

Artificial Intelligence for Satellite Communication and Non-Terrestrial Networks: A Survey

Gianluca Fontanesi, *Member, IEEE*, Flor Ortíz, *Member, IEEE*, Eva Lagunas, *Senior Member, IEEE*, Victor Monzon Baeza, *Member, IEEE*, Miguel Ángel Vázquez, *Member, IEEE*, Juan Andrés Vázquez-Peralvo, *Member, IEEE*, Mario Minardi, *Student Member, IEEE*, Ha Nguyen Vu, *Member, IEEE*, Puneeth Jubba Honnaiah, Clement Lacoste, *Student Member, IEEE*, Youssouf Drif, *Member, IEEE*, Liz Martinez Marrero, *Student Member, IEEE*, Saed Daoud, *Member, IEEE*, Tedros Salih Abdu, *Member, IEEE*, Geoffrey Eappen, *Member, IEEE*, Junaid ur Rehman, Luis Manuel Garcés-Socorrás, *Member, IEEE*, Wallace Alves Martins *Senior Member, IEEE*, Pol Henarejos, *Member, IEEE*, Hayder Al-Hraishawi *Senior Member, IEEE*, Juan Carlos Merlano Duncan, *Senior Member, IEEE*, Thang X. Vu, *Senior Member, IEEE* and Symeon Chatzinotas, *Fellow, IEEE*

Abstract—This paper surveys the application and development of Artificial Intelligence (AI) in Satellite Communication (Sat-Com) and Non-Terrestrial Networks (NTN). We first present a comprehensive list of use cases, the relative challenges and the main AI tools capable of addressing those challenges. For each use case, we present the main motivation, a system description, the available non-AI solutions and the potential benefits and available works using AI. We also discuss the pros and cons of an on-board and on-ground AI-based architecture, and we revise the current commercial and research activities relevant to this topic. Next, we describe the state-of-the-art hardware solutions for developing ML in real satellite systems. Finally, we discuss the long-term developments of AI in the SatCom and NTN sectors and potential research directions. This paper provides a comprehensive and up-to-date overview of the opportunities and challenges offered by AI to improve the performance and efficiency of NTNs.

Index Terms—Satellite Communication, Machine Learning, Artificial Intelligence, Non-terrestrial Networks

I. INTRODUCTION

Satellite Communication and Non-Terrestrial Networks are seeing unbounded growth, offering a new world of enthralling possibilities for global coverage, connectivity, and scalability. Space-based communication systems have been typically used to cover vast airspace and sea areas, supplementing existing

terrestrial networks for global connectivity. With the increasing demand for broadband connectivity, non-terrestrial operators are investing in new orbits and new constellation designs to provide a cost-efficient response to such a market. In addition, Non-Terrestrial Networks technology has been identified as a key component of the upcoming 5G and beyond cellular communication, particularly for backhaul services, backup connectivity, and extending coverage in remote or isolated regions of the globe [1].

While the term Non-Terrestrial Networks includes a huge variety of space-borne and aerial communication networks, such as Geo-Stationary Orbit, Medium Earth Orbit, Low Earth Orbit satellite constellations, High Altitude Platforms systems, Low Altitude Platforms systems, and Air to Ground networks, for the sake of space and clarity, this survey has focused on satellite communications systems limited to Geo-Stationary Orbit, Medium Earth Orbit, Low Earth Orbit.

A. Need for AI/ML in SATCOM

The Satellite Communication ecosystem is currently experiencing a revolution in advanced technological solutions and raising a new space era where lower orbits are emerging as a low-latency alternative to conventional Geo-Stationary Orbit communication systems. The time and geographical dependency of user demands combined with the dynamic movement of low-orbiting satellites have resulted in new antenna architectures with beamforming capabilities and reconfigurable payloads with the capability to adapt the payload configuration in response to traffic needs.

This new communication era brings fundamental operational challenges. Most deployed Satellite Communication systems largely depend on human expertise, and manual intervention [2]. This has two main drawbacks. First, human involvement in system control activity leads to high Operational Expenditure and latency [3]. Second, the rapidly changing radio environment of the new space scenarios claims for autonomously adaptive mechanisms that human intervention cannot offer. Lastly, the multitude of use cases and scenarios served by satellite communication in the following years will produce data in large quantities. Thus, it is beneficial

Manuscript received XXX, XX, 2021; revised XXX, XX, 2021.

The authors from University of Luxembourg have been supported by the Luxembourg National Research Fund (FNR) under the project SmartSpace (C21/IS/16193290). The authors from CTTC have been supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004215 (ATRIA) and by the Spanish ministry of science and innovation under project IRENE (PID2020-115323RB-C31/AEI/10.13039/501100011033) and grant from the Spanish ministry of economic affairs and digital transformation and of the European union – NextGenerationEU [UNICO-5GI+D/AROMA3D-Space (TSI-063000-2021-70).

G. Fontanesi, F. Ortíz, E. Lagunas, V. Monzon Baeza, J. A. Vázquez-Peralvo, M. Minardi, H. N. Vu, P. J. Honnaiah, C. Lacoste, Y. Drif, T. S. Abdu, G. Eappen, J. Rehman, L. M. Garcés-Socorrás, W. A. Martins, H. Al-Hraishawi, J. C. Merlano Duncan, T. X. Vu and S. Chatzinotas are with the Interdisciplinary Centre for Security, Reliability, and Trust (SnT), Luxembourg, L-1855 Luxembourg

M. Á. Vázquez and P. Henarejos are with the Centre Tecnologic de les Telecomunicacions de Catalunya.

(Corresponding author: Gianluca Fontanesi, gianluca.fontanesi@nokia.com)

and necessary to make satellites capable of automatically generating reliable actions using the data produced.

Artificial Intelligence is entering satellite communication with the great promise to solve the abovementioned challenges. Intelligent systems enable automation to process specific input from satellite data and translate them into actions. In addition, data-driven techniques might substantially reduce the Operational Expenditure.

Given the significant momentum and recent activities promoting Artificial Intelligence in satellite communication, it is timely to survey this brand-new field.

Certain literature agrees to classify Machine Learning as a branch of Artificial Intelligence, whose goal is to allow computers to learn from data. In some way, Machine Learning is just one way to achieve Artificial Intelligence. However, most of the time, both terms are used interchangeably, such as the idea of a machine mimicking human intelligence. Therefore, in Section II, we address this dilemma for a better understanding of the reader.

B. Related Surveys and Contribution

Machine Learning has appeared as a promising alternative for dealing with computationally expensive optimization procedures. For this reason, and with the exponential increase in the number of datasets available, ML has become a fundamental technology in different areas of wireless communications [4]. In this context, ML has proven to be an exciting tool for accelerating complex optimization procedures for general wireless communications [5]. Many surveys cover the development of Machine Learning models to support wireless networks.

About the direct application of Machine Learning models in Satellite Communication, there are only a handful of survey papers summarising it. For this reason, in the following subsections, we will present the surveys describing SATCOM and the application of Artificial Intelligence.

C. SATCOM Related Surveys

Due to the increasing importance and interest in satellite communication, several works survey different aspects of satellite communication. The work in [6] deals with the vision of Satellite Communication in conjunction with Unmanned Aerial Vehicles and terrestrial networks to provide a feasible and cost-effective solution for continuous and ubiquitous wireless coverage. The innovation and transformation phase experienced by Satellite Communication, driven by on-board processing capabilities, non-terrestrial networks and space-based data collection/processing, is explained in [7]. The algorithms and applications of Earth Observation are covered in [8]. Cubesats are a special class of miniaturized satellites commonly used as teaching tools, and technology demonstration [9]. The work in [10] describes all the potential future applications of Cubesats for Satellite Communication.

Other papers survey a variety of Satellite Communication aspects, such as handover schemes [11], non Geostationary satellite systems [12], synchronization techniques [13], MIMO techniques [14].

The important case of Satellite Communication as Internet of Things enabler is treated in several surveys [7], [15], [16]. The works [7], [15] categorize the use cases in which satellites serve IoT devices used in global transportation and agriculture applications. In local-area services, a satellite serves a specific set of IoT devices in applications such as smart grid systems.

D. SATCOM-AI Related Survey

While great pieces of work, the above papers neglect to address the important challenge of the application of Machine Learning in the described use cases. While not intended as a tutorial or survey papers, [2], [17] introduce the application of Artificial Intelligence techniques for integrated terrestrial satellite networks, specifically focusing on mega satellite network communications [17] and Satellite Communication operations which are strongly dependent on the human intervention. The work in [18] provides a general overview of Artificial Intelligence, its subfields and algorithms and the potential Artificial Intelligence-based solutions for some satellite challenges. Other works address Deep Learning only [19], [20] or Artificial Intelligence for space missions [21], without however addressing the challenges of onboard or on-ground computing, hardware constraints and limitations.

E. Scope and Contribution

This work complements the above works by surveying all possible Satellite Communication use cases and discussing the practical and implementation aspects of the application of Artificial Intelligence techniques in Satellite Communication. A detailed survey of this type was still missing, and it is of great importance to take a new look at the appealing Artificial Intelligence/Machine Learning approaches for Satellite Communication with an implementation perspective. The contributions of this paper can be summarized as follows:

- We present an extensive classification of the most important use cases in Satellite Communication and Non-Terrestrial Networks, and we group them based on an OSI-oriented architecture.
- We describe each presented use case, the main related challenges, and the conventional tools used to address them. We further present the main Machine Learning techniques applied to each use case, and we classify them by whether the training is onboard or on ground.
- We assess the hardware landscape for current and next-generation processors, presenting a comprehensive review of commercial AI chipsets and comparing their specifications in terms of Satellite Communication applicability.
- We analyze the main challenges for the Satellite Communication and Non-Terrestrial Networks from the perspective of future network deployment. We also present novel and futuristic trends.

This paper consists of six main sections. Specifically, we first briefly introduce the main Artificial Intelligence/Machine Learning learning frameworks. Our survey does not describe how various ML methods operate. There are numerous books and research papers on this topic; we refer the reader to

TABLE I
RELEVANT SURVEYS AND TUTORIALS ON ARTIFICIAL INTELLIGENCE (AI)/MACHINE LEARNING (ML) AND/OR SATELLITE COMMUNICATIONS.

Network	Source	Main Scope	ML Focused	Implementation Oriented
SATCOM	[1], [6], [24]	Survey non-terrestrial networks	No	No
	[7], [25]	Next Generation Satellite Communications	No	No
	[8]	Earth Observation	No	No
	[10]	Cubesat Applications	No	No
	[11]	Handover schemes	No	No
	[12]	non-GEO satellite comm.	No	No
	[13]	Distributed Satellite Systems	No	No
	[14]	MIMO Techniques for Satellite	No	No
	[15]	Satellite Comm in Internet of Things	No	No
	[16]	Satellite Internet of Things	No	No
	[21]	Space mission	Yes	No
ML & SATCOM	[18]	Review of ML techniques for satellite comm.	Yes	No
	[2]	ML for SATCOM operation centers	Yes	No
	[17]	Survey on AI for Mega Constellation	Yes	No
	[19]	Deep Learning in Space	Yes	No
	[20]	Deep Learning in Space	Yes	No
	[26]	ML for Radio Resource Management	Yes	No
	This work	ML for Satellite Communication	Yes	Yes

papers such as [22], [23], for a detailed (although still wireless networking-related) discussion of these methods. We then discuss in Section III the role of Artificial Intelligence in Satellite Communication, highlighting the main differences between an onboard or on-ground implementation. Section IV presents an extensive list of use cases in Satellite Communication and Non-Terrestrial Network. For each use case, we show the benefits of using ML-based techniques. Section V presents an analysis of the HW solutions for AI/ML implementations for Satellite Communication. Last, Section VI on the challenges and future directions of Satellite Communication and Non-Terrestrial Networks concludes the paper.

F. Methodology

The following survey of Satellite Communication is the result of a systematic literature review methodology. First, we identified all the possible use-cases and Satellite Communication research fields and their respective challenges. For each of them, we searched for conventional and Machine Learning solutions in IEEE, MDPI, Elsevier and Wiley libraries. A systematic search was implemented to identify the most important related works. The research is restricted to journal and conference papers only. A second round of research has been computed via manual research and personal knowledge. Table III, to ease the reader understanding, provides a catchy and clear overview of Satellite Communication use cases and related Machine Learning papers.

Additionally, we searched for all relevant surveys. Table I summarizes the past surveys on Satellite Communication, allowing the reader to quickly grasp the primary focus of each of the previously conducted surveys.

ACRONYMS

II. MACHINE LEARNING OVERVIEW

A. Artificial Intelligence, Machine Learning and Deep Learning

A multidisciplinary science that develops theories, techniques, and methods to simulate and extend human intelligence

is Artificial Intelligence [27]. Artificial Intelligence attempts to understand the essence of intelligence and to simulate the human brain's information processing by machines. Machine Learning, a branch of Artificial Intelligence, is related to computational statistics and predictions by exploiting experience and knowledge gained from data [28]. Machine Learning involves learning from data and making decisions or predictions. Essentially, it is based on the assumption that machines can have intelligence that allows them to learn from previous computations and adapt to the environment.

A typical Machine Learning framework includes a training and testing process. Training enables the ML framework to discover the relationships between input and output data. While a tutorial on how to create a dataset and divide it between training and testing is outside the scope of this paper, we highlight that there are two phases of ML techniques for assisting SATCOM at any stage of network communication: training and inference [26]. First, the ML model undergoes the training phase. The goal is to find the optimal model parameters to assist the SATCOM task for the system conditions. Next, a trained model is obtained and used to indicate the best decision in the system.

Available ML models in literature can be classified into classification, regression, and structured learning models [29]. A classification model is used to solve binary or multiple classification problems, and a regression model can be used to make predictions. The structured learning model is widely used in many fields, such as natural language processing.

Deep Learning is essentially a branch of ML, which allows a model to make classifications, predictions, or decisions based on large data sets, without being explicitly programmed. Moreover, ML can be also classified from the training method into supervised, unsupervised, and Reinforcement Learning [29]. The classification of the main ML techniques and the intersections between them are shown in Fig. 1.

