactions on Intelligent Transportation Systems*, 2021. [DOI: 10.1109/TITS.2021.3123890]

[1] The author's bibliometric indices are the following: H-index = 2, total number of citations = 32 (source: Scopus on January 2022).

International Conferences and Workshops

1. **E. d'Afflisio**, P. Braca, L. M. Millefiori, P. Willett. "Maritime Anomaly Detection Based on Mean-Reverting Stochastic Processes Applied to a Real-World Scenario", in *21st International Conference on Information Fusion*, Cambridge (UK), 10-13 July 2018.
[DOI: 10.23919/ICIF.2018.8455854]
2. **E. d'Afflisio**, A. Aubry, P. Braca, A. De Maio, L. M. Millefiori, P. Willett. "Optimal Stealth Trajectory Design to Deceive Anomaly Detection Process", in *MTS/IEEE OCEANS'19 Marseille Conference and Exhibit*, Marseille (France), 17-20 June 2019. [DOI: 10.1109/OCEANSE.2019.8867147]
3. **E. d'Afflisio**, P. Braca, L. Chisci, G. Battistelli, P. Willett. "Maritime Anomaly Detection of Malicious Data Spoofing and Stealth Deviations from Nominal Route Exploiting Heterogeneous Sources of Information", in *24th International Conference on Information Fusion*, Sun City, South Africa, 1-4 November 2021.

Bibliography

[1] "Container ships headed for U.S. poised to worsen port bottleneck." [Online]. Available: https://www.bloomberg.com/news/articles/2021-10-22/container-ships-headed-for-u-s-poised-to-worsen-port-bottleneck

[2] "The cost of the suez canal blockage," *BBC News*. [Online]. Available: https://www.https://www.bbc.com/news/business-56559073

[3] "Freighter and boxship collide off Denmark," *The Maritime Executive*. [Online]. Available: https://www.maritime-executive.com/article/freighter-and-boxship-collide-off-denmark

[4] "How a full moon and a huge lever helped free Ever Given from Suez Canal." [Online]. Available: https://www.theguardian.com/world/2021/mar/30/powerful-tugs-and-an-ebbing-tide-how-the-ever-given-was-freed

[5] "Seven things you should know about AIS," *MarineTraffic Blog*. [Online]. Available: https://www.marinetraffic.com/blog/seven-things-know-ais

[6] "U.S. destroyer collides with tanker off Singapore; 10 missing." [Online]. Available: https://www.nbcnews.com/news/world/navy-destroyer-uss-john-s-mccain-collides-merchant-ship-east-n794386

[7] "USS Fitzgerald crash: Seven navy crew missing off Japan," *BBC News*. [Online]. Available: https://www.bbc.com/news/world-asia-40310563

[8] "VHF data exchange system (VDES) overview," *IALA Guideline on VDES*. [Online]. Available: https://www.iala-aism.org/product/vhd-data-exchange-system-vdes-overview-1117

[9] "National strategy for maritime security: National plan to achieve maritime domain awareness," U.S. Department of Homeland Security, Tech. Rep., Oct. 2005.

[10] "Rules of navigation," Accessed May 6, 2021, Suez Canal Authority, Tech. Rep., 2015. [Online]. Available: https://www.suezcanal.gov.eg/English/Navigation/Pages/RulesOfNavigation.aspx

[11] "Functional services for command and control of maritime operations (TRITON), Book II - Part IV SOW Annex A: System Requirements Specification (SRS)," NATO Communications and Information Agency, Tech. Rep., Jun 2016.

[12] "Annual overview of marine casualties and incidents 2020," Accessed April 30, 2021, EMSA, Tech. Rep., 2020. [Online]. Available: http://www.emsa.europa.eu/newsroom/latest-news/item/4266-annual-overview-of-marine-casualties-and-incidents-2020.html

[13] "Collision between container vessel ANL Wyong and gas carrier King Arthur," Accessed May 9, 2021, UK Marine Accident Investigation Branch, Tech. Rep., 2020. [Online]. Available: https://www.gov.uk/maib-reports/collision-between-container-vessel-anl-wyong-and-gas-carrier-king-arthur

[14] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716 – 723, 1974.

[15] H. Akaike, *On the Likelihood of a Time Series Model*. New York, NY: Springer New York, 1998, pp. 289 – 307.

[16] J. Alava and F. Paladines, "Illegal fishing on the Galápagos high seas," *Science*, Sep. 2017.

[17] J. J. Alava, M. J. Barragán-Paladines, J. Denkinger, L. Muñoz-Abril, P. Jiménez, F. Paladines, and et al., "Massive Chinese fleet jeopardizes threatened shark species around the Galápagos Marine Reserve and waters off Ecuador: Implications for national and international fisheries policy," *International Journal of Fisheries Science and Research*, 2017.

[18] E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: a neural network based database mining system for credit card fraud detection," in *IEEE/IAFE Computational Intelligence for Financial Engineering*, 1997, pp. 220 – 226.

[19] C. R. Alya Itani, "Introducing the Pattern of Life (PoL) concept for maritime traffic," in *GAST Workshop*, 2019.

[20] F. Amato, M. Fiorini, S. Gallone, and G. Golino, "Fully solid state radar for vessel traffic services," in *11th International Radar Sympsium*, Jun. 2010, pp. 1 – 5.

[21] M. Anneken, Y. Fischer, and J. Beyerer, "Evaluation and comparison of anomaly detection algorithms in annotated datasets from the maritime domain," in *SAI Intelligent Systems Conference (IntelliSys)*, 2015, pp. 169 – 178.

[22] V. F. Arguedas, F. Mazzarella, and M. Vespe, "Spatio-temporal data mining for maritime situational awareness," in *OCEANS 2015 - Genova*, 2015, pp. 1 – 8.

[23] V. F. Arguedas, G. Pallotta, and M. Vespe, "Automatic generation of geographical networks for maritime traffic surveillance," in *17th International Conference on Information Fusion*, 2014, pp. 1 – 8.

[24] A. Aubry, P. Braca, E. d'Afflisio, A. De Maio, L. M. Millefiori, and P. Willett, "Optimal opponent stealth trajectory planning based on an efficient optimization technique," *IEEE Transactions on Signal Processing*, vol. 69, pp. 270 – 283, 2020.

[25] M. Balduzzi, A. Pasta, and K. Wilhoit, "A security evaluation of AIS automated identification system," *30th Annual Computer Security Applications Conference*, 2014.

[26] R. Bale, "Thousands of sharks found on boat in huge illegal haul," *National Geographic*, Aug. 2017. [Online]. Available: https://www.nationalgeographic.com/animals/article/wildlife-watch-galapagos-illegal-shark-fishing

[27] Y. Bar-Shalom and X. Rong-Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[28] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[29] Y. Bar-Shalom, P. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT: YBS Publishing, Apr. 2011.

[30] O. E. Barndorff-Nielsen and N. Shephard, "Integrated OU processes and non-Gaussian OU-based stochastic volatility models," *Scandinavian Journal of Statistics*, vol. 30, no. 2, pp. 277 – 295, Jun. 2003.

[31] G. Battistelli, L. Chisci, S. Morrocchi, F. Papi, A. Farina, and A. Graziano, "Robust multisensor multitarget tracker with application to passive multistatic radar tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3450 – 3472, Oct. 2012.

[32] G. Battistelli, A. Benavoli, and L. Chisci, "Data-driven communication for state estimation with sensor networks," *Automatica*, vol. 48, no. 5, pp. 926 – 935, 2012.

[33] A. Ben-Tal and A. S. Nemirovskiaei, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2001.

[34] A. Benavoli, L. Chisci, A. Farina, S. Immediata, L. Timmoneri, and G. Zappa, "Knowledge-based system for multi-target tracking in a littoral environment," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 1100 – 1119, 2006.

[35] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.

[36] P. Blackwell, "Random diffusion models for animal movement," *Ecological Modelling*, vol. 100, no. 1, pp. 87 – 102, 1997.

[37] N. A. Bomberger, B. J. Rhodes, M. Seibert, and A. M. Waxman, "Associative learning of vessel motion patterns for maritime situation awareness," *9th International Conference on Information Fusion, Florence, Italy*, pp. 1 – 8, 2006.

[38] E. Bonaccorso, N. Ordonez, D. Pazmiño, A. Hearn, D. Páez-Rosas, S. Cruz, J. Muñoz-Pérez, E. Espinoza, J. Suárez, L. Muñoz-Rosado, A. Vizuete, J. Chaves, M. Torres, W. Bustos, D. Rueda, M. Hirschfeld, and J. Guayasamin, "International fisheries threaten globally endangered sharks in the Eastern Tropical Pacific Ocean: the case of the Fu Yuan Yu Leng 999 reefer vessel seized within the Galápagos Marine Reserve," *Scientific Reports*, vol. 11, Jul. 2021.

[39] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[40] P. Braca, S. Maresca, R. Grasso, K. Bryan, and J. Horstmann, "Maritime surveillance with multiple over-the-horizon HFSW radars: An overview of recent experimentation." *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 12, pp. 4 – 18, Dec. 2015.

[41] P. Braca, P. Willett, K. LePage, S. Marano, and V. Matta, "Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1864 – 1878, 2014.

[42] N. Brax, E. Andonoff, and M.-P. Gleizes, "A self-adaptive multi-agent system for abnormal behavior detection in maritime surveillance," in *KES-AMSTA LNAI. Springer-Verlag Inc.*, 2012, pp. 174 – 185.

[43] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, "Ship surveillance with TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 1092 – 1103, Apr. 2011.

[44] K. P. Burnham and D. R. Anderson, *Model Selection and Multi-model Inference*. New York: Springer-Verlag New York, 2002.

[45] S. Capobianco, L. M. Millefiori, N. Forti, P. Braca, and P. Willett, "Deep learning methods for vessel trajectory prediction based on recurrent neural networks," *IEEE Transactions on Aerospace and Electronic Systems*, 2021. Accepted. [Online]. Available: https://arxiv.org/abs/2101.02486

[46] G. Capraro, A. Farina, H. Griffiths, and M. Wicks, "Knowledge-based radar signal and data processing: a tutorial review," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 18 – 29, Jan. 2006.

[47] J. Carson-Jackson, "Satellite AIS – developing technology or existing capability?" *Journal of Navigation*, vol. 65, no. 2, pp. 303 – 321, 2012.

[48] C. Carthel, S. Coraluppi, and P. Grignan, "Multisensor tracking and fusion for maritime surveillance," in *10th International Conference on Information Fusion*, 2007, pp. 1 – 6.

[49] F. Castaldo, F. A. N. Palmieri, and C. Regazzoni, *Application of Bayesian Techniques to Behavior Analysis in Maritime Environments*. Switzerland: Springer International Publishing, 2015, pp. 175 – 183.

[50] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, Jul. 2009.

[51] S. Chawla, Y. Zheng, and J. Hu, "Inferring the root cause in road traffic anomalies," in *IEEE 12th International Conference on Data Mining*, 2012, pp. 141 – 150.

[52] C. Chen, D. Zhang, P. Castro, N. Li, L. Sun, and S. Li, "Real-time detection of anomalous taxi trajectories from GPS traces," in *Proceedings of MobiQuitous*, vol. 104, 01 2012.

[53] M. Chen, Y. Li, Z. Zhang, C.-H. Hsu, and S. Wang, "Real-time, large-scale duplicate image detection method based on multi-feature fusion," *Journal of Real-Time Image Processing*, vol. 13, no. 3, pp. 557 – 570, Oct. 2017.