Opposite to conventional ML algorithms based on predefined features, DL can extract essential features from the raw data through multiple layers of nonlinear processing units to make predictions or perform target-based actions. DL's

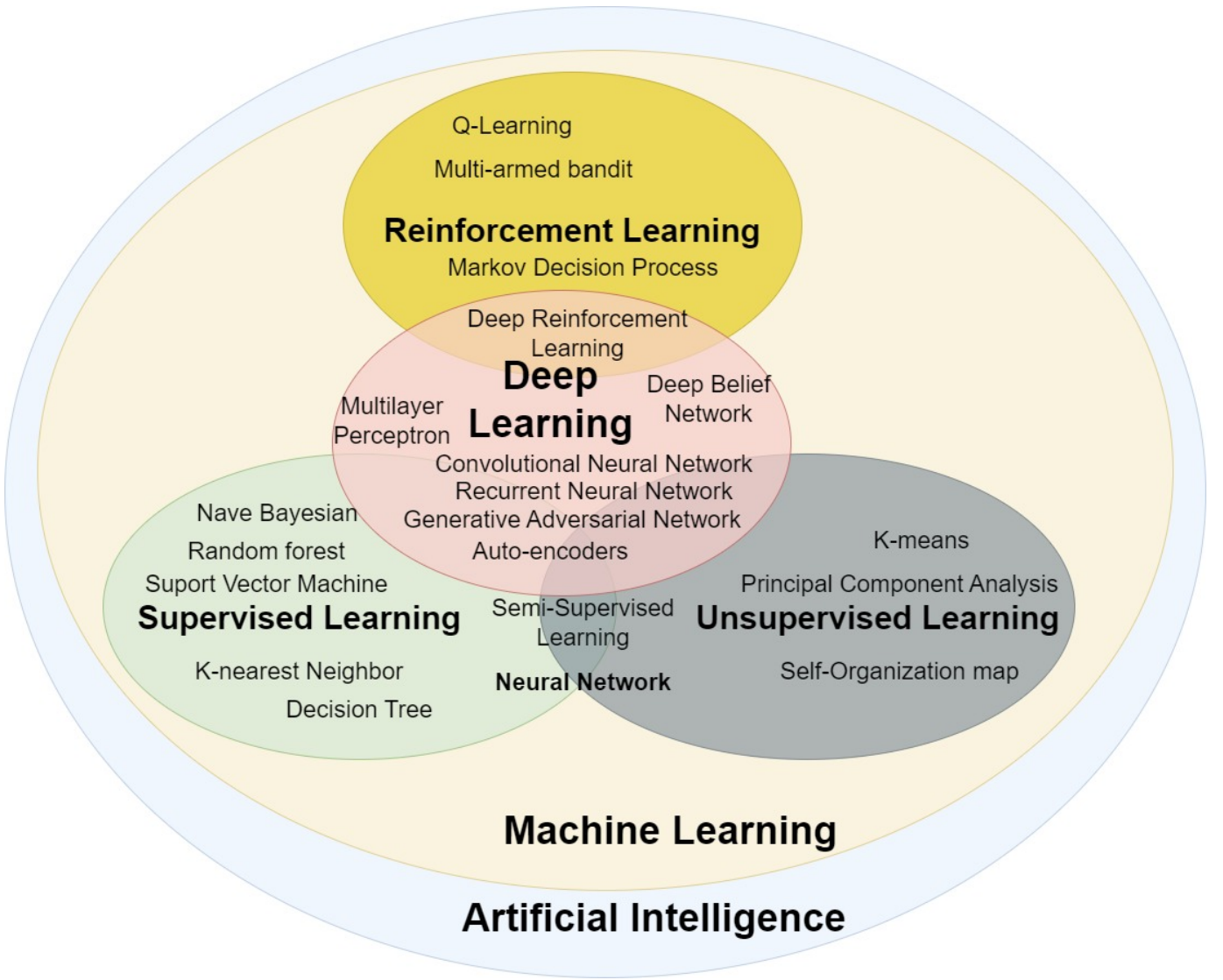


Fig. 1. Artificial Intelligence, Machine Learning, and Deep Learning relationship, as well as their main methods.

main advantages are to extract features automatically. The best-known DL models are neural networks with a sufficient number of hidden layers. Even though the multilayer neural network was proposed several decades ago, unprecedented interest has only recently arisen due to the advancement of backpropagation-based training as well as the success of GPU. Deep Neural Networks aims to approximate any complex function by a composition of simple operations on neurons. Deep Neural Networks can automatically extract essential features from input data with complex structures. No human-designed learning process is needed, significantly reducing the effort in feature elaboration. DL may learn meaningful patterns from unlabeled data in an unsupervised manner. A key challenge of employing DL in mobile communication systems lies in designing optimal deep neural networks for different scenarios just so that the model can be effectively trained in the offline stage and achieve good performance in the online testing stage. Representative typical DL models include multilayer perceptron (MLP), deep belief network (DBN),

autoencoders, Convolutional Neural Network, recurrent neural network (RNN), and generative adversarial network (GAN) [30]–[32].

B. Learning Frameworks

Machine Learning learning approaches can be grouped into supervised learning, unsupervised learning, and reinforcement learning (RL).

1) *Supervised Learning*: A supervised learning approach needs a supervisor to label the input and output data. During the training step, a learning algorithm is provided with labeled training data containing known input and output data to train a model representing the relationships between the input and output data. A novel test data set is introduced into the learned model during the testing step to obtain the expected output. Usually, supervised learning is employed in scenarios with sufficient labeled data. Many fields, such as object recognition, speech recognition, and spam detection, have used this method extensively. Algorithms typical of this category are Naive

Bayes, K-nearest neighbor (KNN), random forest, Neural Networks), Support Vector Machine, and Decision Tree [33], [34].

2) *Unsupervised Learning*: Unsupervised learning algorithms use a set of unlabeled inputs to infer the output correctly. These techniques are frequently used for clustering and aggregation. Some typical unsupervised algorithms include K-mean, self-organizing maps (SOM), hidden Markov model (HMM), and restricted Boltzmann machine (RBM) [35]–[37].

3) *Semi-supervised Learning*: Semi-supervised learning involves a combination of supervised learning and unsupervised learning. A large amount of labeled data is expensive and time-consuming to generate in many real-world scenarios, while it is relatively inexpensive and convenient to collect enough unlabeled data. Consequently, semi-supervised learning develops to fully use the unlabeled data to improve the performance of the trained model. Descriptive pseudo labeling, a typical example of the semi-supervised learning algorithm, is simple and efficient. The constrained labeled data are used to train a model. Next, the trained model generates the pseudo labels for the unlabeled data. And finally, the labeled data and the pseudo-labeled data are used to retrain the model. To fully use the unlabeled data, semi-supervised learning algorithms usually require certain assumptions on the dataset, such as manifold assumption, low-density assumption, clustering assumption, and smoothness assumption. For example, expectation maximization is based on the clustering assumption, while transductive SVM requires the low-density separation assumption [38]–[40].

4) *Reinforcement Learning*: Reinforcement learning (RL) mimics the brain's trial-and-error learning process [41]. As opposed to learning the structure of the training data set, RL attempts to explore the best actions during a dynamic process. Being able to understand the environment through actions and feedback makes it suitable for solving decision-making problems. RL framework can be classified into two types: model-based and model-free. The framework of model-based RL includes an agent, a state space, and an action space. Interacting with the environment, the agent tries to represent the model of the environment and learn the best action to maximize its long-term reward, which is a cumulative discounted reward and is related to current and future rewards. On each step, the agent monitors a state and takes action from the action space, then receives an immediate reward that indicates the effect of the action, followed by the system moving to another state. In model-based approaches, the agent attempts to learn a policy.

In the state transition process, the agent learns the best policy, mapping from the state to the action space to maximize the long-term reward. When determining the long-term reward of the action in a state space, the value function is applied. A value function most commonly used is the Q function, used by the Q-learning method to learn a table for maintaining state-action pairs and associating long-term rewards. RL based on models is more suitable for dynamic systems due to the difficulty of constructing an accurate model of dynamic networks [42]. Compared to other learning techniques, the main advantage of reinforcement learning is that it does not

rely on an accurate mathematical model of the environment.

Furthermore, the approach addresses long-term rewards, including immediate and future rewards, which allows for long-term optimization results. The key challenge in applying RL to wireless communication systems is designing the system state, action, and reward in different scenarios to achieve optimal performance. During recent years, RL has been widely employed to solve decision-making problems in wireless communications, such as user scheduling, spectrum sharing, or radio access technology handover [43]–[46]. Nevertheless, the technique faces some challenges when dealing with problems with a large state or action space since it is challenging to model each state-action pair directly. Consequently, Reinforcement Learning is hardly used in practice [46].

III. MACHINE-LEARNING ASSISTED SATELLITE SYSTEM

A. Introduction

The satellite communications world is also progressing towards having all the ingredients making ML suitable for particular satellite-related use-cases. First, in response to the increasing telecommunication's market demand for flexibility, satellite payloads are being digitalized. In particular, the digital processing of the uplink signals combined with digital routing and filtering flexibility can be highly efficient to address the emerging broadband market [47].

Full digital payloads are paving the way for satellites with a higher level of flexibility to efficiently support peaks in demand. These payloads include capabilities like onboard channelization, interference mitigation techniques, digital beamforming and/or time flexibility. In principle, the upcoming satellite systems should enable to dynamic allocate much higher capacity to designated areas, while other areas do not require as much bandwidth and are just being scanned to gauge demand. For most satellite operators, automatic dynamic resource allocation algorithms are the key innovation aspect to unlocking the value of the aforementioned technological investments. Therefore, to unveil the full potential of such new added flexibility, ML can help accelerate the complex procedures typically encountered in system optimization and minimize the human decision-making over such procedures. Furthermore, ML is known to be a great tool for network load prediction, which is also a key aspect when deciding on when and how to reconfigure the satellite payload.

A side from the flexibility of next generation satellite systems, satellite payloads are known by its nonlinear effects, which cause significant headaches to engineers in order to countermeasure their impact on the final performance. Non-linear effects and general impairments introduced by hardware components are part of the channel information, which is typically estimated by the on-ground receiver by measuring pilot signals. Although different models for such impairments exist, such models may not accurately present the actual signal. To deal with such complex end-to-end system model, ML can be exploited to either predict the true model directly from experience or to estimate the parameters to fine-tune the already available model.

Focusing on the ground segment, which has been typically represented by few gateways (GWs) for GEO systems, it is

currently evolving towards a multi GW environment. In the case of GEO, this is because the congestion of the Ka band is pushing feeder links to operate in higher bands, e.g. Q/V bands. Given the strong impact of weather impairments on higher frequency bands, the multi-location diversity achieved by multiple distributed GW becomes an attractive solution. In the case of NGSO, the need of multi-GW is dictated by the movement of the satellites and the need to have often a connection to Earth for the uplink / downlink of the signals. The multi-GW environment resembles the promising Centralized Radio Access Network (C-RAN) architecture of cellular systems [48], where the control plane is logically centralized gathering the whole system intelligence providing a global view of the entire system status and enabling cross-layer system-wide optimization. In this context, machine-learning-based automation applied at the central node can autonomously optimize the network while reducing the critical processing time for decision-making.

Embracing cloud-based architectures, satellite operators look forward to managing and coordinating the different parts of the network in a centralized mode. The benefit for operators is the availability of very large data sets from their networks and the possibility to apply a range of data processing, geo-spatial, ML and other tools to manipulate and analyse such data, with the main goal to optimize the system operation.

Last but not least, the emergence of piggy-back launches, nano and microsat technology and new higher-risk accepting development approaches and the availability of venture capital, new innovative mission concepts have been envisaged and implemented. This trend of having space-based mesh network of small satellites opens up a plethora of opportunities for massive self-organized, reconfigurable and resilient NGSO satellite constellations, which can operate as a global network instead of a single relay.

To exploit NGSO satellite signals, frequency and time misalignment caused by the fast system mobility shall be compensated. For this, accurate estimation and prediction of ephemeris is a key aspect and has been the subject of extensive research, e.g. [49], [50]. On the other hand, the distributed nature of NGSO constellations translates into highly dynamic and overlapped coverage areas on ground, which bring many operation challenges, e.g. the dynamic handover of users from satellite to satellite, the edge-processing in space, the cooperative transmission from multiple satellites, etc. Each of those relatively new challenges can have potential benefits on the application of ML, as we will see in the next parts of this paper. Before that, we present in the following an overview of the conventional satellite architecture and how ML can be integrated into such systems.

B. Communications Architecture

Satellite network architecture comprises space, ground, control, management, and user segments, as shown in Figure 2. (i) Space segment comprises the satellites organized in the constellation. It supports routing, adaptive access control, and spot beam management. (ii) The ground segment consists of satellite gateways (GWs) interconnected by optical backbones

and satellite terminals (STs) that provide connections for end-user devices. The GWs and STs are interconnected through the space segment. (iii) The control and management segment consists of the network operations centers (NOCs) of the ground segment and the payload operations centers (SOCs) of the space segment. The NOCs and SOCs provide real-time management and control functions for the satellite networks. They perform connection establishment, tracking and release, admission control, resource allocation, satellite network element configuration, security, fault, and performance management. Co-located GWs, NOCs, and SOCs are commonly referred to as satellite hubs. (iv) The user segment comprises all end-user devices used by end-users to consume satellite-based services, fixed or mobile. They access satellite networks directly or through terrestrial access points [51].

C. ML for SATCOM Dilemma: Onboard or Onground

One of the main dilemmas from the space segment architecture perspective is whether to perform operations onboard or on-ground. Onboard strategies increase the complexity and mass/power of the satellite, which has to be ultra-reliable and robust since there is very little chance of repairing/replacement after the asset is put in orbit [7]. Thus, due to the limited satellite process capabilities, onboard applications are challenging in Satellite Communication. On the contrary, on-ground techniques alleviate the onboard complexity. Recent advances in the efficiency of power generation and the energy efficiency of radio frequency and digital processing components enhanced onboard processing, which can enable innovative communication technologies, such as flexible routing, Radio Resource Management, and beamforming.

As an example of the onboard/on-ground dilemma for the latest mentioned case, [52] classifies beamforming methods into onboard and on-ground. While no perfect solution exists, different satellite playmakers have opted for different techniques, onboard (e.g. Inmarsat4, Alphasat or Thuraya) or on-ground techniques, such as US satellite operators.

The same dilemma persists and is emphasized in the application of Machine Learning in Satellite Communication, where Satellites' ability to capture vast amounts of data from space represents an excellent opportunity to apply ML.

From the onboard/on-ground perspective, Machine Learning techniques for Satellite Communication can be first categorized depending on whether the Artificial Intelligence chipset is onboard the satellite, or at the ground station. AI algorithms can be very computationally intensive, consume significant power, and are slow when running on standard processors. Thus, running ML models on-ground would be preferable to avoid high payload complexity and reduce cost. However, this would represent a longer delay if the inference waits for the return channel to decide based on the new conditions.

The industry has developed AI-specific processors capable of executing complex algorithms in real-time and consuming a fraction of the power to enable the practical use of AI for satellite applications in terrestrial and onboard scenarios. These processors are now available as commercial AI chipsets and will be presented in Section V.

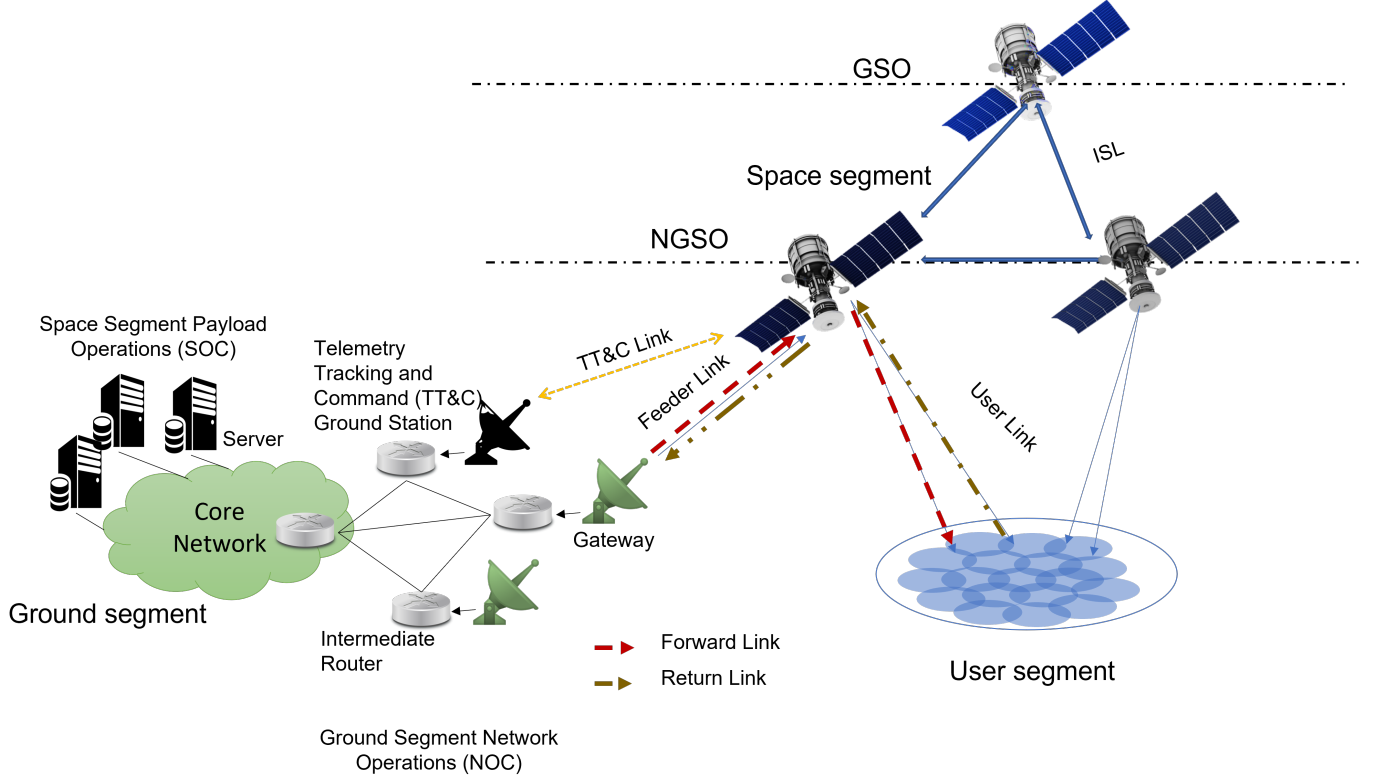


Fig. 2. Communications Architecture

In addition, any AI techniques for Satellite Communication must go through a training, where the model parameters are computed, and an inference phase. Thus, an essential decision is where to perform the algorithm's training. More specifically, the term online refers to performing the training phase when the Satellite Communication system is active regardless of whether the Machine Learning hardware can be on board satellites in a space environment or on the ground. However, due to the computational complexity, it is often performed on the ground to avoid the limitations of power, mass and cost of the payload when training is performed online. Still, thanks to the optimized hardware explained in section V, it is hoped that it can help mitigate this challenge to make onboard training more feasible. In contrast, offline training of the ML algorithm is done when the system is not active in the ground segment with a training database and large servers with powerful GPUs. Nevertheless, the AI chipset could go on board the satellite using the model for inference [3].

Based on these criteria, in this survey we categorize Machine Learning techniques in Satellite Communication as offline or online Machine Learning. In what follows, we are interested in showing each configuration's main benefits and challenges.

The main advantage of online training is that the system will have greater flexibility to adapt to unforeseen network changes, e.g., a drastic and unexpected change in traffic demand, because the learning is constantly updated. Additionally, onboard methods avoid downloading data to servers on the ground,

leading to lower latency. Consequently, the data to train the model can be collected and the training can be performed without incurring in outages and network disconnections. Last, reducing the need for server communication, the data collected and used for Machine Learning training, that often captures sensitive information, is less exposed to privacy issues and the risk to be leaked.

However, while onboard training provides several benefits, implementing such training can be challenging due to the relatively high convergence length and the limited computing capacity onboard [53]. Consequently, Machine Learning models in an onboard setting need to be efficient and often have to trade off against processing time, memory, number of computational operations and energy consumption.

In offline training, once the model has been trained, the Artificial Intelligence chipset could be located onboard the satellite for inference. Consequently, computing/processing times are reduced since inference on a trained ML model requires a low computational cost compared to the training phase. The ML model might thus act as an intelligent switch that selects the configuration based on the current system requirements. The main advantage of this architecture is the reduced processing time. However, this architecture strongly depends on the training data and models used. An extensive database may be necessary to obtain successful training that considers multiple parameters, and the model may present significant errors in unexpected situations in the system.

Recent advances to solve the on-board/off-board dilemma

TABLE II
BENEFITS AND CHALLENGES FOR AN ONLINE OR OFFLINE ML TRAINING IN SATCOM

ML Training in SatCom	Benefits	Challenges
<i>Online</i>	Lower Latency Improved Autonomy Flexibility to adapt to unforeseen changes Higher Privacy	Limited Computing Capacity on-board Longer Processing Time Limited Power Radiation Tolerance
<i>Offline</i>	Reduced processing time Large computational resources	Not flexible to sudden network' changes Inaccurate in unexpected situations Higher Latency

includes hybrid architectures, where the model training is initially performed online/offline and then updated by collecting new data [26]. Therefore, this trade-off strongly impacts the AI Chipset implementation in the system. To ensure successful deployability, designing lightweights Neural Network, which might arrive to have hundreds of parameters and layers, is an essential step in deploying resource-constrained computing platforms. Architectures as MobileNets, ShuffleNet, PeleeNet, and techniques as Model compression, Neural Architecture Search are indicated to reach this goal [54]. Another interesting solution is to run the Machine Learning on-board, but adjusting the Machine Learning algorithm to the satellite limited capacity [55]. In this case, due to the lower compute and lower memory resources onboard, the ML models deployed on the satellite are designed for compute and memory efficiency. Typical techniques to reach increase efficiency are compressing high-performing pre-trained models or designing ad-hoc and application-oriented architectures that are parameter and compute-efficient. Alternatively, hardware acceleration techniques applied in co-design with software can be used to exploit the model compression algorithms at a hardware level. The reader will find in III the indication whether the literature ML models are run online or offline. We will discuss the more advanced solutions and visions to solve the onboard/on-ground dilemma in Section VI.

D. Ongoing Activities

This section will provide a brief overview of the most relevant activities carried out by private and public institutions worldwide on the development of ML in Satellite Communication, focusing on the most advanced technology readiness level.

1) *ML Testbeds*: MIRSAT testbed [56] for Non-Geostationary Orbit (NGSO) constellations proposes an experimentation software platform based on Software Defined Network for validation of the new autonomous network-slicing algorithms and their performance evaluation in integrated terrestrial-satellite systems. The laboratory testbed¹ has been developed and validated [57], consisting of a mininet-based simulator, a Ryu SDN controller with an End-to-End (E2E) Traffic Engineering (TE) application for the Virtual Networks (VNs) establishment, a satellite simulator STK, a real traffic generator OSTINATO and the VNE algorithm implemented in Matlab (Fig. 3).

Late 2020, SPIRE company (with support from the ESA's Earth Observation Science for Society Programme) launched

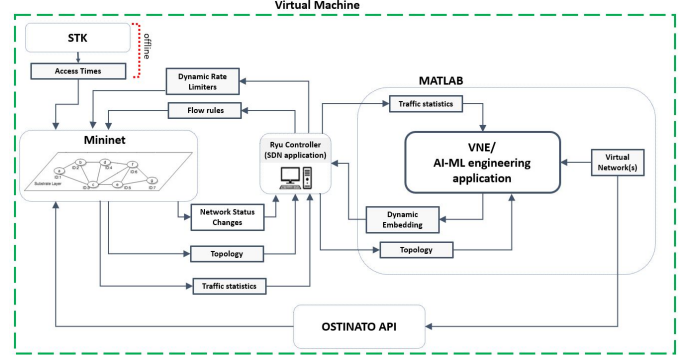


Fig. 3. Laboratory Testbed

the so-called Brain in Space project². Based on SPIRE's LEMUR 3U platform, Brain in Space provides an on-the-ground simulated testbed including multiple new embedded ML modules for users to schedule, upload and test their ML-powered applications.

On the other hand, INTEL has funded the Intel Neuromorphic Research Community, which includes academic, government, and corporate groups worldwide engaged in advancing the science and application of ML-based neuromorphic computing using Intel's Loihi neuromorphic processor. A discussion on neuromorphic computing is provided in Section VI of this survey. Herein, we highlight the willingness of industrial partners to make the technology available for testing and a better understanding of the ML challenges in the upcoming future.

2) *ESA*: The European Space Agency opened an AI-related call for SATCOMs for the first time in 2019 to investigate the applicability of AI techniques in satellite communications. Several potential use cases were shortlisted during these projects. A preliminary evaluation of a small number of them was carried out to provide guidelines for future research; (i) SATAI - Machine Learning and Artificial Intelligence for Satellite Communications and (ii) MLSAT - Machine Learning and Artificial Intelligence for Satellite Communication.

In 2020 European Space Agency pioneered the deployment of Machine Learning into space for earth observation onboard a 6U Cubesat called Φ -Sat-1 using an Intel Movidius chip [54], with the intent of demonstrating and validating the state-of-the-art Deep Learning technology applied in-orbit for autonomously processing Earth Observation data.

¹https://www.wen.uni.lu/layout/set/print/snt/research/sigcom/sdn_lab

²<https://spire.com/blog/space-services/brain-in-space-new-satellite-testbed/>

More recently, ESA is planning a new activity for 2023 named 6G Satellite Precursor, whose objective is to develop an in-orbit laboratory that allows for the R&D process to be realised early in 6G adoption so that the satellite industry can adapt their technologies products, and use-cases to work side by side with terrestrial communication infrastructure. As part of the in-orbit laboratory, an AI chipset will be considered.

3) *NASA*: NASA has been actively investigating the Cognitive Radio framework in satellite communications within the John H. Glenn Research Center testbed radios aboard the International Space Station. In addition, a custom edge computing solution for DNA sequencing on the ISS on HPE's Spaceborne Computer-2 eliminates the need to move the massive data produced on the ISS by the DNA Sequencing project by presenting containerized analytical code where the data is being produced [58]. By leveraging the local computing available on ISS, the dependence on Earth and the time to get results can be reduced. The custom solution utilizes Red Hat CodeReady Containers, a single-node OpenShift cluster. This solution connects back on the ground with IBM Cloud, where researchers will develop, test and make their code ready to be pushed to the ISS. A ground-based solution on IBM Cloud will permit users to submit jobs, via a VPN connection, to the HPE ground systems that communicate securely with the HPE Spaceborne Computer-2 systems on the ISS, where Red Hat CodeReady Containers are installed. This solution will not only help expedite the research being performed on the ISS but will also open doors for many new explorations on the ISS and future missions in space.

4) *European Commission*: From the European Commission side, the Smart Networks and Services Joint Undertaking (SNS JU) recently awarded the first batch of projects to facilitate and develop industrial leadership in Europe in 6G networks and services [59]. Among the awarded projects, some of them included ML-powered optimizations for the integrated 6G and Non-Terrestrial network. The goal of such activities is to improve performance and automation, achieve zero-touch network management and timely monitor and predict overall network behaviour.

On the other hand, the European Union's Horizon 2020 research and innovation programme already granted a few space-related projects, e.g. ATRIA [60] or DYNASAT [61], both including the investigation of ML for SatCom systems.

IV. USE CASES ANALYSIS

This section presents in detail the main Satellite Communication use cases. We have divided the used cases in three layers, that group the conventional OSI layers (Fig. 4). For each use case we present (i) Motivation, (ii) Detailed Description, (iii) Conventional solutions and (iv) ML solutions. Referenced articles are categorized and summarized in Table III. Note that Neural Network approaches have a dedicated column to include literature that do not refer to supervised learning.

TABLE III
USE CASES ANALYSIS FOR ML AND SATELLITE COMMUNICATION: IN MAGENTA WE HIGHLIGHT AN OFFLINE ML TRAINING, IN CYAN AN ONLINE ML TRAINING

	Use Case	ML Framework				
		Supervised	Unsupervised	NN & DL	RL	DRL
ML for Satellite Communication	<i>Low Layers</i>					
	Beamforming	[62]	[63] [66], [67]	[64], [65] [66], [68], [69], [70]		
	Antenna Design					
	Antenna Beamforming					
	Beam Hopping			[71], [72]		[73], [74]
	Power			[75], [76]		[77]–[79]
	Flexible Payload			[80]		[74], [81]–[83]
	Bandwidth					[85]
	Beamwidth		[84]		[84]	[89]
	Link Adaptation	[86]		[87]–[89] [90]		
	Spectrum Sensing Classification					
	Interference Management			[91], [92]		
	Detection			[93]		
	Classification					
	Intersatellite Synchronization	[94]–[96] [99]	[97]	[98] [99]		
	Precoding			[100]–[103]		
	Link Quality Prediction			[104]		
	Predistortion					
	Coding		[105]	[105]–[110]		
	<i>Medium Layers</i>					
	User Scheduling		[111], [112]			
	NOMA Access	[113], [114]	[113]	[113], [115]–[117] [118]		[119], [120]
	Rate Splitting			[121], [122]		[123]
	Constellation Routing					[124]–[126]
	IoT Channel Access Scheduling					[82], [127]
	Fixed Access					[119], [120] [132]
	Random Access			[118]	[128]–[131]	
	<i>Upper Layers</i>					
	Traffic/Congestion Prediction					
	Satellite-Terrestrial Integration	[133]	[133]	[134]	[135]	
	Edge Computing and Caching					[126], [136], [137] [138]
	Network Security	[139]–[141]	[142]–[144]	[145]		
	Quantum Key Distribution		[146], [147]	[148], [149]		
	<i>Other</i>					
	ISAC		[150]	[151]–[154]		[150]
	Cooperative SATCOM					
	Img Processing			[155]–[160]		
	Ground Segment Dimensioning					

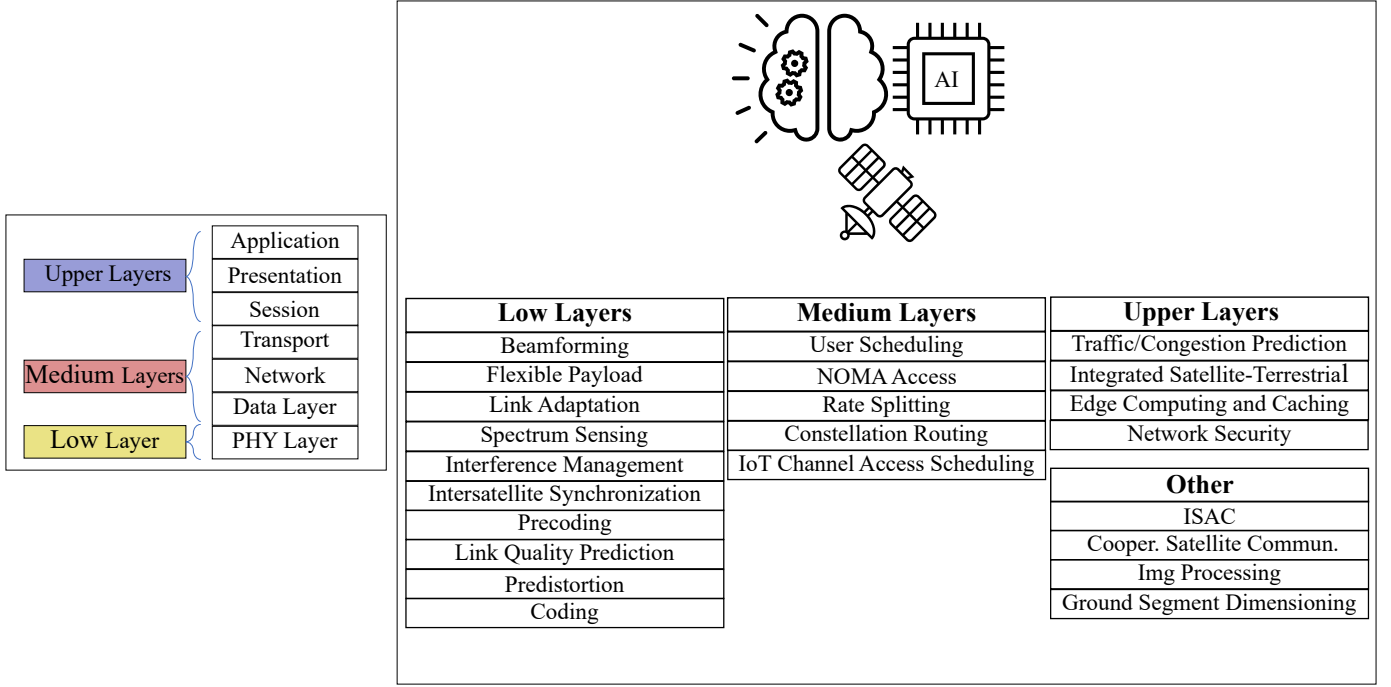


Fig. 4. Classification OSI-oriented for the use cases presented in Section IV

A. Low Layers

1) Antenna Design and Fail Detection:

a) Motivation: The advance in antenna design simulation tools has been notoriously significant in the past 30 years. This advancement has helped the research antenna community to design antennas accurately without deep knowledge of computational electromagnetics. However, this advantage comes with the high price of requiring a long computational time to simulate antenna models that use more than one radiating element, as in the case of antenna arrays for satellite communications. Moreover, suppose the designed and manufactured antenna array has problems due to bad connections or broken RF chains; in that case, more time will be necessary to address such problems, which in satellite communications cannot be easily afforded. Therefore, machine learning can help to overcome these time-consuming procedures, reducing antenna design and troubleshooting time.

b) Description: Antenna design nowadays uses full-wave electromagnetic software that applies different electromagnetic computational approaches to obtain the antenna responses like S-parameters, far-field, near-field, E-field, H-field, and current distribution results. Some solvers are very handy for satellite communications and can offer optimal solution time depending on the antenna type. For instance, to model large reflector antennas like antenna reflectors, it is very convenient to use the asymptotic solvers Geometric Optics Geometric optics and Physical optics [161]. Now, if we consider small-size antennae used in CubeSats, for instance, helicoidal, dipole, or patch antennas, then solvers based on the Finite element method [162], Finite-difference time-domain method [163], or Method of Moments [164] are the most suitable ones in terms of computational time and resources. Moreover, for

antenna arrays, some resources like the array factor can significantly reduce the computational time without considering mutual coupling, network distribution losses, radome effects, or spillover. On the other hand, antenna fail detection is crucial for phased-array antenna applications. This is because one or more faulty elements can distort the radiation pattern by modifying the main beam gain, beamwidth, Effective Isotropic Radiated Power, and scanning direction. Moreover, suppose a defective component is present; in that case, the array has to compensate for this by varying the amplitude of the phase of the rest of the element, which may take time using conventional optimization methods.

c) Conventional Solutions and Issues: Limitations of using a combined approach (full-wave simulation, asymptotic solvers, and array factor formula) for phase-array simulation and reflect-array design include the lack or complete absence of mutual coupling, feeding losses, axial ratio results, and cross-polarization. To overcome this limitation, a full-wave simulation that considers the entire structure will solve the problem. Still, at the same time, it will exponentially increase computational time and resources.