[54] B. Choi, *ARMA Model Identification*. New York: Springer-Verlag New York, 1992.

[55] H. L. Christensen, J. Murphy, and S. J. Godsill, "Forecasting high-frequency futures returns using online langevin dynamics," *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 366 – 380, Aug. 2012.

[56] S. Coraluppi and C. Carthel, "Stability and stationarity in target kinematic modeling," in *IEEE Aerospace Conference*, Mar. 2012, pp. 1 – 8.

[57] ——, "Stochastic data association in multi-target filtering," in *Signal and Data Processing of Small Targets 2012*, vol. 8393, International Society for Optics and Photonics. SPIE, 2012, pp. 269 – 275.

[58] S. Coraluppi, C. Carthel, P. Braca, and L. Millefiori, "The mixed Ornstein-Uhlenbeck process and context exploitation in multi-target tracking," in *19th International Conference on Information Fusion*, Jul. 2016, pp. 217 – 224.

[59] S. Coraluppi and C. Carthel, "A hierarchical MHT approach to ESM-radar fusion," in *15th International Conference on Information Fusion*, 2012, pp. 677 – 683.

[60] ——, "Advances in multi-target filtering of evasive targets," *IEEE Aerospace Conference*, pp. 1 – 8, Jun. 2015.

[61] P. Coscia, P. Braca, L. M. Millefiori, F. A. N. Palmieri, and P. Willett, "Multiple Ornstein-Uhlenbeck processes for maritime traffic graph representation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 5, pp. 2158 – 2170, Oct. 2018.

[62] P. Coscia, F. A. N. Palmieri, P. Braca, L. M. Millefiori, and P. Willett, "Unsupervised maritime traffic graph learning with mean-reverting stochastic processes," in *21st International Conference on Information Fusion*, 2018, pp. 1822 – 1828.

[63] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[64] E. d'Afflisio, P. Braca, L. M. Millefiori, and P. Willett, "Detecting anomalous deviations from standard maritime routes using the Ornstein-Uhlenbeck process," *IEEE Transactions on Signal Processing*, vol. 66, no. 24, pp. 6474 – 6487, Dec. 2018.

[65] ——, "Maritime anomaly detection based on mean-reverting stochastic processes applied to a real-world scenario," in *21st International Conference on Information Fusion*, Jul. 2018, pp. 1171 – 1177.

[66] E. d'Afflisio, P. Braca, and P. Willett, "Malicious AIS spoofing and abnormal stealth deviations: A comprehensive statistical framework for maritime anomaly detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2093 – 2108, 2021.

[67] E. d'Afflisio, P. Braca, L. Chisci, G. Battistelli, and P. Willett, "Maritime anomaly detection of malicious data spoofing and stealth deviations from nominal route exploiting heterogeneous sources of information," in *24th International Conference on Information Fusion*, Jun. 2021.

[68] A. Daranda, "A neural network approach to predict marine traffic," *Baltic Journal of Modern Computing*, vol. 4, no. 3, pp. 483 – 495, 2016.

[69] G. K. D. De Vries and M. Van Someren, "Machine learning for vessel trajectories using compression, alignments and domain knowledge," *Expert Systems with Applications*, vol. 39, no. 18, pp. 13 426 – 13 439, 2012.

[70] N. Draper and H. Smith, *Applied regression analysis.* New York: Wiley, 1966.

[71] F. Y. Edgeworth, "On discordant observations," *Philosophical Magazine*, vol. 23, pp. 364 – 375.

[72] A. Einstein, "On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat," *Annalen der Physik*, vol. 17, pp. 549 – 560, 1905.

[73] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *2nd International Conference Knowledge Discovery and Data Mining*, 1996, pp. 226 – 231.

[74] L. Etienne, C. Ray, and G. Mcardle, "Spatio-temporal visualisation of outliers," in *International workshop on Maritime Anomaly Detection (MAD)*, vol. 1, Tilburg, Netherlands, 2011, pp. 1 – 2.

[75] C. Fantacci, B. N. Vo, B. T. Vo, G. Battistelli, and L. Chisci, "Robust fusion for multisensor multiobject tracking," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 640 – 644, May 2018.

[76] *Fishing Operations. 1. Vessel Monitoring Systems. FAO Technical Guidelines for Responsible Fisheries*, FAO (Food and Agriculture Organization) Supplement 1, 1998.

[77] A. Farina and F. Studer, *Radar data processing. Vol. II - Advanced topics and applications*. New York: Wiley, 1986.

[78] M. Fiorini, A. Capata, and D. D. Bloisi, "AIS data visualization for maritime spatial planning (MSP)," *International Journal of e-Navigation and Maritime Economy*, vol. 5, pp. 45 – 60, 2016.

[79] N. Forti, L. M. Millefiori, and P. Braca, "Hybrid Bernoulli filtering for detection and tracking of anomalous path deviations," in *21st International Conference on Information Fusion*, 2018, pp. 1178 – 1184.

[80] ——, "Unsupervised extraction of maritime patterns of life from Automatic Identification System data," in *OCEANS 2019 - Marseille*, Jun. 2019.

[81] N. Forti, L. M. Millefiori, P. Braca, and P. Willett, "Anomaly detection and tracking based on mean-reverting processes with unknown parameters," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 8449 – 8453.

[82] ——, "Random finite set tracking for anomaly detection in the presence of clutter," in *IEEE Radar Conference*, 2020.

[83] N. Forti, G. Battistelli, L. Chisci, S. Li, B. Wang, and B. Sinopoli, "Distributed joint attack detection and secure state estimation," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 96 – 110, 2018.

[84] N. Forti, G. Battistelli, L. Chisci, and B. Sinopoli, "A Bayesian approach to joint attack detection and resilient state estimation," in *IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 1192 – 1198.

[85] ——, "Joint attack detection and secure state estimation of cyber-physical systems," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 11, pp. 4303 – 4330, 2020.

[86] N. Forti, E. d'Afflisio, P. Braca, L. M. Millefiori, P. Willett, and S. Carniel, "Maritime anomaly detection in a real-world scenario: Ever Given grounding in the Suez Canal," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[87] N. Forti, L. Millefiori, P. Braca, and P. Willett, "Prediction of vessel trajectories from AIS data via sequence-to-sequence recurrent neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Feb. 2020, pp. 8936–8940.

[88] N. Forti, L. M. Millefiori, P. Braca, and P. Willett, "Bayesian filtering for dynamic anomaly detection and tracking," *IEEE Transactions on Aerospace and Electronic Systems*, 2021. Accepted.

[89] R. Fujimaki, T. Yairi, and K. Machida, "An approach to spacecraft anomaly detection problem using kernel feature space," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Jan. 2005, pp. 401 – 410.

[90] A. Gad and M. Farooq, "Data fusion architecture for Maritime Surveillance," in *5th International Conference on Information Fusion*, vol. 1, 2002, pp. 448 – 455.

[91] D. Gaglione, P. Braca, and G. Soldi, "Belief propagation based AIS/radar data fusion for multi-target tracking," in *21st International Conference on Information Fusion*, Jul. 2018, pp. 2143 – 2150.

[92] D. Gaglione, G. Soldi, F. Meyer, F. Hlawatsch, P. Braca, A. Farina, and M. Z. Win, "Bayesian information fusion and multitarget tracking for maritime situational awareness," *IET Radar, Sonar & Navigation*, vol. 14, pp. 1845 – 1857, 2020.

[93] C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences*. Berlin: Springer-Verlag, 2004, vol. 13.

[94] D. T. Gillespie, "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral," *Physical Review E*, vol. 54, pp. 2084 – 2091, Aug. 1996.

[95] S. Giompapa, F. Gini, A. Farina, A. Graziano, R. Croci, and R. Distefano, "Maritime border control multisensor system," *IEEE Aerospace and Electronic Systems Magazine*, vol. 24, no. 8, pp. 9 – 15, Aug. 2009.

[96] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.

[97] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[98] A. K. Gostar, T. Rathnayake, R. Tennakoon, A. Bab-Hadiashar, G. Battistelli, L. Chisci, and R. Hoseinnezhad, "Cooperative sensor fusion in centralized sensor networks using cauchy-schwarz divergence," *Signal Processing*, vol. 167, p. 107278, 2020.

[99] K. Granström, A. Natale, P. Braca, G. Ludeno, and F. Serafino, "Gamma Gaussian inverse wishart probability hypothesis density for extended target tracking using X-band marine radar data." *IEEE Geoscience and Remote Sensing Society*, vol. 53, no. 12, pp. 6617 – 6631, Jul. 2015.

[100] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95 – 110, http://stanford.edu/~boyd/graph_dcp.html.

[101] ——, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[102] R. Grasso, P. Braca, J. Osler, J. Hansen, and P. Willett, "Optimal asset network planning for counter piracy operation support, part 1: Under the hood," *IEEE Aerospace and Electronic Systems Magazine*, vol. 29, no. 5, pp. 4 – 11, May 2014.

[103] H. Greidanus and C. Santamaria, *First analyses of Sentinel-1 images for maritime surveillance*. JRC Science and Policy Reports, 2014.

[104] H. Greidanus, "Satellite imaging for maritime surveillance of the European Seas," in *Remote Sensing of the European Seas*. Springer, Dordrecht, 2008.

[105] M. Guerriero, S. Coraluppi, and C. Carthel, "Analysis of AIS intermittency and vessel characterization using a Hidden Markov Model," NURC, Tech. Rep. NURC-FR-2010-002, NATO UNCLASSIFIED, Jan. 2010.

[106] W. Guo, J. Li, G. Chen, Y. Niu, and C. Chen, "A PSO-optimized real-time fault-tolerant task allocation algorithm in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 12, pp. 3236 – 3249, 2015.

[107] D. O. D. Handayani, W. Sediono, and A. Shah, "Anomaly detection in vessel tracking using Support Vector Machines (SVMs)," *International Conference on Advanced Computer Science Applications and Technologies*, pp. 213 – 217, Dec. 2013.

[108] A. Harati-Mokhtari, A. Wall, P. Brooks, and J. Wang, "Automatic Identification System (AIS): A human factors approach," *Journal of Navigation*, vol. 60, no. 3, pp. 373 – 389, Jan. 2008.

[109] G. K. Høye, T. Eriksen, B. J. Meland, and B. T. Narheim, "Space-based AIS for global maritime traffic monitoring," *Acta Astronautica*, vol. 62, no. 2, pp. 240 – 245, 2008.

[110] E. R. Hruschka, E. R. Hruschka, and N. F. F. Ebecken, "Bayesian networks for imputation in classification problems," *Journal of Intelligent Information Systems*, vol. 29, pp. 231 – 252, 2006.

[111] "International Convention of Safety of Life at Sea (SOLAS), 1 November 1974," IMO (International Maritime Organization). [Online]. Available: https://www.imo.org/en/About/Conventions/Pages/International-Convention-for-the-Safety-of-Life-at-Sea-(SOLAS),-1974.aspx

[112] "Revised Performance Standards and Functional Requirements for LRIT," IMO (International Maritime Organization), 2008.

[113] C. Iphar, A. Napoli, and C. Ray, "Detection of false AIS messages for the improvement of maritime situational awareness," in *OCEANS 2015 Washington*, Oct. 2015.

[114] *Technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band*, ITU (International Telecommunications Union) Recommendation ITU-R M.1371, Rev. 5, Feb. 2014.