Considering the case of antenna failure detection, multiple approaches require the antenna near-field or far-field to obtain the surface currents and identify the faulty element [165]. Other methods use a genetic algorithm to detect the faulty element [166]. The first method needs the antenna near-field or far-field measurement, which makes it impossible for antennas placed on an in-orbit satellite, whereas the second method takes a lot of computational and processing time to find the failure element.

d) Proposed ML Solutions: In the antenna design field, some application of Machine Learning has been used, especially for reflect-array antennas that can be extrapolated to

transmit-array and phased-array antennas. In [64] a Neural Network is used as a tool able to perform near-field to far-field transformation, directly or through the application of the Theorem of Equivalence. Support Vector Machine algorithm is used in [167] in the design of reflect-array antennas for Direct Broadcast Satellite. Specifically, Support Vector Machine is adopted to create a surrogate model and substitute the full-wave simulation. A very comprehensive and complete compilation has been presented in [168], where the author gathers some applications of Machine Learning in antenna design to improve the acceleration design process, error minimization, antenna behavior prediction, and antenna optimization.

On the other hand, there are plenty of Machine Learning algorithms for detecting faulty antennas. In [65], the authors focus on faulty detection of phase arrays for satellite applications using Artificial Neural Network, and compare its efficiency with radial basis function neural network and Probabilistic neural network. Similarly, [62] considers a Deep Learning-based algorithm for fault detection. An Support Vector Machine classifier is instead proposed in [63] for planar array failure diagnosis.

While the above works show some preliminary effort on the use of Machine Learning in the antenna design, we have not identified relevant works in compensation of faulty elements using Machine Learning, which represents a gap to be filled by the research community in the next years.

2) Antenna Beamforming:

a) Motivation: Beamforming directs the transmitting antenna energy in a particular direction. Beamforming techniques have evolved over the years with increasing complexity. Being an excellent solution to a variety of challenging fields, Artificial Intelligence has already begun to be used by researchers as a possible solution to realistic beamforming problems in Satellite Communication. Implementing Artificial Intelligence algorithms to beamforming allows recognizing underlying patterns, which would be inefficient to identify manually. This "black-box" characteristic of Artificial Intelligence methods, especially Neural Network, is an advantage since the performance of the deployed model does not depend on the antenna characteristics. The main disadvantage of Neural Network is the lengthy training process to be conducted onboard or on-ground. However, despite the significant time required for training an Neural Network in a multi-beam satellite scenario, the fast temporal response of Neural Networks can be exploited for adaptive beamforming. In fact, after training, Neural Network can perform cumbersome operations in real-time, representing a promising solution for adaptive beamforming, where weights must be repeatedly calculated and conventional solutions are lengthy.

b) Description: An antenna array consists of multiple same-characteristic antennas separated by a specific distance, determined by the lattice, which can be square, triangular or irregular. The radiation pattern maxima, beamwidth, Side Lobe Level, and null pattern are functions of the number of radiating elements, amplitude taper, lattice, and progressive phase shift [169]. Each antenna element, often called a unit cell, will determine mainly the scanning losses in the overall array, the complexity of the feeding network, associate losses, and

production cost. Now, considering a fixed-size array employed for a given service in a certain region on the Earth, the beamforming characteristics control of the array are mainly a function of the complex-valued weight matrix, whose entries' magnitudes control the beamwidth in both planes and the Side Lobe Level; on the other hand, the entries' phases of the matrix control the antenna beam scanning, and nulls [170].

c) Conventional Solutions and Issues: The conventional solution applied to beamforming for satellite multi-beam application is an optimization problem, which consists of calculating the beamforming weights accurately to match the desired beam pattern [171], focusing the available transmit power on locations of interest while placing nulls at other directions, achieving a predetermined 3-dB main-beam width, and minimizing the Side Lobe Level. In the literature, there are multiple solutions to beam synthesize [172]–[177]. Usually, a simple beamforming scenario, which is controlling the gain, can be easily simplified to a sum of the contributions of each antenna element with the required weight matrix as indicated in 1

$$\text{AF} = \sum_{m=1}^{M_x} \sum_{n=1}^{N_y} W_{mn} e^{j(m-1)(\kappa d_x \sin(\theta) \cos(\phi))} \times e^{j(n-1)(\kappa d_y \sin(\theta) \sin(\phi))}, \quad (1)$$

where M_x is the number of elements in the x -direction, N_y is the number of elements in the y -direction, κ is the wave number, d_x is the period in the x -direction, d_y is the period in the y -direction, θ and ϕ are the evaluating angles, and W_{mn} is the weight matrix entry of index (m, n) . Now, if we consider a beam scanning angle (θ_0, ϕ_0) scenario, then the computation of the weight matrix can be done via

$$W_{mn} = e^{-j\kappa(m d_x \sin \theta_0 \cos \phi_0 + n d_y \sin \theta_0 \sin \phi_0)}. \quad (2)$$

Moreover, the previous weight matrix can be modified by applying an amplitude taper like Hamming [178], Taylor [179], Hann [180], Chebyshev [181], or others to reduce the Side Lobe Level and adjust the beamwidth. Complementarily, to avoid interference from another adjacent beam, it is convenient to add a nulling in a certain direction using, for example, [170]. Therefore, combining all the previously analyzed constraints to synthesize the beams in a multi-beam scenario will require proper optimization tool, computational time, computational resources, and power consumption, being limited in an in-orbit satellite. A diagram covering all the beamforming inputs and outputs is presented in Figure 5.

d) Proposed ML Solutions: Literature offers several contributions to the implementation of AI techniques and especially deep learning Deep Learning. Due to their fast response, rapid convergence rates, successful failure detection, and proactive decision capability, most adaptive beamforming adaptive beam forming) techniques are now based on Deep Learning realizations. In [70], the neural beamformer utilizes Recurrent Neural Network to search the optimal between three digital beamforming strategies. The work in [68], even if not explicitly targeting a Satellite Communication scenario, proposes a feedforward Neural Network and Recurrent Neural Network to compute the complex feeding weights for a linear

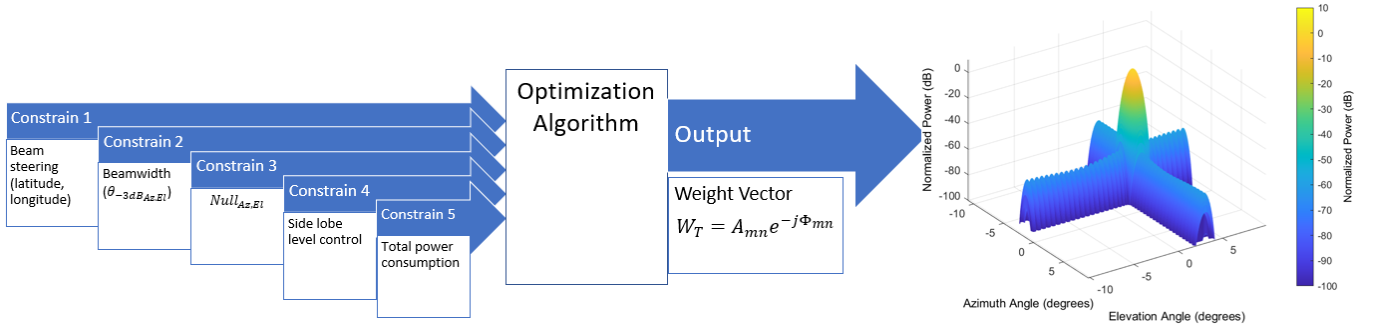


Fig. 5. Diagram with the inputs and outputs of a beamforming optimization.

microstrip antenna array. The work [66] exploits Convolutional Neural Networks with offline training to match a specific radiation pattern and computes the array antenna excitations. Convolutional Neural Network represents a faster alternative to conventional array synthesis techniques. [67] utilize a linear regression model to steer to predict the directions of arrival of the incoming signal to steer the array beam on the ground in the satellite direction. [69] implement a Deep Neural Network to perform beamforming in a Low Earth Orbit system at sub-THz, where the atmospheric absorption is compensated by very large antenna arrays, that need to achieve very high gain in the desired direction via very narrow beamforming. The DNN beamformer is trained to mimic the actual output voltage generated by a TTD beamformer as the angle of arrival (AoA) changes at the receiver due to the rapid movement of the LEO satellite.

3) Flexible Payload-Beam Hopping:

a) Motivation: The main idea of the High Throughput Satellites technology breakthrough is generating multiple small beams and re-using the available satellite spectrum numerous times across the coverage area to increase the total satellite capacity significantly. To cope with the critical adjacent-beam interference in High Throughput Satellites systems, one excellent solution has been introduced is to divide the whole spectrum into various bands so-called “colors”, which are separated by frequency allocation and polarization. Recently, in the new Satellite Communication generation of Ultra High Throughput Satellites, the current coverage area is foreseen to be divided into a higher number of smaller beams. Herein, the reduction in the beam pattern angle focuses one beam on a smaller spot and enhances a higher capacity gain even at a lower CAPEX cost. However, dividing the coverage into smaller spots comes at the expense of a high probability of no demand within some beam spots at a specific time [182]. In such cases, the corresponding no-demand beams can be turned off to reduce the operating cost. That is the key idea of beam hopping technology in the new generation of satellite systems [183].

b) Description: Beam Hopping helps Satellite Communication adapting to the time-varying communication demands which dynamically change due the time, mobility, and weather condition. Different from the quasi-static lighting in traditional multi-beam satellite systems, the satellite circularly illuminates

a set of specific beams according to the generated schedule. The Beam Hopping technology refers to the use of less numerous active beams in accordance with the design of the BH pattern to serve the ground terminals within the satellite coverage area, herein, the available satellite resources can be further optimally used to provide services with the heterogeneous demands [183], [184]. The main difficulty in Beam Hopping pattern design is the large search space for identifying the optimal patterns. That is, in order to find the optimal solution, the number of possible Beam Hopping patterns to be searched increases exponentially with the number of beams [72]. In this context, Machine Learning appears as a promising technique that offers an alternative to design optimization algorithms for complex resource management in wireless networks.

c) Conventional Solutions and Issues: The SATCOM systems realize the BH technology the the flexible payload where the design parameters can flexible configured. The existing BH-related works aim to compose the beam switching and jointly optimizing beam pattern illumination mechanism with power/spectrum allocation and precoding/beamforming design. The work in [185] propose an iterative algorithm for Beam Hopping illumination design while Beam Hopping can also be optimized to minimize the transmitting power, as shown in [186]. Lin Chen et. al. in [187] have addressed a dynamic precoding-regarded beam-hopping design to minimize the cross interference and also avoid precoding whenever possible while it can satisfy all beam demands. The numerical simulations have illustrated the superior of the dynamic BH strategy to the traditional cluster-based BH mechanisms. [188] aims to jointly determine Linear Precoding vectors, Beam Hopping, and discrete DVB-S2X transmission rates for the GEO satellite communication systems to minimize the payload power consumption and satisfy ground users’ demands within a time window. Regarding the constraint on the maximum number of illuminated beams per time slot, the technical requirement is formulated as a sparse optimization problem in which the hardware-related beam illumination energy is modeled in a sparsity form of the LP vectors. To cope with this problem, an iterative compress-sensing-based algorithm is employed to transform the sparsity parts into the quadratic form of precoders. Authors in [189] design a dynamic beam hopping and beam position division for Low Earth Orbit Satellite Communication systems to shorten the packet queue-

ing delay. In particular, this work aims to develop a novel mechanism to cover all users with the least number of beam positions in a set of beams which can be illuminated. The problem is first transferred into a p -center one, then the beam positions among the footprint of Low Earth Orbit satellites are determined dynamically by the user distribution and the traffic distribution. The work in [190] focuses on developing a novel cooperative beam-hopping mechanism for Non Geo-Stationary constellation consisting of multiple satellites. Regarding the load-balancing issue, the proposed algorithm in this work is designed to manage both intra-satellite interference and inter-satellite interference by designing beam-hopping patterns with spatial isolation characteristics.

d) Recent ML-based BH solutions and Future Directions:

A number of recent works have also regarded employing the Machine Learning tools to several Beam Hopping-enabled SATCOM schemes. Lei Lei et. al. investigates in [71] a Deep Learning method for Beam Hopping optimization and to predict the number of elements in the beam pattern for Beam Hopping and speed up the process of Beam Hopping pattern selection and allocation. The same authors in [72] have explored a combined learning-and-optimization approach to provide an efficient, feasible, and near-optimal solution. The investigations are from the following aspects: 1) Integration of BH optimization and learning techniques; 2) Features to be learned in BH design; 3) How to address the feasibility issue incurred by machine learning. The numerical results have illustrated that the learning feature can enable a very high accuracy in predictions for the size of beam clusters that should be illuminated together to meet users' demands. In [73], the authors investigated the optimal fairness policy for beam hopping in DVB-S2X satellite regarding two main goals: minimizing the delay of real-time services transmission and maximizing the throughput of non-instant services transmission. To cope with the time-varying and unpredictable wireless channel issues, and differentiated service arrival rates in the multi-beam satellite environment, this work employed the model-free multi-objective deep reinforcement learning approach to learn and retrieve the optimal policy through interactions with the situation. To solve the problem with action dimensional disaster, a novel multi-action selection method based on Double-Loop Learning (DLL) is proposed. Moreover, the multi-dimensional state is reformulated and obtained by the deep neural network. Under realistic conditions achieving evaluation results demonstrate that the proposed method can pursue multiple objectives simultaneously, and it can also allocate resources intelligently adapting to the user requirements and channel conditions. The DRL-based approach is also exploited in [74] where a dynamic beam pattern and bandwidth allocation scheme based on DRL, which flexibly uses three degrees of freedom of time, space and frequency. Considering that the joint allocation of bandwidth and beam pattern will lead to an explosion of action space, a cooperative multi-agents deep reinforcement learning (MADRL) framework is presented in this paper, where each agent is only responsible for the illumination allocation or bandwidth allocation of one beam. The agents can learn to collaborate by sharing the same reward to achieve the common

goal, which refers to maximize the throughput and minimize the delay fairness between cells. Regarding future works on applying ML/DL in BH design for the next generation of SATCOM, it can consider developing the payload with high-flexible adjustment and control capabilities at different levels such as beam shape and network topology. This can be done by employing more sophisticated processing capabilities of satellite digital devices, as well as ML/DL technologies. The work may be also enlarged by linking it some critical technical problem is multi-connectivity SATCOM schemes such as user scheduling, beam placement, and traffic prediction.

4) Flexible Payload-Power Allocation:

a) Motivation: The next generation of satellite communication systems will incorporate flexible digital payloads, allowing advanced resource management techniques to be implemented. In this case, the digital payload can be configured to change the bandwidth, carrier frequency, and transmit power of the system in response to heterogeneous traffic demands [191]. Notably, the payloads of the system are equipped with multiport amplifiers so that the power of the system can be divided across multiple beams. Hence, it is possible to use the remaining power in beams with lower traffic loads to those with higher traffic loads [192]. However, these flexible payloads require advanced power allocation algorithms to optimize the total transmit power based on each beam's traffic load [193].

b) Description: In satellite systems, power consumption is a major limitation due to its significant impact on satellite mass and operation lifetime. Additionally, the upcoming on-board beamforming technology, which is highly power-hungry, presents a major challenge for on-board power optimization. In order to reduce power consumption, it is necessary to implement innovative techniques, and one feasible approach is to employ advanced power optimization techniques to reduce radiated power. In addition, conventional satellites provide connectivity for users through a multibeam footprint with uniform power allocation per beam. Uniform power allocation, however, is inefficient when there is a distribution of non-uniform demand, which may cause beams with low demands to be assigned too much power while beams with high demands to be assigned too little. Consequently, we may be unable to meet some users' demands. Thus, it is necessary to implement an advanced power allocation algorithm in the next generation of satellite communication systems to ensure efficient power management [194], [195].

c) Conventional Solutions and Issues: Several studies have been conducted to develop power allocation algorithms using classical methods such as analytical optimization and heuristic/metaheuristic methods. With analytical optimization, the algorithm provides an optimal or nearly optimal solution. In this context, in [196], an energy-efficient power allocation scheme has been designed to maximize the average rate ratio over the total consumed power. In addition, energy-efficient power allocation over multibeam satellite downlinks with imperfect Channel State Information has been proposed in [197]. Furthermore, an energy-aware power allocation algorithm has been developed to jointly minimize unmet system capacity (USC) and total radiated power in [198]. In [199], a power

optimization approach based on demand and channel quality has been evaluated to minimize the difference between offered capacity and demand while maintaining reasonable fairness for all users.

A metaheuristic method may not guarantee optimality, but it is well-suited to nonlinear, multi-objective, and NP-hard problems. In this context, Particle Swarm Optimization is used in [200] to solve the power allocation problem in the multibeam satellite system, providing the minimum signal-to-noise plus interference ratio that the user terminal requires for reliable communication. Furthermore, a power allocation technique has been proposed in [201] to simultaneously minimize unmet capacity and power consumption in multibeam satellite systems. In [202], a game-based dynamic power allocation (AG-DPA) algorithm has been developed for multibeam satellite systems to match the offered capacity and per-beam demand.

The above-proposed method lacks flexibility in bandwidth allocation, limiting the ability to fully exploit flexible on-board payloads. This is addressed in [203]–[207] using analytical optimization and in [208], [209] using a metaheuristic technique to jointly optimize the power and bandwidth to match the capacity offered with the per-beam demand. Nevertheless, using analytical optimization and metaheuristic techniques for power allocation involves many optimization parameters, so these optimizations are very time-consuming. Furthermore, the algorithm needs to be performed frequently if there are slight changes in traffic requirements or channel conditions. Thus, the proposed power allocation algorithms may not be suitable for real-time applications. In this case, machine learning techniques can be vital to provide a low-complexity power allocation algorithm that adapts to the change of demand and/or channel conditions.

d) Proposed ML Solutions: Several researchers have proposed machine learning-based solutions to overcome the limitations of conventional power optimization techniques. Using a neural network, [75], [76] proposes a classification algorithm that determines the power required for each beam to minimize the error between the offered capacity and the requested demand. Data is generated from a linear approximation function of spectral efficiency to train this neural network. Since neural network training is performed offline, the main advantage of this method is that it performs resource allocation at a low computational cost. However, the exponential relationship between the number of classes and the number of beams may increase the complexity of the computation and require further investigation. In [84], a Convolutional Neural Network is used to minimize the error difference between the offered capacity and the per-beam demand while saving unused satellite resources. For this, training data is generated from a realistic traffic model. The proposed algorithm in [84] shows better performance on the tradeoff of reducing capacity error and power consumption when compared with traditional resource allocation approaches. However, the algorithm depends on training data which is required to train the neural network if there is a change in the traffic model.

A power allocation algorithm using Deep Reinforcement Learning to maximize the system transmit data has been investigated in [77]. In this method, the satellite acts as an

agent, while the traffic between beams acts as an environment. Additionally, the DRL state contains a combination of buffered data among beams, power allocations, and beam geographical distributions. Accordingly, the agent adjusts the power among the beams to maximize data transmission. The simulation results in [77] demonstrate that the proposed method outperforms heuristic power allocation methods in terms of the system's throughput. However, the power of each beam is adjusted to some discrete values, which may limit the system's performance.

In contrast, DRL for continuous power allocation in flexible high throughput satellites to minimize the overall unmet system demand and power consumption has been proposed in [78]. In this DRL-based power optimization method, the agent takes action (power allocation) given the state of the environment, which contains both the satellite model and demand per beam. For each action, a reward is assigned, modeled as a function of the difference between unmet system demand and power consumption. The proposed method is evaluated using time-series demand data provided by SES, and the authors demonstrate that the algorithm can match beam demand. Furthermore, the same DRL model as [78] has been used in [79] to compare with metaheuristic techniques. This comparison shows that DRL is the most appropriate option when time is critical and a solution is needed in seconds.

5) Flexible Payload-Bandwidth Allocation:

a) Motivation: Future satellite networks are expected to offer extremely high data rates, enabling seamless integration with large satellite-terrestrial networks. However, due to spectral limitations, the main challenge will be providing high data rates while reducing the cost (satellite launch and operation) per bit. The key to solving this challenge is to maximize the system's spectral efficiency by utilizing the limited spectrum effectively [210]. Thanks to the emerging on-board digital payloads, the system not only offers flexibility in power optimization but also allows flexibility in bandwidth utilization. Typically, a flexible payload consists of a channelizer, which digitally routes channels within beams by digitizing the signal. It has a variable number of ports, each capable of processing a specific amount of bandwidth (channel bandwidth) while providing switching capability at the subchannel level [192]. However, an optimal bandwidth allocation technique is required to configure these digital payloads. Consequently, we can manage the limited spectrum to enhance the system's spectral efficiency.

b) Description: A satellite with multi-beam technology generates narrow beams, each covering a specific geographical area. As the beams represent a specific area, the same frequency can be reused on non-adjacent beams. Therefore, several beams can share the same frequency band, resulting in a significant increase in spectral efficiency. In traditional satellite systems, bandwidth is divided uniformly among the beams, with each beam sharing the total bandwidth equally. While some beams may have different demands, uniform bandwidth allocation may not be efficient in this case. As a result, bandwidth may be allocated excessively to beams with low demand and insufficiently to beams with high demand. Consequently, we may not satisfy some beam demands. Thus,

advanced bandwidth allocation techniques must be developed while leveraging flexible payloads to allocate bandwidth per beam demand.

c) Conventional Solutions and Issues: Based on the characteristics of the beam demand, bandwidth allocation techniques can be categorized into three types:

- 1) Orthogonal frequency band assignment across adjacent beams: One example is a four-color scheme in which the total bandwidth is divided into four frequency bands (subchannels). Accordingly, adjacent beams receive orthogonal frequency bands, while non-adjacent beams receive the same frequency band [201], [211], [212]. However, this method is inefficient for heterogeneous demand, where some high-demand beams require more bandwidth while others need less. Furthermore, assigning orthogonal frequencies to some adjacent beams may not be necessary if their respective terminals are located near the center of the beams and their demands are low.
- 2) Semi-orthogonal frequency band assignment across adjacent beams: In this case, adjacent beams may reuse the same frequency band. In [185], [203], subchannels (frequency bands) are iteratively assigned to each beam according to the traffic request while minimizing co-channel interference. Similarly, in [213], subchannels are allocated to each beam to maximize system throughput. This method uses the same subchannels between adjacent beams within a range where inter-beam interference is minimal, while other ranges use different subchannels. Alternatively, analytical optimization in [206], [207], [214], [215] and metaheuristics in [208], [209], [216] have been considered for bandwidth allocation based on the per-beam demand. However, it may be challenging to meet all beams that have a high demand using the techniques mentioned above.
- 3) Full frequency reuse can be applied to satisfy high beam demand while managing interference among the beams using precoding techniques [217], [218].

Most of the methods mentioned above use combinatorial optimization, resulting in longer computation times as the number of beams or subchannels increases. Furthermore, re-executing the algorithm may be required if traffic requirements or channel conditions change slightly. In this situation, developing an algorithm based on machine learning that can adapt to changing traffic and/or channel conditions could be useful.

d) Proposed ML Solutions: Several researchers have proposed machine learning-based bandwidth allocation techniques. In [81], a dynamic channel allocation algorithm for multibeam satellite systems is presented based on deep reinforcement learning, where the total bandwidth is divided into subchannels and assigned to each beam in order to reduce the blocking probability of the system. A discrete-time event system based on service arrival events is assumed and the satellite checks which channel to assign to the new arrival user terminal. In this case, the satellite serves as the agent, the action determines which channels to allocate, and a positive reward is provided when the new service is satisfied. On the other hand, a deep reinforcement learning-based approach for energy-efficient channel allocation in the satellite internet

of things has been studied in [82]. This method involves finding an optimal channel allocation strategy that allocates the limited channels to users on the ground while saving transmitting power in the long term. In [82], the satellite is viewed as an agent, and the user requirement and location represent the environment's state. Based on the environment's state, the agent allocates channels. In this method, the reward is considered the sum of the system's power consumption and service-blocking rate. Based on the numerical results in [82], the proposed method has better energy efficiency than traditional algorithms.

In [83], an algorithm for time-frequency resource allocation with deep reinforcement learning has been investigated to maximize the number of users and system throughput. In [83], the ground gateway is considered as an agent, as well as the number of time-frequency resource blocks and the user requirements as the environment. Consequently, the agent allocates the resource block while examining the state of the environment. Based on the simulation results presented in [83], the proposed deep reinforcement learning method gives better resource utilization with low computational complexity compared to the Genetic Algorithm (GA) and Ant Colony Optimization (ACO). Similarly, cooperative multi-agent deep reinforcement learning to allocate bandwidth with flexibility in time, space, and frequency has been investigated in [74].

Generally, several studies have investigated the use of machine learning applications to maximize system quality of service [219], spectral efficiency [220], transmission efficiency [221], and resource utilization [222], [223]. While [224] applies multi-objective reinforcement learning to manage multiple systems' performance attributes such as throughput, bandwidth, spectral efficiency, bit error rate, power efficiency, and power consumption. On the other hand, in [80], a demand-aware bandwidth and power allocation algorithm have been proposed based on analytical optimization and Deep Learning (DL). Thus, analytical optimization enables bandwidth and power allocation, while DL can speed up computation.

6) Flexible Payload-Beamwidth Allocation:

a) Motivation: User demand requirements are generally met in satellite industry through multiple spot beams of high-throughput satellites with fixed multiple beam pattern and footprint planning. However, mobile broadband users such as air (airplanes) and water (ships) located in remote geographic locations have dynamic traffic demand [225]. Additionally, in next generation Satellite Communication the traffic demand is expected to become less heterogeneous due to non-uniformly distributed user terminals with varying traffic demand. This cannot be faced by the current fixed footprint management and requires satellites equipped with beamforming capabilities and beam pattern plan [226], [227].

b) Description: The main objective of beamwidth allocation is minimizing a proposed primary KPI (Key Performance Indicator), such as the gap between the offered and required capacity or the Coverage Error. The system characteristics, such as service area, satellite orbital position, bandwidth, and power per beam, must be defined. Together with the cost function, technical requirements or constraints must be specified to define the design parameters, such as the possible

beamwidth values that can be assigned or the possible antenna orientations that can be provided [226].

Once all the system conditions have been defined, the distribution of the average traffic demand in the service area must be studied and divided into K regions. In addition, the beamwidth of k -th region is assigned to meet the predicted traffic demand distribution. Generally, higher traffic demand require narrower beams, and larger beams should be assigned for lower traffic demand.

The coverage is thus designed for each possible configuration of the beam orientation, defined by the technological constraints. All Satellite Communication link conditions, such as the resulting number of beams, the capacity offered on each beam, the provided capacity of the overall satellite, and the KPIs to calculate the cost function for time instant t , must be evaluated.

c) Conventional Solutions and Issues: A first contribution for a more efficient beam pattern planning of a multibeam antenna with reflector and overlapping beams has been made by the authors of [228]. To compare conventional contoured beam antennas with the multibeam antenna, the authors of [229] propose a fixed beam for a specified number of beams to coverage using overlapping spot beams with a hexagonal grating arrangement. Additionally, the authors of [169], [230], [231] analyze the antenna requirements for planning a good beam pattern for fixed beam footprint plans. Hence, all the first works consider only the geometrical feasibility of beams in the beam pattern and footprint plan guaranteeing global coverage. Consequently, the coverage region and the offered throughput remain the same for all multiple spot beams, regardless of traffic demand.

In [226], a proposal is presented to optimize the coverage and pattern of irregular beams by minimizing the cost per Gbps in orbit, the Normalized Coverage Error, and the Offered Capacity Error per beam. The analysis and performance of the case study are presented, and a comparison is made with a previous algorithm for a uniform coverage area.

Recent work has focused on adaptive schemes to accommodate fluctuations in traffic demand. Global resource management for dynamic beam steering due to changes in Quality of Service requirements and user channel conditions is discussed by the authors of [232]. The authors of [233] optimize the beam directivity and transmit power based on the traffic demand to improve the overall system performance. Also, the authors of [234] adjust the satellite transmit antenna aiming point to maximize the satellite ground station's signal-to-noise power ratio (SNR) and minimize interference with terrestrial networks. An optimization tool is proposed by the authors of [235] to jointly optimize the number of beams, beamwidth, power, and bandwidth allocation to match the provided data rate to the expected traffic demand. Till they optimize the beam positions and beam shape. A mathematical model to derive multipoint beam arrangements is proposed by the authors of [236], and they discuss the relationship between multipoint beam placement and user throughput. Further, they only study the effect of geographical distances between spot beams in the same frequency band and geographical spaces between adjacent spot beams in different frequency bands

on the overall system performance. However, they do not implement adaptivity in terms of flexible beam size and beam position in the beam plan. So far as we are aware, the authors have yet to consider the mobility aspects of non-uniformly distributed users and dynamically to change beam demand during beam pattern and footprint planning.

In [227], advanced fixed multibeam patterns and footprint plans are studied to show their drawbacks to supporting non-uniformly distributed user terminals and varying traffic demands. The authors propose an adaptive multibeam pattern and footprint plan in which we design point beams with flexible sizes and positions based on the spatial clustering of users to increase the flexibility of the high-throughput satellite system.

d) Proposed ML Solutions: In [84], the use of a CNN to solve the resource management problem in SATCOMs, including beamwidth management using a suitable cost function and a realistic traffic model, is analyzed. The suggested cost function aims to minimize the error between the offered and required capacity while minimizing the number of resources used in the satellite and obtaining the adaptive beamwidth at each instant. One of the limitations of CNN in this work is the dependence on the traffic model used during training. Thus, in a real system with changes in traffic behavior not conforming to the model, the CNN will have to be retrained.

Authors also evolve the same cost function in [85], analyzing the use of Deep Reinforcement Learning and Reinforcement Learning algorithm to manage the available resources in flexible payload architectures, including beamwidth. These algorithms are Q-Learning (QL), Deep Q-Learning (DQL), and Double Deep Q-Learning (DDQL), which are compared based on their performance, complexity, and added latency. On the other hand, this work demonstrates the superiority of a decentralized cooperative multi-agent (CMA) distribution over a single agent for multibeam scenarios.

However, implementing ML to enhance flexibility in geographical positioning has yet to be explored as another possible resource in addition to beamwidth. In that sense, it is expected that ML can enhance the gain in the cost function, a reduction in the time and complexity required because the search space is huge for conventional optimization techniques.