[115] M. Jakob, O. Vaňek, S. Urban, P. Benda, and M. Pěchouček, "Adversarial modeling and reasoning in the maritime domain," Agent Technology Center, Department of Cybernetics, FEE Czech Technical University in Prague, Tech. Rep., Dec. 2009.

[116] R. L. Kashyap, "Optimal choice of AR and MA parts in autoregressive moving average models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-4, no. 2, pp. 99 – 104, 1982.

[117] F. Katsilieris, P. Braca, and S. Coraluppi, "Detection of malicious AIS position spoofing by exploiting radar information," in *16th International Conference on Information Fusion*, Jul. 2013, pp. 1196 – 1203.

[118] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall PTR, 1993.

[119] ——, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Prentice Hall PTR, 1998.

[120] S. Kazemi, S. Abghari, N. Lavesson, H. Johnson, and P. Ryman, "Open data for anomaly detection in maritime surveillance," *Expert Systems with Applications*, vol. 40, no. 14, pp. 5719 – 5729, 2013.

[121] I. Kononenko and M. Kukar, *Machine Learning and Data Mining: Introduction to Principles and Algorithms*. Horwood Publishing Limited, 2007.

[122] W. Koppe, K. Bach, and P. Lumsdon, "Benefits of TerraSAR-X - PAZ constellation for maritime surveillance," in *10th European Conference on Synthetic Aperture Radar (EUSAR)*, Jun. 2014, pp. 1 – 4.

[123] K. Kowalska and L. Peel, "Maritime anomaly detection using Gaussian process active learning," in *15th International Conference on Information Fusion*, Jul. 2012, pp. 1164 – 1171.

[124] M. Krüger, "Detection of AIS spoofing in fishery scenarios," in *22th International Conference on Information Fusion*, 2019, pp. 1 – 7.

[125] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79 – 86, 1951.

[126] V. Kumar, "Parallel and distributed computing for cybersecurity," *IEEE Distributed Systems Online*, vol. 6, no. 10, 2005.

[127] R. O. Lane, D. A. Nevell, S. D. Hayward, and T. W. Beaney, "Maritime anomaly detection and threat assessment," in *13th International Conference on Information Fusion*, Jul. 2010.

[128] P. Langevin, "Sur la théorie du mouvement brownien," *Comptes-rendus de l'Académie des sciences*, pp. 530 – 533, Aug. 1908.

[129] R. Laxhammar, "Anomaly detection for sea surveillance," in *11th International Conference on Information Fusion*, Jan. 2008, pp. 1 – 8.

[130] R. Laxhammar and G. Falkman, "Online Detection of Anomalous Subtrajectories: A Sliding Window Approach Based on Conformal Anomaly Detection and Local Outlier Factor," in *8th International Conference on Artificial Intelligence Applications and Innovations (AIAI)*, ser. Artificial Intelligence Applications and Innovations, vol. AICT-382, no. Part II. Halkidiki, Greece: Springer, Sep. 2012, pp. 192–202.

[131] R. Laxhammar, G. Falkman, and E. Sviestins, "Anomaly detection in sea traffic - a comparison of the Gaussian Mixture Model and the Kernel Density Estimator," *12th International Conference on Information Fusion*, pp. 756 – 763, 2009.

[132] R. Laxhammar and G. Falkman, "Sequential conformal anomaly detection in trajectories based on Hausdorff distance," in *14th International Conference on Information Fusion*, 2011, pp. 1 – 8.

[133] K. Lee and K. M. Gilleade, "Generic processing of real-time physiological data in the cloud," *International Journal of Big Data Intelligence*, vol. 3, pp. 215 – 227, 2016.

[134] E. Lehmann, *Elements of Large-Sample Theory.* New York: Springer Science & Business Media, 2004.

[135] P.-R. Lei, "A framework for anomaly detection in maritime trajectory behavior," *Knowledge and Information Systems*, vol. 47, pp. 189 – 214, 2015.

[136] P.-R. Lei, T.-H. Tsai, and W.-C. Peng, "Discovering maritime traffic route from AIS network," in *18th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Oct. 2016, pp. 1 – 6.

[137] S. Li, G. Battistelli, L. Chisci, W. Yi, B. Wang, and L. Kong, "Multi-sensor multi-object tracking with different fields-of-view using the LMB filter," in *21st International Conference on Information Fusion*, 2018, pp. 1201 – 1208.

[138] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333 – 1364, Oct. 2003.

[139] ——, "Survey of maneuvering target tracking. Part II. Motion models of ballistic and space targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 96 – 119, Jan. 2010.

[140] B. Liu, E. N. de Souza, S. Matwin, and M. Sydow, "Knowledge-based clustering of ship trajectories using density-based approach," in *IEEE International Conference on Big Data*, 2014, pp. 603 – 608.

[141] S. Liu, L. M. Ni, and R. Krishnan, "Fraud detection from taxis' driving behaviors," *IEEE Transactions on Vehicular Technology*, vol. 63, pp. 464 – 472, 2014.

[142] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xing, "Discovering spatio-temporal causal interactions in traffic data streams," in *17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* New York, NY, USA: Association for Computing Machinery, 2011, pp. 1010 – 1018.

[143] F. Maggi, A. Mooij, and W. Aalst, *Analyzing Vessel Behavior Using Process Mining.* Springer International Publishing, Oct. 2013.

[144] R. Mahler, *Statistical Multisource-Multitarget Information Fusion.* Norwood, MA: Artech House, 2007.

[145] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, "Maritime surveillance using multiple high-frequency surface-wave radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 5056 – 5071, Aug. 2014.

[146] E. Martineau and J. Roy, "Maritime anomaly detection: Domain introduction and review of selected literature," *Technical Memorandum Defence Research and Development, Canada – Valcartier TM 2010-460*, Oct. 2011.