7) Link Adaptation:

a) Motivation: Adaptive Coding and Modulation techniques are among the most successful fade mitigation techniques for wireless links. The channel conditions in SATCOM links, such as rain, scintillation and delay, make it difficult for the estimation signal for feedback to choose the most suitable Modulation and Coding (MODCOD) at all times. ACM is easy for a single link without interference, where a pilot signal can measure the channel and report back to the receiver. However, the innate delay of satellite links changes the problem, especially in multi-beam satellites with many channels to estimate. There is an evident "dead time" between the fade event and its corresponding ACM reaction. This dead time may last between 600 and 800 ms and is caused by the two hops over the satellite and additional averaging delays. Frame errors might appear during this period if the SNR decreases. In addition, emerging SATCOM scenarios with NGSO and GEO increase

the interference between systems, whereas ACM can improve communication performance by avoiding interference as far as possible. ACM employment includes as benefits greater distances free of errors, using smaller antennas which save on mast space, and higher availability to provide better link reliability. The usage of ACM techniques aims to improve SATCOM links' operational efficiency by increasing network capacity over the existing infrastructure – while reducing sensitivity to environmental interferences.

b) Description: ACM or Link Adaptation (LA) is a technique used in communications to match the modulation and coding schemes among other signal and protocol parameters according to the quality of the radio channel. LA dynamically and automatically changes the modulation and/or the Forward Error Correction code rate as the radio link conditions change. The receiver estimates the instantaneous SNR and signals, and the estimated Channel State Information is transmitted to the transmitting end through a feedback channel. The transmitting end keeps the transmit power unchanged and changes the coding and modulation mode according to the channel state. High-order MODCOD schemes are selected during good channel states, while lower MODCOD schemes are set when the channel states are bad. This means the system can transmit data at high speed under favourable channel conditions and reduce the throughput when the channel is degraded, thereby providing high spectral efficiency without sacrificing system power and Bit Error Rate, maximizing the throughput obtained under momentary propagation conditions. In the uplink, the on-board ACM controller selects an appropriate modulation method and channel coding mode based on CSI sent to the satellite via the feedback channel.

c) Conventional Solutions and Issues: Typical LA requires knowledge of the channel. This is easy for a single link without interference, where the pilot signal can measure the channel and report back to the receiver. However, the channel is stochastic, and the reported realization may differ from the actual one. In addition, the channel can substantially change during the so-called feedback delay (time between pilots being transmitted, feedback received and sent signal). The problem becomes more challenging when an interference channel is considered, or multiple channels are present, including Multiple-Input Multiple-Output (MIMO) systems. The authors in [237] and [238] have considered the power adaptation for interference channels when the complete CSI is unavailable at each transmitter. Remarkably, in [238], different levels of coordination between transmitters are evaluated. From the point of view of emerging scenarios, an AMC scheme suitable for integrating the satellite in 5G networks is proposed in [239] based on 21 MODCOD schemes and different thresholds for SNR grouping. ACM techniques for high-rate transmitters are developed on FPGA in [240] and proposed for LEO satellites. This proposal reaches a spectral efficiency of 7.068 bps/Hz with a code rate of 0.8889 in the highest rate mode. For the optical feeder link, time-packing is used as ACM considering low-order M-ary Pulse Amplitude Modulation (M-PAM) as shown in [241] where time-packing enables a finer granularity on the LA capability of the optical link, enabling adjust its SE according to the moderate attenuation that thin cloud layers

introduce. [242] uses precoding at the gateway side and SINR estimation at the terminal to address the issue of LA or ACM in a multi-user and multibeam satellite. This strategy includes a scheduler among different SNIR values and with imperfect CSI to maximize the similarity of two channels in the same frame. The literature shows two main problems to be solved in LA with the new SATCOM paradigm. On the one hand, the wide margin that the Q/V and W frequency bands require for new mega-constellations and, on the other, the amount of information that must be estimated in multibeam satellites.

d) Proposed ML Solutions: An improvement of the CSI prediction in the satellite network is expected with the ML algorithms that will help realize ACM, predicting the most appropriate ACM to optimize system capacity for given channel conditions. This strategy will enable us to develop a more efficient ACM mechanism. The application of ML for LA has been considered in [87] and [88] using Bayesian methods focused on Neural Network, which shows performance gains over the commonly used ACM selection based on effective SINR. The work presented in [90] focuses on applying ML to CSI prediction in SATCOM and using the improved prediction results to optimize system capacity. The different tested ML techniques demonstrate the improvement in system performance and the feasibility of deploying an ML-based prediction framework. The ML-based CSI prediction model provides an average capacity increase of up to 10.9% with acceptable overhead. The authors in [89] propose a neural episodic control (NEC) algorithm from the DRL group as an ACM scheme for Inter-Satellite Communication Link. The proposed scheme adjusts the modulation and coding scheme region boundaries with a differentiable neural dictionary of the NEC agent, which enables the effective integration of the previous experience. The ACM proposed in [86] utilizes an online random regression forest (ORRF) to predict time series of SNR values aiding the ACM switching decisions, introducing adaptive techniques for satellite links on Q/V-Band. The ORRF-based approach could outperform the classical approaches in terms of SE, and the parameterization of the ORRF is simpler and needs less knowledge of the channel properties than the discussed classical ACM approaches.

8) Spectrum Sensing, and Interference Detection and Classification:

a) Motivation: The usage of spectrum in current day Satellite Communication follows a rigid and fixed spectrum allocation management paradigm. However, since the number of satellite and terrestrial services that require an ever-increasing bandwidth grows while the spectrum resource is fixed, the existing paradigm is inefficient. A direct implication is that the available spectrum has to be shared and reused among those services. Besides interference, which is due to sharing the spectrum, there are intentional jamming sources that try to disrupt the Quality of Service of satellite communications systems. Both types of interference require detecting and classifying interference signals to take the appropriate measures. This will require a two-step process: 1) observing/collecting data, and 2) processing these data to decide on the current status of the spectrum.

In conventional approaches, the data-processing stage hap-

pens online and is designed by domain experts based on the specific knowledge of the signals' characteristics. On the other hand, in ML approaches, the algorithm learns the characteristics of the signals by training on a sample training data offline. Once the algorithm learns the characteristics well enough to differentiate between the different states of the spectrum, the resulting model is deployed online.

b) *Conventional Solutions and Issues:* To address the issue of spectrum sharing and the entailed interference management issue, many works in the literature focus on Cognitive Radio SATCOM, where the coexisting systems are divided into primary and secondary, and the secondary system has to operate within limits specified by the Quality of Service of the primary system. Cognitive Radio is divided into overlay CR and underlay Cognitive Radio networks [243]. In the former, the secondary system relies on detecting and using white spaces of the spectrum (parts of the spectrum that are not currently being used by the primary system), whereas in the latter, the secondary system relies on adapting the transmission parameters based on the maximum interference the primary system allows at its receivers to maintain the required Quality of Service [244]. These systems, however, require collaboration and sharing of information between the primary and secondary systems, which is usually done via a third-party (centralized) manager, which adds complexity, overhead, and cost. Another approach to manage interference while sharing the spectrum is interference detection and mitigation, where conventionally, a receiver analyzes the received signal based on some *a priori* statistical knowledge of the desired and interference signals as well as the channels [245], and takes actions whenever interference is detected, like changing the transmission carrier frequency and subband. In [246], the authors studied distributed Dynamic Spectrum Sharing for a system where Geosynchronous Orbit, Non Geo-Stationary, and terrestrial networks share the same spectrum in both downlink Ka-band (17.7-19.7) GHz, and uplink Ka-band (27.5-29.5) GHz, where terrestrial networks are the incumbent (primary) users that consist of 5G fixed wireless access and microwave links. The multiple access technique assumed for satellite-related links is Multi Frequency-Time Division Multiple Access (MF-TDMA), whereas for terrestrial networks, it is Frequency Division Multiple Access. Two techniques were proposed: collaboration protocol, where nodes from different systems share information they capture from other nodes participating in the collaboration, and spectrum sensing, where each receiver makes decisions on spectrum allocation based on locally acquired energy maps of the spectrum by processing the raw I/Q received data. An advantage of the latter is that the receiver can detect known and unknown interference, while only interference sources shared through CP are detected in the former technique. Another advantage of decentralized spectrum sensing is that information exchange happens only between communicating entities for resource re-allocation or release, which reduces the overhead of sharing information. Although the spectrum sensing technique does not require exchanging information from entities of other systems, it requires changes in the radio modem to integrate spectrum sensing. Some satellite links, such as Global Naviga-

tion Satellite Services (GNSS) systems, require *anti-jamming* capabilities. Conventionally, GNSS uses postprocessing techniques that rely on the statistical characterization of the signals to activate interference mitigation procedures since different countermeasures are appropriate to different jamming signals [93]. These approaches can be effective when the jamming signals are known, and the number of jamming sources is limited. However, estimating the probability density functions for all hypotheses becomes more complex as the number of classes increases.

c) *Proposed ML Solutions:* Data-driven approaches employing Machine Learning algorithms are considered an alternative way of detecting interference [91], [92], where complexity is transferred offline in training the algorithm instead of the online complexity in processing the received signal in the conventional approaches. In [92], the authors studied the problem of wireless signal identification that covers the problems of modulation recognition, and identifying the interference source, using machine learning algorithms. Three signal representations were considered: I/Q representation, amplitude/phase representation, and frequency domain representation. The authors found that amplitude/phase representation gives better accuracy for modulation recognition than the other two representations. In comparison, for the problem of interference detection, they found frequency domain representation to attain better accuracy. In [91], the authors considered a two-stage system based on the Deep Learning model, wherein in the first stage, an autocorrelator is used to detect the interference signal. In the second stage, Deep Neural Network classifier is employed to classify the interference signal from three different well-known standards, namely: LTE, UMTS, and GSM, while the incumbent system is a DVB-S2 system. This means that one classifier can detect and classify the interference signals instead of different classifiers for different signals. Regarding *Anti-Jamming* capabilities, in [93], the authors studied *k*-nearest neighbor (KNN) ML algorithm in detecting and classifying chirp jamming signals to GNSS systems, which is a challenging task using traditional approaches due to their variety, that requires post-processing of the signals and human-driven analysis, thus limiting their timely awareness and response jamming attacks. On the other hand, ML-based approaches can detect and classify the jamming signals automatically, which makes them better candidates in scenarios where timely awareness and response are required.

9) Intersatellite Synchronization:

a) *Motivation:* Inter-satellite synchronization is a new open challenge triggered by the boom of small satellites grouped in Distributed Satellite Systems. Currently, Distributed Satellite Systems are considered in multiple applications, from space communications to remote sensing. In the first case, the main advantages of distributed over monolithic satellite communication systems could be the increment of the system capacity and the reduction of the transmission power. Meanwhile, distributed Earth observation satellite missions can achieve higher spatial resolution requirements than single-platform architectures. However, the Distributed Satellite Systems should satisfy strict time and phase synchronization

requirements to provide such performance enhancement.

b) Description: Synchronization algorithms are critical blocks in any communication system. They are key in the establishment of the communication link and have strong impact on the system performance. Conventional wireless communication systems implement synchronization functions in an iterative way, starting from the frame synchronization, then time synchronization and carrier synchronization. The first of these steps allows the receivers to calibrate several parameters such as accuracy of its local oscillator frequency, the desired timing phase and frequency of the digital samples, receiver's gain, etc [247]. To this end the transmitters send pre-established signals that are known by the receivers. In addition, the time synchronization allows the receivers to sample each symbol at the perfect time while the carrier synchronization refers to the estimation and compensation of the carrier frequency and phase offset introduced by the channel.

During the frame synchronization step, the signal detector searches for the signal energy peak by correlating a known transmitted symbol sequence with the received signal samples at different sampling times. These algorithms are known as data-aided synchronization methods. Thus, the signal peak found is used as the initial timing phase estimate for the fine-tune carrier phase and frequency synchronization.

Timing synchronization is usually achieved by forming a feedback loop and thus it is often called Timing Locked Loop or timing Control Loop. Once the Timing Locked Loop has converged, its output is synchronous to the remote transmitter timing, i.e., the remote transmitter's channel symbol clock [247]. Carrier synchronization also rely on feedback loops that can be implemented in the analog or digital domain. These systems are based on phase or frequency offset detectors, loop filters and Voltage-Controlled Oscillator or Numerically-controlled Oscillator [248]. There is a huge number of data-aided and non-data-aided algorithms for frequency and phase offset detectors [249]. Generally, non-data-aided methods are used when the frequency offset is comparable with the symbol rate. After the frequency offset is reduced, higher accuracy can be achieved using data-aided methods.

c) Conventional Solutions and Issues: Conventional solutions for inter-satellite synchronization can be classified as Closed-loop or Open-loop methods based on the feedback from the receptor or target. The Closed-loop methods require a communication channel to transmit the feedback information between the external node and the Distributed Satellite System, which makes these methods more suitable for communication systems. Whereas in Open-loop, synchronization is achieved without the participation of any node other than the distributed satellites, as typical remote sensing applications require. In addition, we consider the communication between the elements of the Distributed Satellite Systems as an additional way of classification. It includes (I) synchronization algorithms based on the exchange of information among the distributed satellites, which can be done as a two-way message exchange or (II) as a broadcast, i.e., one-way communication. (III) Another option is to synchronize without any communication among the elements of the Distributed Satellite Systems by

relying the control on a node out of the Distributed Satellite System. Both classifications can be superimposed or combined

d) Proposed ML Solutions: Even when there has not yet been any distributed satellite mission performing inter-satellite synchronization with Machine Learning techniques, they have been used to address other problems related to synchronization in end-to-end-communication systems. For instance, frame synchronization [94], [98], sampling frequency and time offset compensation [95], carrier synchronization [96], and the characterization of phase noise [97], [250] have shown promising results by using Machine Learning techniques.

A Convolutional Neural Network-based synchronization model with softmax activation function was proposed in [98] to achieve better frame synchronization and compensate for the sampling timing errors caused by sampling time offsets. The Machine Learning detection results showed an increment of 2 dB over the direct correlation detection in terms of correctly detecting the actual position of a frame header. Furthermore, the traditional frame synchronization techniques based on maximum likelihood and correlation present hardware implementation constraints related to frequency deviation problems which can be avoided using Machine Learning techniques such as multi-instance learning to solve the frame synchronization problem under different frequency ranges [94].

In addition, Machine Learning techniques increase the accuracy of the amplitude and phase noise characterization of frequency references, which are essential elements for the synchronization of Distributed Satellite Systems. In this regard, a Bayesian filtering-based framework combined with expectation-maximization was used to characterize the noise of lasers in [97]. The carrier synchronization was experimentally demonstrated using the proposed framework in low Signal-to-Noise Ratio scenarios.

On the other hand, time synchronization attacks to cyber-physical systems, which mainly focus on modifying the measurements' sampling time or the time stamps, can be effectively detected by residual-based lousy data detection and conventional supervised Machine Learning-based detectors. However, new Machine Learning-based solutions have been proposed in [95]. The "first difference aware" Machine Learning classifier presented in [95] can detect two types of time synchronization attacks, namely, direct time synchronization, which only modifies some time stamps, and stealth time synchronization attack, which modifies all the timestamps at a specific time. The First Difference Machine Learning (FDMachine Learning) technique utilizes the backward first difference of the time-series data to process the input data stream before employing a Machine Learning method.

In the context of integrated 5G and satellite communications utilizing Orthogonal Frequency-division Multiplexing, authors in [250] proposed a Cyclic-Prefix-based multi-symbol merging blind synchronization algorithm to enhance the timing accuracy. Besides, the authors proposed an improved synchronization method to accurately correct time-frequency errors in 5G SATCOM integration scenarios. Another example of Machine Learning used in frequency hopping-frequency division multiple access satellite communications is the challenging synchronization between Dehop-rehop Transponder and ground

equipment using different hopping sequences for the uplink and downlink. In this regard, authors in [96] proposed a novel Machine Learning-based method to synchronize the Frequency Hopping signal by utilizing serial search for coarse acquisition and Long Short-term Memory network for fine acquisition. The main objective of the proposed method was to reduce the synchronization time.

10) Network Security-Device Authentication:

a) *Motivation:* Satellite Communication Systems can be an attractive target for attackers due to their importance in global communication systems. Thus, a hacker can launch cyber-attacks at scale by targeting a few key components that constitute the backbone of the global Satellite Communication System. The risk of attacks on Satellite Communication System is further aggravated due to a wide threat surface caused by the ease of spoofing wireless devices and limited on-board computing power, which prohibits the implementation of complex cryptographic techniques [251]–[253]. One of the most common attacks in this context is spoofing where a hacker impersonates a legitimate device in the network to provide tampered or falsified information to the network users [252]. These attacks put Satellite Communication Systems on a high risk of service outage, ransomware attacks, and even political and financial losses [254]. With the advent of distributed learning techniques, e.g., Federated Learning, device authentication becomes vital to deter various attack vectors including model and data poisoning, membership inference, and reconstruction attacks by the unauthorized agents [255].

b) *Description:* The first line of defense against these spoofing and impersonation attacks is the device authentication procedure [256]–[258]. In traditional approaches, the device authentication and tracking may be carried out by the Network Control Centre. Alternatively, individual devices may carry out additional mutual device authentication procedures to include another layer of security in case the Network Control Centre is compromised [259].

c) *Conventional Solutions and Issues:* Traditional approaches for device authentication may be based on Secret-key Cryptosystem [258] or Public-key Cryptosystem [260]. The server maintaining the secret-key tables in Secret-key Cryptosystem-based authentication schemes become an attractive target. On the other hand, high complexity of Public-key Cryptosystem-based schemes is a major challenge for their use in low-powered devices. Some other issues in these approaches include high transmission delays and session dependence, i.e., if the key of a session is compromised the entire subsequent communication and authentication is compromised. Recent approaches for device authentication include blockchain [261], orbit-based authentication using time-difference of arrival [262], and hierarchical group key distribution [263]. Some common demerits of these traditional approaches include model dependence, low reliability, and low cost efficiency due to an increasing number of network nodes.

d) *Proposed ML Solutions:* In the recent years, intelligence-based approaches for authentication have emerged providing cost efficiency, high reliability, and a fairly high level of model independence [264]. Another major advantage of Machine Learning-based approaches is the departure

from session-based authentication to the *continuous* protection throughout the duration of communication.

In [139] and [140], authors proposed a supervised Machine Learning approach for detecting Global Navigation Satellite System spoofing signals. The authors used code, phase, doppler, and signal strength measurements in different frequency bands as data features sampled at 1Hz. The spoofing was performed as an intermediate timing attack aimed to affect the receiver's clock divergence by emulating a satellite clock drift. The specific algorithm utilized in this approach was Support Vector Machine, which resulted in a high accuracy of more than 98.5% for different datasets.

The authors of [141] investigated similar spoofing attacks on Global Navigation Satellite System. They compared the performance of multiple Machine Learning methods, namely, Support Vector Machine with radial basis function, k -nearest neighbors, adaptive boosting (AdaBoost), decision tree, and random forest. They found that the simple classification based on decision tree outperformed other methods in their considered scenario providing high accuracy with a very low false-alarm ratio.

Machine Learning-assisted antenna fingerprinting is another attractive approach for authenticating the wireless communication devices. This approach exploits the antenna manufacturing defects that are unique to each antenna and cause identifiable waveform imperfections [142]–[144]. One major challenge in antenna fingerprinting is degradation of antenna due to different environmental factors that might alter the antenna fingerprint [265]. The authors of [145] used a Recurrent Neural Network to capture temporal variations and provide robust antenna fingerprinting of different Low Earth Orbit scenarios with a high accuracy of 99.34% for 198 days despite antenna degradation.

These Machine Learning-assisted physical layer authentication methods complement the device authentication employed at higher layers. Highly reliable and model-free gate-keeping against spoofing and impersonation attacks at the physical layer greatly reduces the number of possible attacks that can be carried out at the higher layers. This makes the satellite systems more reliable and difficult to intrude by the malicious agents.

11) Network Security-Quantum Key:

a) *Motivation:* The wide coverage and broadcasting nature of satellite signals make them susceptible to eavesdropping and information leakage to malicious parties [266], [267]. Data security protocols aim to eliminate the possibility of information leakage to eavesdroppers. Two main approaches to ensure data security are 1) physical-layer security and 2) data encryption in higher layers [268], [269].

b) *Description:* physical-layer security methods operate under the premise that the wiretap channel to the eavesdropper is inferior to that of the intended receiver. To this end, the randomness and time-varying nature of wireless channels are exploited [270], [271]. New techniques are being developed that utilize multiuser Multiple-Input Multiple-Output channels to ensure that eavesdropper ends up with a worse channel than the intended users due to interference [272], [273]. On the other hand, data encryption in higher layer relies on a

secret key shared between transmitter and receiver to encrypt data. Data encryption in higher layers can be preferred over physical-layer security due to its modular nature and independence from operating/channel conditions. However, the biggest demerit of data encryption techniques is that they can be compromised if an eavesdropper has strong computational powers or a reasonably big quantum computer [266], [274].

c) Conventional Solutions and Issues: The advent of quantum key distribution methods has eliminated the biggest demerit of key-based encryption by providing unconditionally secure keys that can ensure information-theoretic security [275]–[281].

In the context of satellite communication, quantum key distribution methods become more attractive due to a unique symbiotic relation between the two. That is, not only quantum key distribution enables to achieve information-theoretic security of satellite communication but also Satellite Communication Systems have a vital role in achieving global-scale quantum key distribution [282], [283]. To demonstrate the approach feasibility, the work in [284] presents the configurable finite-size secret key rates that protocols with continuous variable quantum key distribution may achieve for both downlink and uplink. Quantum key distribution systems are very sensitive to the environmental factors, e.g., phase drift caused by internal as well as external factors, channel noise, and photon losses. Satellite channels can have high variability in these factors due to weather, satellite mobility, and imperfect internal devices. Accurate prediction and tracking of these factors is crucial to maintaining correct operation of quantum key distribution systems. Traditional approaches for tracking the relevant parameters perform a scan before transmission, which amounts to a significant “down-time” for calibration of quantum key distribution equipment [148], [285].

d) Proposed ML Solutions: The authors of [148] employed Long Short-term Memory network to predict relevant physical parameters in real-time and actively control quantum key distribution devices to ensure a stable operation over the course of 10 days. The performance of resulting system was comparable to that of traditional “scanning-and-transmitting” schemes in terms of quantum-bit error rates and key transmission rates. Lack of down-time for scanning resulted in a high system efficiency over the course of continuous operation. The authors of [149] targeted the quantum key distribution over free space on moving platforms including satellites. They employed a neural network to predict optimal parameters for operation of quantum key distribution in variable environments. The resulting system required low computational power in its operation and was demonstrated on hardware devices including Raspberry Pi 3 and a mobile phone consuming less than 5 W. As compared to a brute-force search for the optimal parameters, the proposed approach provided up to two to four orders of magnitude speedup while providing a 95% to 99% of the optimal secure key rate for a given protocol.

Machine Learning was employed in [147] for real-time selection of optimal quantum key distribution protocol in a given environment characterized by dark count rate, efficiency of single photon detectors, misalignment error rate, and transmission distance. The authors employed random forest

and compared it with other approaches e.g., support vector machine, K-nearest neighbors algorithm, and convolutional neural networks. Testing results showed an accuracy of over 98% for selecting the most optimal quantum key distribution protocol in a given scenario by random forests.

Carrier recovery is another critical task in modern continuous-variable quantum key distribution systems. The authors of [146] employed an unscented Kalman filter for carrier recovery in noise over a 20 km fibre-optic link. The experimental results showed low variance and high stability in excess noise even at low pilot powers.

These Machine Learning-based operations and optimization of quantum key distribution systems demonstrate a high potential of Machine Learning in achieving a global-scale quantum key distribution network. Considering the highly variable nature of satellite communication channels, it is clear that Machine Learning-based solutions provide an efficient solution for system-parameter tracking and prediction, which in turn achieves optimal performance of quantum key distribution that is an important component for realizing the goal of unconditionally secure global-scale communication over Satellite Communication Systems.

12) Precoding:

a) Motivation: Multibeam system design is based on the concept of frequency reuse, and currently, multibeam systems typically exploit 4-color frequency reuse schemes to avoid aggressive levels of co-channel interference. In this sense, more aggressive frequency reuse schemes, such as full frequency reuse, can fully exploit the available bandwidth and achieve significantly higher performance levels. However, we need several solutions to increase system throughput when providing terabit connectivity over satellite systems. In this context, interference between adjacent beams strongly limits the overall system throughput. Consequently, it is of utmost importance to apply advanced interference mitigation techniques at the receiver, e.g., multiuser detection, or at the transmitter, i.e., precoding.

Due to the popularity of multiuser multiple-input multiple-output (MU-MIMO) techniques in terrestrial communications, together with the launch of the superframe structure in the DVB-S2X standard, the SATCOM community has started to evaluate and implement precoding techniques in multibeam SATCOM systems.

While the concept of precoding in satellite networks have widely studied in theoretical scientific papers, e.g. [286], an actual live-based demonstration supported by the European Space Agency (ESA) has been carried out in ESA project LiveSatPreDem (2020), validating the feasibility of such technique considering the recently amended DVB-S2X specifications to support it [287], [288]. Precoding is generally embedded at the gateway, thus keeping the complexity of the payload and user terminal (UT) infrastructure low.

b) Description: The most popular precoding design in satellite communications is the low-complexity Regularized Zero Forcing precoder, which is designed as follows:

$$\mathbf{W}_{ZF} = \eta \cdot \tilde{\mathbf{H}}^H \left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \alpha \mathbf{I} \right)^{-1}, \quad (3)$$

where η is a normalization factor ensuring that the output signal stays within the power limits. For a system-level power constraint, $\eta = \sqrt{P_{\text{tot}}/\text{Trace}\{\mathbf{W}_{ZF}\mathbf{W}_{ZF}^H\}}$ such that the sum of the norm of the precoder vectors in \mathbf{W}_{ZF} is equal to P_{tot} , i.e. $\sum_{n=1}^N \|\mathbf{w}_n\|^2 = P_{\text{tot}}$. The regularization factor α is an arbitrary number, which is usually considered equal for all users and proportional to the inverse of the expected signal-to-noise ratio [286].

In Regularized Zero Forcing, the precoding matrix is calculated once per frame period, and the computation of \mathbf{W}_{ZF} is mainly driven by the matrix inversion process.

c) Conventional Solutions and Issues: The asymptotic complexity when assuming the Cholesky decomposition method is given by $\frac{1}{3}N^3$ flops [289], where N is the number of beams. Even if the number of beams is high, the standard is limited to 32 unique Walsh-Hadamard (WH) sequences. Therefore, the receiver can discriminate the signals coming from the 31 nearest interfering beams, and the complexity is upper-bounded by $\frac{1}{3}32^3$. It is to be noted that such inverse operation needs to be re-calculated as soon as the scheduled users change (i.e. when the channel matrix changes). All in all, such computation is considered computationally intensive for conventional systems. Different approaches have been investigated to mitigate this challenge, e.g. precoding matrix nullification, which consists of nulling-out irrelevant coefficients.

d) Proposed ML Solutions: While ML for multi-antenna transmission has been well explored in terrestrial systems [290], [291], there are very few works published on the SATCOM domain. The European Space Agency (ESA) funded the 'MLSAT - Machine Learning and Artificial Intelligence for Satellite Communication' in 2020 [292], where the GEO satellite precoding matrix calculation was studied and evaluated using a two-step procedure involving autoencoders. The first step tries to learn the relationship between the real channel and the estimated channel matrix, i.e. understand the proper parametrizations to cover all the relevant information about the real channel matrix and the estimated channel matrix, respectively. The second step starts with the already inverted, estimated channel matrix from the simulation framework and predicts the alternative precoding matrix. Therefore, the second step does not need to learn the complex matrix inversion performed in the simulator framework but can focus on the remaining projection. Once the individual steps are solved, the combined approach estimates the machine learning-based precoding matrix directly from the estimated channel state matrix of the terminals. According to ESA MLSAT project reports, the proposed AI-based approach showed a 3dB gain in mean SINR.

Some works like [80], [99] have considered precoded GEO satellite systems and learning-based techniques but not directly to calculate the precoding matrix but to optimize the resource allocation assuming a fixed linear precoding matrix.

Boosting the precoding performance with proper user scheduling (i.e. clustering) is discussed in Section IV-B1 of this survey.

There is potential for ML techniques to improve the conventional precoding in SATCOM systems, possibly exploiting

the advances made in the terrestrial domain, e.g. [293].

13) Link Quality Prediction:

a) Motivation: Satellite services are being deployed in higher frequency bands to provide higher capacities. Ka and Q/V show substantial excess attenuation in atmospheric events, especially rain. As a result, techniques for mitigating satellite link unavailabilities, such as automatic power control, gateway switching and adaptive modulation and coding schemes, have been studied in recent years. As the implementation of the mentioned fade mitigation approaches requires a non-negligible delay, link quality prediction techniques become mandatory. Indeed, for the efficient deployment of pre-emptive fade mitigation techniques, satellite systems shall accurately track the link quality.

While the terminal segment predictions require a time window similar to the propagation delay, the ground segment usually needs longer time windows for performing mission-level operations such as gateway switching. Bearing this in mind, while the terminal channel predictions can rely on former and current channel values, addressing longer time windows usually require external information such as weather predictions. Channel samples become uncorrelated with time lags similar to a few seconds.

Mobile satellite services vendors have addressed the terminal channel variations for many years as it was the main shortcoming of terminal capacity. Those systems generally rely on margins so that modulation, coding scheme, and transmitted power are selected considering a potential decay of the channel magnitude. On the contrary, ground segment systems, whose cost severely relies on high power amplifier and antenna size, focus on switching mechanisms when their fade mitigation techniques do not circumvent the channel degradation. We describe how this later operation is performed in the next paragraph.

b) Description: As mentioned earlier, ground segment equipment is designed considering a certain channel decay tolerance and targetting a constraint cost of the teleports based on the target EIRP. To offer a 99.99% ground location availability, teleports are located in areas with annual rain values below a certain threshold. The automatic power control can combat up to 10-12 dBs of fade. Satellite operators must rely on redundant gateways if extreme atmospheric events exceed the mentioned 12 dBs. This is traditionally coined as $N + P$ gateway diversity [294]–[296] where N nominal gateways can be eventually replaced by any of the P redundant ones. With this, overall mission outage is substantially reduced.

This ideal overall outage mission target cannot be attained in the system's regular operation due to the delays in performing the gateway switch. While the delay between the gateway switch decision has been made and the actual switch taking place requires a few minutes in nowadays network operation centres, after the switch, all terminals have to resynchronize again with the new ground station leading to a non-negligible additional delay. This operational delay is critical as outages with a temporal duration below this delay shall not involve a gateway switch.

Bearing this in mind, operators have to predict the gateway's link quality to reach the ideal capacity outage value conceived

with the $N + P$ ground mission design. In particular, it is relevant to determine atmospheric events that yield strong fading values with a duration larger than the operational gateway switching delay.

c) Conventional Solutions and Issues: The propagation community in different papers carefully addressed the prediction of channel values [297]–[299]. Mathematical models behave well in short-term predictions and offer good results in temporal horizons below one minute. However, the operational gateway switch requires a few minutes, so channel predictions shall consider a longer prediction window than those in the mentioned mathematical models.

The first attempt to consider several minutes of gateway prediction can be found in [300]. This seminal paper utilizes a Bayesian approach to consider a priori forecast weather data to attempt to predict the channel values in several minutes time windows. Prediction accuracy is linked to weather forecast accuracy as rainfall rate is the determining feature of strong satellite link fading events.

d) Proposed ML Solutions: Due to the exponential interest in ML solutions, channel prediction and, in general, radio propagation modelling techniques have been revisited in [301]. The goal is to move from a model-based approach to a data-based approach which can eventually embrace high-complex scenarios while preserving a relatively low computational complexity.

This is the case of the recent works in [100]–[103]. The goal of these works is to rely on deep learning time series prediction known approaches and apply them to channel prediction. As a general statement, the model boils down to

$$g(\mathbf{h}, \mathbf{w}; \theta) = \hat{h}', \quad (4)$$

where \hat{h}' is the predicted channel value, \mathbf{h} is a vector containing the staked previous channel values, \mathbf{w} is the external weather features, and θ is the deep neural parameters to be learnt in order to perform the prediction efficiently. Remarkably, the external weather information results to be a critical aspect in the works [100], [101], [103], which can be either a set of RADAR images or the current rainfall rate.

While (4) shows a good performance in below few minutes time horizon prediction, addressing longer time windows could eventually require weather predictions. This is reported in [302], where external rainfall prediction techniques are utilized as a priori information for performing the channel quality prediction in long-term periods.

14) Predistortion:

a) Motivation: Optimizing size and weight of the communication payload in a spacecraft is essential for reducing launch and operational costs, as well as for improving the utilization of the rather limited available space on board [303]. Conventional payload architectures usually rely on the separation among the antenna feeding segments, such as Power Amplifier, transmission line, and passive unit cell. This scheme has many drawbacks, including losses in each interconnection between segments, increased payload size, high cost, high risk of failure, and high noise figures. For all these reasons, active antennas, which centralize the active device, transmission

lines, and radiating element [304] in a single device, are a promising solution for future High Throughput Satellites.

b) Description: Current and future High Throughput Satellites payloads need to simultaneously generate hundreds to thousands of beams over the coverage area, requiring a very large aggregated capacity (in the order of a few THz). In the context of Digital Beam Forming with Direct Radiating Active Antennas, the only feasible solution for the amplification of the signals is the use of Solid-State Power Amplifier which are inherently less efficient than the conventional Traveling-Wave Tube Amplifiers typically used in Satellite Communication.