[147] M. Martorella, F. Berizzi, D. Pastina, and P. Lombardo, "Exploitation of Cosmo SkyMed SAR images for maritime traffic surveillance," in *IEEE RadarCon*, 2011, pp. 113 – 117.

[148] M. Martorella, D. Pastina, F. Berizzi, and P. Lombardo, "Spaceborne radar imaging of maritime moving targets with the Cosmo-SkyMed SAR System," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 7, pp. 2797 – 2810, Jul. 2014.

[149] S. Mascaro, A. E. Nicholson, and K. B. Korb, "Anomaly detection in vessel tracks using Bayesian networks," *International Journal of Approximate Reasoning*, vol. 55, no. 1, pp. 84 – 98, 2014.

[150] G. Máttyus, "Near real-time automatic marine vessel detection on optical satellite," in *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. 40, 2013, pp. 233 – 237.

[151] F. Mazzarella, M. Vespe, A. Alessandrini, D. Tarchi, G. Aulicino, and A. Vollero, "A novel anomaly detection approach to identify intentional AIS on-off switching," *Expert Systems with Applications*, vol. 78, Feb. 2017.

[152] F. Meyer, T. Kropfreiter, J. Williams, R. Lau, F. Hlawatsch, P. Braca, and M. Z. Win, "Message passing algorithms for scalable multitarget tracking." *Proceedings of IEEE*, vol. 106, no. 2, pp. 221 – 259, Feb. 2018.

[153] L. M. Millefiori, P. Braca, and G. Arcieri, "Scalable distributed change detection and its application to maritime traffic," in *IEEE International Conference on Big Data*, Dec. 2017, pp. 1650 – 1657.

[154] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2313 – 2330, Oct. 2016.

[155] L. M. Millefiori, P. Braca, and P. Willett, "Consistent estimation of randomly sampled Ornstein-Uhlenbeck process long-run mean for long-term target state prediction." *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1562 – 1566, Nov. 2016.

[156] L. M. Millefiori, P. Braca, D. Zissis, G. Spiliopoulos, S. Marano, P. K. Willett, and S. Carniel, "COVID-19 impact on global maritime mobility," *Nature Scientific Reports*, 2021. Under review.

[157] L. M. Millefiori, G. Pallotta, P. Braca, S. Horn, and K. Bryan, "Validation of the Ornstein-Uhlenbeck route propagation model in the Mediterranean Sea," in *OCEANS 2015 - Genova*, 2015, pp. 1 – 6.

[158] L. M. Millefiori, P. Braca, and G. Arcieri, "Scalable distributed change detection and its application to maritime traffic," *IEEE International Conference on Big Data*, pp. 1650 – 1657, 2017.

[159] N. A. Miller, A. Roan, T. Hochberg, J. Amos, and D. A. Kroodsma, "Identifying global patterns of transshipment behavior," *Frontiers in Marine Science*, vol. 5, 2018.

[160] A. Moreira, G. Krieger, I. Hajnsek, D. Hounam, M. Werner, S. Riegger, and E. Settelmeyer, "TanDEM-X: a TerraSAR-X add-on satellite for single-pass SAR interferometry," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, 2004, pp. 1000 – 1003.

[161] E. S. Nelson, "Maritime terrorism and piracy: Existing and potential threats," *Global Security Studies*, vol. 3, no. 1, 2012.

[162] D. Nevell, "Anomaly detection in white shipping," *Mathematics in Defence*, 2009.

[163] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*. Berlin, Heidelberg: Springer-Verlag, 1992.

[164] E. Osekowska, S. Axelsson, and B. Carlsson, "Potential fields in maritime anomaly detection," in *3rd International IEEE Conference on Models and Technologies for Intelligent Transportation Systems*, 2013.

[165] E. Osekowska, H. Johnson, and B. Carlsson, "Grid size optimization for potential field based maritime anomaly detection," *Transportation research procedia*, vol. 3, pp. 720 – 729, 2014.

[166] ——, "Maritime vessel traffic modeling in the context of concept drift," *Transportation Research Procedia*, vol. 25, pp. 1457 – 1476, Dec. 2017.

[167] R. L. Paes, J. A. Lorenzzetti, and D. F. M. Gherardi, "Ship detection using TerraSAR-X images in the campos Basin (Brazil)," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 3, pp. 545 – 548, 2010.

[168] G. Pallotta, S. Horn, P. Braca, and K. Bryan, "Context-enhanced vessel prediction based on Ornstein-Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results," in *17th International Conference on Information Fusion*, 2014, pp. 1 – 7.

[169] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218 – 2245, 2013.

[170] G. Pallotta and A.-L. Jousselme, "Data-driven detection and context-based classification of maritime anomalies," *18th International Conference on Information Fusion*, pp. 1152 – 1159, 2015.

[171] G. Pallotta, M. Vespe, and K. Bryan, "Traffic knowledge discovery from AIS data," in *16th International Conference on Information Fusion*, 2013, pp. 1996 – 2003.

[172] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, 11 2013, pp. 344 – 353.

[173] L. Pang, S. Chawla, W. Liu, and Y. Zheng, "On mining anomalous patterns in road traffic streams," in *International Conference on Advanced Data Mining and Applications*, vol. 7121, 12 2011, pp. 237 – 251.

[174] ——, "On detection of emerging anomalous traffic patterns using GPS data," *Data & Knowledge Engineering*, vol. 87, pp. 357 – 373, 09 2013.

[175] G. Papa, P. Braca, S. Horn, S. Marano, V. Matta, and P. Willett, "Multisensor adaptive Bayesian tracking under time-varying target detection probability." *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2193 – 2209, Oct. 2016.

[176] J. Pearl, *Bayesian Networks: A Model of Self-activated Memory for Evidential Reasoning*, ser. Report. University of California (Los Angeles). Computer Science Department, 1985.

[177] L. P. Perera, P. Oliveira, and C. Guedes Soares, "Maritime traffic monitoring based on vessel detection, tracking, state estimation, and trajectory prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1188 – 1200, 2012.

[178] W. Pitz and D. Miller, "The TerraSAR-X satellite," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 615 – 622, 2010.

[179] B. Porat and B. Friedlander, "Performance analysis of a class of transient detection algorithms - a unified framework," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2536 – 2546, Oct. 1992.

[180] A. N. Radon, K. Wang, U. Glässer, H. Wehn, and A. Westwell-Roper, "Contextual verification for false alarm reduction in maritime anomaly detection," in *IEEE International Conference on Big Data*, 2015, pp. 1123 – 1133.

[181] C. Ray, R. Dréo, E. Camossi, A.-L. Jousselme, and C. Iphar, "Heterogeneous integrated dataset for Maritime Intelligence, surveillance, and reconnaissance," *Data in Brief*, vol. 25, 2019.

[182] C. Ray, C. Iphar, and A. Napoli, "Methodology for real-time detection of AIS falsification," in *Maritime Knowledge Discovery and Anomaly Detection Workshop*, Ispra, Italy, Jul. 2016, pp. 74 – 77.

[183] B. Rhodes, "Taxonomic knowledge structure discovery from imagery-based data using the neural associative incremental learning (NAIL) algorithm," *Information Fusion*, vol. 8, no. 3, pp. 295 – 315, 07 2007.

[184] B. J. Rhodes, N. A. Bomberger, M. Seibert, and A. M. Waxman, "Maritime situation monitoring and awareness using learning mechanisms," *IEEE Military Communications Conference (MILCOM)*, vol. 1, pp. 646 – 652, 2005.

[185] B. J. Rhodes, N. A. Bomberger, and M. Zandipour, "Probabilistic associative learning of vessel motion patterns at multiple spatial scales for maritime situation awareness," *10th International Conference on Information Fusion*, pp. 1 – 8, 2007.

[186] L. Rikard and F. Göran, "Inductive conformal anomaly detection for sequential detection of anomalous sub-trajectories," *Annals of Mathematics and Artificial Intelligence*, vol. 74, pp. 67 – 94, 2015.

[187] B. Ristic, B. L. Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in AIS data: Anomaly detection and motion prediction," in *11th International Conference on Information Fusion*, Jun. 2008, pp. 1 – 7.

[188] B. Ristic, "Detecting anomalies from a multitarget tracking output," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, pp. 798 – 803, 2014.

[189] M. Riveiro, G. Pallotta, and M. Vespe, "Maritime anomaly detection: A review," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 5, 2018.

[190] E. Ronchetti, "Robust model selection in regression," *Statistics & Probability Letters*, vol. 3, no. 1, pp. 21 – 23, 1985.

[191] J. Roy, "Anomaly detection in the maritime domain," in *Optics and Photonics in Global Homeland Security IV*, vol. 6945, International Society for Optics and Photonics. SPIE, 2008, pp. 180 – 193.

[192] C. Santamaria, M. Alvarez, H. Greidanus, V. Syrris, P. Soille, and P. Argentieri, "Mass processing of Sentinel-1 images for maritime surveillance," *Remote Sensing*, vol. 9, no. 7, p. 678, 2017.

[193] R. E. Schnurr and T. R. Walker, "Marine transportation and energy use," in *Reference Module in Earth Systems and Environmental Sciences*. Elsevier, 2019.

[194] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, no. 2, pp. 461 – 464, Mar. 1978.

[195] H. Y. Shahir, U. Glässer, N. Nalbandyan, and H. Wehn, "Maritime situation analysis: A multi-vessel interaction and anomaly detection framework," *IEEE Joint Intelligence and Security Informatics Conference*, pp. 192 – 199, 2014.

[196] H. Y. Shahir, U. Glässer, A. Y. Shahir, and H. Wehn, "Maritime situation analysis framework: Vessel interaction classification and anomaly detection," *IEEE International Conference on Big Data*, pp. 1279 – 1289, 2015.

[197] A. Sidibé and G. Shu, "Study of automatic anomalous behaviour detection techniques for maritime vessels," *Journal of Navigation*, vol. 70, pp. 1 – 12, Mar. 2017.

[198] S. K. Singh and F. Heymann, "Machine learning-assisted anomaly detection in maritime navigation using AIS data," in *IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2020, pp. 832 – 838.

[199] J. Smith, I. Nouretdinov, R. Craddock, C. R. Offer, and A. Gammerman, "Anomaly detection of trajectories with kernel density estimation by conformal prediction," in *AIAI Workshops, IFIP International Federation for Information Processing*, 2014, pp. 271 – 280.

[200] M. Smith, S. Reece, S. Roberts, I. Psorakis, and I. Rezek, "Maritime abnormality detection using Gaussian processes," *Knowledge and Information Systems*, vol. 38, pp. 717 – 741, 2014.

[201] M. Smith, S. Reece, S. Roberts, and I. Rezek, "Online maritime abnormality detection using Gaussian processes and extreme value theory," Dec. 2012, pp. 645 – 654.

[202] G. Soldi, D. Gaglione, N. Forti, A. Di Simone, F. C. Daffinà, G. Bottini, D. Quattrociocchi, L. M. Millefiori, P. Braca, S. Carniel *et al.*, "Space-based global maritime surveillance. Part I: Satellite technologies," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, Nov. 2020.

[203] ——, "Space-based global maritime surveillance. Part II: Artificial intelligence and data fusion techniques," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, Nov. 2020.

[204] B. H. Soleimani, E. N. De Souza, C. Hilliard, and S. Matwin, "Anomaly detection in maritime data based on geometrical analysis of trajectories," in *18th International Conference on Information Fusion*, 2015, pp. 1100 – 1105.

[205] C. Spence, L. Parra, and P. Sajda, "Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model," in *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*, 2001, pp. 3 – 10.

[206] M. Stasolla, J. J. Mallorqui, G. Margarit, C. Santamaria, and N. Walker, "A comparative study of operational vessel detectors for maritime surveillance using satellite-borne synthetic aperture radar," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 6, pp. 2687 – 2701, Jun. 2016.

[207] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36 – 47, 2004.

[208] L. D. Stone, T. L. Corwin, and C. A. Barlow, *Bayesian Multiple Target Tracking*. USA: Artech House, Inc., 1999.

[209] A. Thorpe and L. Scharf, "Data adaptive rank-shaping methods for solving least squares problems," *IEEE Transactions on Signal Processing*, vol. 43, no. 7, pp. 1591 – 1601, Jul. 1995.

[210] E. Tu, G. Zhang, L. Rachmawati, E. Rajabally, and G.-B. Huang, "Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1559 – 1582, 2018.

[211] H. Tuckwell, F. Wan, and J.-P. Rospars, "A spatial stochastic neuronal model with Ornstein-Uhlenbeck input current," *Biological Cybernetics*, vol. 86, no. 2, pp. 137 – 145, Feb. 2002.

[212] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Physical Review*, vol. 36, pp. 823 – 841, Sep. 1930.

[213] "Long Range Identification and Tracking (LRIT) Overview," U.S. Coast Guard Navigation Center. [Online]. Available: https://www.navcen.uscg.gov/?pageName=lritMain

[214] J. van Laere and M. Nilsson, "Evaluation of a workshop to capture knowledge from subject matter experts in maritime surveillance," in *12th International Conference on Information Fusion*, Jul. 2009, pp. 171 – 178.

[215] A. Vandecasteele, R. Devillers, and A. Napoli, "A semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behavior analysis," in *11th International Symposium for GIS and Computer Cartography for Coastal Zones Management*, 2013.

[216] D. Velotto, C. Bentes, B. Tings, and S. Lehner, "First comparison of Sentinel-1 and TerraSAR-X data in the framework of maritime targets detection: South Italy case," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 4, pp. 993 – 1006, 2016.

[217] J. Venskus, M. Kurmis, A. Andziulis, Z. Lukošius, M. Voznak, and D. Bykovas, "Self-learning adaptive algorithm for maritime traffic abnormal movement detection based on virtual pheromone method," in *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, 2015, pp. 1 – 6.

[218] M. Vespe, I. Visentini, K. Bryan, and P. Braca, "Unsupervised learning of maritime traffic patterns for anomaly detection," in *9th IET Data Fusion Target Tracking Conference: Algorithms Applications*, May 2012, pp. 1 – 5.

[219] G. Vivone, L. M. Millefiori, P. Braca, and P. Willett, "Performance assessment of vessel dynamic models for long-term prediction using heterogeneous

data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 11, pp. 6533 – 6546, Nov. 2017.

[220] G. Vivone, P. Braca, and J. Horstmann, "Knowledge-based ship tracking applied to HF surface wave radar data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2015, pp. 326 – 329.

[221] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*. Springer, 2005.

[222] T. Walker, O. Adebambo, M. Del, A. Feijoo, E. Elhaimer, T. Hossain, S. Johnston Edwards, C. Morrison, J. Romo, N. Sharma, S. Taylor, and S. Zomorodi, "Environmental effects of marine transportation," in *Word Seas: An Environmental Evaluation (Second Edition) Volume III: Ecological Issues and Environmental Impacts*. Elsevier, 2018, ch. 27, pp. 505 – 530.

[223] H. Wang, T. Kirubarajan, and Y. Bar-Shalom, "Precision large scale air traffic surveillance using IMM/assignment estimators," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 1, pp. 255 – 266, Jan. 1999.

[224] X. Wang, X. Liu, B. Liu, E. N. de Souza, and S. Matwin, "Vessel route anomaly detection with Hadoop MapReduce," in *IEEE International Conference on Big Data*, 2014, pp. 25 – 30.

[225] Y. Wang, Z. Zhang, N. Li, F. Hong, H. Fan, and X. Wang, "Maritime surveillance with undersampled SAR," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 8, pp. 1423 – 1427, 2017.

[226] S. Weisberg, *Applied Linear Regression*. New York: Wiley, 1980.

[227] Z. Xiao, X. Fu, L. Zhang, and R. S. M. Goh, "Traffic pattern mining and forecasting technologies in maritime traffic service networks: A comprehensive survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1796 – 1825, 2020.

[228] T. Xu, X. Liu, and X. Yang, "Ship trajectory online prediction based on BP neural network algorithm," in *International Conference of Information Technology, Computer Engineering and Management Sciences*, vol. 1, 2011, pp. 103 – 106.

[229] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338 – 353, 1965.

[230] M. J. Zaki and W. Meira, Jr, *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*. Cambridge University Press, 2020.

[231] L. Zhao and G. Shi, "Maritime anomaly detection using density-based clustering and recurrent neural network," *Journal of Navigation*, vol. 72, no. 4, pp. 894 – 916, 2019.

[232] R. Zhen, J. Yongxing, Q. Hu, Z. Shao, and N. Nikitakos, "Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes classifier," *Journal of Navigation*, vol. 70, pp. 1 – 23, Jan. 2017.

[233] Y. Zheng, "Trajectory data mining: An overview," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, no. 3, 05 2015.

[234] H. Zhong, X. Song, and L. Yang, "Vessel classification from space-based AIS data using random forest," in *5th International Conference on Big Data and Information Analytics (BigDIA)*, 2019, pp. 9 – 12.

[235] W. Zucchini, G. Claeskens, and G. Nguefack-Tsague, *Model Selection*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 830 – 833.