Radiofrequency Power Amplifiers are essential building blocks in Satellite Communication payloads. They are required to drive the transmitting antennas with an output power high enough to overcome all the losses and attenuation of the RF signals on their way to the receiver. In fact, the Power Amplifiers are among the most power-hungry components in communications systems. Therefore, Power Amplifiers should work with high efficiency, consuming as little energy as possible in addition to that to be delivered. This is even more critical in power-limited devices, such as battery-operated handhelds and satellite payloads. In this regard, Radio Frequency Power Amplifiers are most energy-efficient when operated at their highest output power, close to saturation. This gives rise to nonlinear effects upon the transmitted signals. The nonlinearities in the transmitter tend to degrade performance due to gain compression, clipping, phase distortion, intermodulation, adjacent channel interference, spectral regrowth, increased out-of-band emissions, etc, thus reducing the maximum throughput delivered by the payload. Particularly in digital communications, nonlinear distortion causes constellation warping and clustering, thus significantly complicating signal reception/detection.

c) Conventional Solutions and Issues: Selecting the Power Amplifiers' operation point poses significant challenges for properly balancing efficiency and linearity [305]. Limiting the output signal's Peak Envelope Power away from the Power Amplifier's saturation level (i.e., to back-off the output) improves linearity at the cost of reducing the amplifier efficiency. This approach is particularly inconvenient when dealing with high Peak-to-Average Power Ratio values, as those present in many current multicarrier-modulated signaling to achieve higher power spectrum efficiencies [306]. Choosing a conservative Output Back-Off value to accommodate the Peak Envelope Power in the intrinsic linear part of the amplifier's characteristic implies moving the average operating point further away into the low energy efficiency regime. In other words, the higher the back-off, the lower the Power-Added Efficiency of the Power Amplifier. This justifies the need for applying additional linearizing techniques to energy-efficient but highly nonlinear amplifiers.

The basic design principle behind pre-distortion is to measure the non-linearity at the PA output side and provide a complementary non-linearity at the PA input side via analog methods in the RF domain or digital in the base band domain. SATCOM systems with digital signal processing and active beamforming capabilities could employ Digital Predistortion techniques as part of the on-board signal processing. Applying

Digital Predistortion will either increase the amplifier's useful output power or increase the PAE for the same linearity requirement. A moderate increase in processor power requirement could be traded-off against either increased Power-Added Efficiency, leading to a reduction of the mass of the thermal management system, or reduced beamforming and intermodulation interference at the same Power-Added Efficiency. These improvements have the potential to provide substantial power and mass reduction in different missions (e.g., MEO/GEO). Additionally, the use of Digital Predistortion techniques will allow the use of different types of waveforms, which are currently prohibitive for Satellite Communications applications, such as the Orthogonal Frequency-division Multiplexing based 5G-NR standard.

d) Proposed ML Solutions: Machine Learning algorithms are attractive for Digital Predistortion domain due to their ability of modeling non linear behaviours. While strictly not related to Satellite Communication, relevant works on Machine Learning and Digital Predistortion has been published in literature in recent years. Differentiating based on the dataset form used as an input for Deep Neural Network, the work in [307] proposes a novel neural network model based on a JANET architecture, [308] includes the physical features of the Power Amplifier in the input of an MLP neural network Digital Predistortion. Paper [309] presents a Digital Predistortion scheme based on a dual-time delayed neural network (TDNN) learning architecture. [104] proposes a Neural Network to compensates the nonlinearity given by the onboard satellite payload, considered as a whole and including Power Amplifier non linearity, and linear distortion caused by equipment such as filters and multiplexers.

While the application of Machine Learning techniques in the field of Digital Predistortion techniques in the field of satellite communications is still limited, we expect a widespread rise in the next future.

15) Coding:

a) Motivation: Coding schemes are a crucial element of modern digital communications systems. Convolutional Codes, Turbo codes, Low-Density Parity-Check (LDPC), Reed-Solomon or Polar codes are examples of Forward Error Correction algorithms used in different communications standards such as DVB or 3rd Generation Partnership Project (Fourth Generation, Fifth Generation and future Sixth Generation) for channel coding. Forward Error Correction decoding algorithms are typically the element of the receivers that require more complex computational capacity regarding resources. This computing problem is increased and aggravated in high data rate communications. Reducing the implementation complexity in communication receivers and the power consumption of FEC algorithms has become essential to evolving future communication systems to increase throughput. A demand for enormous resources in current communication receiver implementations. These resources can be the number of computations or processes to perform in each task, the computation time or the number of hardware components used in the implementation as logic gates. From the point of view of SATCOM, reducing power consumption and processing time is vital due to its repercussion on the payload size and,

consequently, on the cost. Also, satellites must reduce latency and delay to integrate with terrestrial networks. Reducing the processing time favours the previous reduction or compensates for the least latency and delay innate to SATCOM's nature. All these challenges are the objective for future regenerative payloads, reducing complexity in HTS and, on the other hand, reducing the sizes of the payloads and improving the NGSO small satellites.

b) Description: Coding schemes are grouped into source and channel schemes. Source coding removes redundancy from the bitstream. Instead, channel coding adds redundancy to the bitstream; this redundancy or repeating of code is introduced to reduce SER.

c) Conventional Solutions and Issues: Typical decoding relies on maximum a posteriori decoding, which consists in computing the probability that a specific bit was 0 or 1 and selecting the hypothesis with a higher probability. The two main drawbacks of a typical decoding approach are (i) the computational complexity and (ii) the unknown distribution of the channel noise. The classical Viterbi algorithm is used for decoding a bitstream that has been encoded using FEC based on a convolutional code. The Hamming distance is used as a metric for hard-decision Viterbi decoders. The squared Euclidean distance is used as a metric for soft decision decoders. Regarding power consumption, several studies have evaluated the consumption of FPGA as hardware to support LDPC coding [310]–[313]. In the DVB standard with a coding rate of (64800, 29160), the power consumption by LDPC is 12.92W for a throughput of ~ 4 Gbps, which for a satellite with a maximum power consumption of 50W, the decoding process account for 25.84% of the total of the consumption of the payload for a single reception chain [310]. In the worst case, for small satellites like CubeSatKit, decoding at such a throughput rate is impossible since this task's consumption is larger than the power budget (2W) [311]. In future HTS, where there could be up to thousands of parallel receiving chains on board the regenerative satellite, power consumption is increased with the number of beams up. For example, suppose an HTS satellite has a capacity of 200 beams, each being decoded by a FEC decoding chain with power consumption similar to SmallGEO. In that case, the total power consumption will be proportionally incremented by 200, $\sim 25.8\%$. Therefore, to overcome these challenges, it is possible to consider a ML approach to learn from data.

d) Proposed ML Solutions: The application of ML to FEC decoding is generally restricted to shortcodes due to the exponential training complexity. For instance, a message with " k " bits gives 2^k possible codewords. Fortunately, the ACM codes of DVB-S2X are generally limited, favouring such a scenario. Two typical decoding approaches based on NN are data-driven and model-driven schemes, as shown in [106]. One of the earliest works on ML for FEC decoding is [107], which showed that the belief propagation decoding algorithm might be equipped with learnable multiplicative weights and trained as a NN to improve error correction performance. Recently, any academic works have been published on the topic with the objective of wireless communication receivers, e.g. [108]–[110]. Furthermore, [105] presents a novel approach

for decoding 5G data frames (a combined demapper and decoder) based on a combination of autoencoders and deep neural networks. The performance results of the proposed system compared with traditional implementations based on constellation demapping and LDPC decoders. The proposed approach can obtain a gain of 3 dB of SNR, which may be expanded by considering spatial-domain diversity through a MIMO approach. This scheme proposed in [105] may be a candidate for satellite integration in the 5G network.

B. Medium Layers

1) User Scheduling:

a) Motivation: Current multi-beam high throughput satellite systems use popular Orthogonal Multiple Access (OMA) schemes such as multi-frequency Time Division Multiple Access (MF-TDMA) to exploit frequency flexibility. Furthermore, the traditional four-colour frequency reuse scheme is adapted to handle adjacent beam interference. The term "colour" in the four-colour frequency reuse scheme refers to the four different sets of non-overlapping channels assigned to each cell, represented by different colours. This scheme helps minimize interference between adjacent cells by using different sets of channels, allowing for efficient use of the available frequency spectrum.) However, when non-orthogonal users are scheduled together, even with the four-colour frequency reuse scheme, system performance can degrade due to user interference. This interference can arise due to the non-orthogonal use of frequency resources, which can lead to overlapping signals and reduce the quality of the received signals. Nevertheless, the extent of performance degradation will depend on various factors, such as the number of users scheduled together, the modulation and coding schemes used, and the amount of interference in the system. Hence, the system performance is agnostic when non-orthogonal users are scheduled together. As a result, scheduling orthogonal users is crucial to reducing interference and maximizing the capacity offered.

In recent years, the full frequency reuse scheme has replaced the traditional four-colour frequency reuse scheme. Precoding has been used as an effective co-channel interference mitigation technique to enhance satellite spectral efficiency. In such an aggressive frequency reuse environment, orthogonal user scheduling is increasingly significant in precoding performance. There is a significant reduction in throughput when adjacent beams have collinear user channel vectors. Hence, a proper user scheduling strategy must be designed to select users with orthogonal channel vectors to be served simultaneously to achieve optimal precoding performance. Similarly to terrestrial systems [314], [315], multibeam precoding systems are sensitive to user scheduling. As a general statement, the attainable system rates increase when co-scheduled users have close-to-orthogonal channel vectors. In any case, user data flows are subject to end-to-end delays, frame encapsulation problems and user traffic prioritization, which makes this channel-vector-based scheduling challenging to be deployed in a real system.

b) Description: A user scheduler establishes the relation between users, beams and scheduling time. In the unicast case, one user is scheduled per beam at a given scheduling time. As a result, the DVB-S2X [316] defined XFECFRAME contains data that belongs to a single user. The scheduler's task is to schedule one user per beam for each scheduling period by considering factors such as orthogonality (better spectrum efficiency), fairness, and demand satisfaction.

As a result of the huge number of users to be served in a single beam in broadband satellite scenarios, frames are filled with data from more than one user. In this context, the resulting precoding operation becomes the so-called multigroup-multicast [317]. This communication framework requires users embedded in the same frame to have similar channel vectors and C/I . This latter requirement comes from frame encapsulation restrictions that impose that all users served on the same frame will be assigned the same modulation and coding scheme.

In the multicast scenario, a group of users are scheduled per beam at a given scheduling time. As a result, multiple users share the same physical frame defined by DVB-S2X and are usually grouped based on similar SINR values. The challenge in multicast scheduling consists of two steps: the first is to group users using the same DVB XFECFRAME, and the second is to schedule these groups on time. The first step groups highly non-orthogonal users to ensure that all users sharing an XFECFRAME can decode the frame correctly. This is because the modulation and coding scheme will be selected based on the SINR of the weakest user. In the second step, the multicast scheduler schedules one group per beam at every scheduling time to make the scheduled groups' channel vectors as orthogonal as possible. Furthermore, the second step of the multicast scheduler should also consider fairness and demand satisfaction.

c) Conventional Solutions and Issues: Scheduling users in multicast multibeam precoding is a coupled problem. Indeed, the ground station can only compute the precoding matrix once the scheduled users have been determined. To avoid iterative methods, the scheduler has to provide scheduling policies without knowing the resulting attainable rates. In this context, scheduling must involve heuristic approaches that yield to resulting scenarios where users within the same group have similar C/I yet different channel vectors with co-scheduled users served at other beams to attain high data rates.

As the line-of-sight component strongly dominates satellite channels, one could resort to latitude and longitude positions to perform the scheduling. This is a substantial complexity reduction of the scheduler as it no longer requires access to reported CSI values. In [318], the authors propose a location-based scheduling algorithm that groups users with similar geographical locations and assigns them to the same beams. An alternative approach is suggested by the authors in [319] who suggest using Euclidean distance to correlate channel vectors and impose channel orthogonality.

In satellite systems, cellular scheduling can be used to increase the number of users that can be supported and improve system performance. However, as the number of cells increases, the likelihood of co-channel interference also

increases. From the authors of [320], it is also possible to avoid interference using graph theory in cellular networks to reduce intercell interferences (ICIs) using the Least Beam Collision (LBC) algorithm. This method avoids the simultaneous scheduling of two adjacent cells that might interfere with each other by refraining from scheduling them simultaneously.

Despite geographical scheduling results into the same result as when using perfect CSI in clear sky conditions and perfect satellite feed element frequency precompensation, differences appear when a close-to-real operation is analyzed and fading is considered. In those cases, grouping users considering their Euclidean distance seems to behave well [321] even when mobility is considered [322]. Furthermore, the authors of [321] propose a user grouping scheme based on random preprocessing and the previously discussed Euclidean norm.

A method to minimize the simultaneous scheduling of users in interference beams is also described in [323]. This method uses partial channel state information (CSI) to avoid scheduling users in interference beams simultaneously. A study in [324] investigates user scheduling for multicast transmissions with full frequency reuse and multicast precoding. Some authors have, however, enhanced their study by using the Euclidean norm and cosine similarity to examine channel characteristics. Also, in [325], the authors sequentially select users with orthogonal channel vectors based on the cosine similarity metric. Consequently, channel orthogonality and semi-orthogonal scheduling, as defined in [314], have been adopted widely by the satellite industry to address scheduling challenges. On a different approach, the authors of [326] use user scheduling to achieve better demand satisfaction by considering both co-channel interference and user demands. Furthermore, considering the joint nature of precoding and user scheduling, the works in [327]–[329] suggest an alternative solution involving suboptimal solutions that approach stationary points.

d) Proposed ML Solutions: In recent years, unsupervised machine learning techniques have been used to automatically group users into clusters with similar characteristics, which helps make the scheduling process more efficient. These clustering techniques can also take into account temporal variations in user traffic, which is particularly beneficial in the context of multibeam satellite systems. In particular, [111] focus on user clustering for multicast precoding in multibeam satellite systems. Two clustering algorithms have been proposed in this context: (a) fixed-size clustering, aimed at minimizing the impact of outlier users, and (b) variable-size clustering based on the K-means++ algorithm. Through numerical simulation, it was shown that the rate achieved with the proposed fixed-size clustering algorithm consistently outperforms the solutions available in the literature, and almost meets the upper bound performance. Also, variable-size clustering is more efficient than existing solutions. However, fixed-size clustering should be preferred when the number of users increases and the channel coefficient similarity increases. Nevertheless, the study does not consider the impact of non-uniform traffic requests on scheduling decisions.

Additionally, [112] approaches the user scheduling problem in the context of clustering users as a function of the feature

vector of the users, taking into account not only their geographical position but also channel information. Eventually, they evaluate three different algorithms based on UL: K-means (Km), Hierarchical clustering (Hc), and SelfOrganization (SO). However, further investigation is necessary to assess the wider implications of the research. This is because several issues remain to be addressed, such as finding the optimal number of clusters, studying other user characteristics for clustering, and evaluating its impact. In this sense, unsupervised learning techniques are ideal for exploring the structure of evaluated users.

2) NOMA Multiple Access:

a) Motivation: The current satellite communication systems use popular Orthogonal Multiple Access schemes such as Time Division Multiple Access, Frequency Division Multiple Access, Code Division Multiple Access and Spaced Division Multiple Access, where orthogonal users share time, frequency, code and space, respectively. However, the performance of such OMA schemes is confined by the physical limitations of the resources. Conversely, Non-orthogonal Multiple Access differs fundamentally from OMA in that all users share the same resources (time and frequency), but their power levels are distinct. Hence, the concept of NOMA has recently received considerable attention due to its ability to improve power-domain flexibility in managing resources and achieve higher spectral efficiency than Orthogonal Multiple Access. Furthermore, Non-orthogonal Multiple Access has numerous advantages over conventional OMA, such as increased spectral efficiency, increased connectivity, reduced transmission latency and signalling costs. Despite such benefits, the resource allocation for satellite-Non-orthogonal Multiple Access systems still needs to be fully explored, and many open questions exist.

b) Description: In power-domain Non-orthogonal Multiple Access, by implementing superposition coding at the transmitter, different users can send their respective signals through the same time-frequency block without interfering with each other. The SIC receiver then uses successive interference cancellations to decode the different signals.

On the transmission side, the channel conditions determine the power coefficients of users in an inversely proportional manner, i.e. transmission power for users with poor channel conditions is higher than for those with good channel conditions.

In the receiver, since the user with the highest transmission power considers the signals of other users as noise, it immediately recovers its signal without requiring any SIC processing. Nevertheless, other users are required to perform the SIC process. As part of the SIC, each user's receiver detects the strongest signals before detecting their desired ones. This process continues until a user's signal is determined by subtracting those signals from the received signal. The final step involves each user decoding its signal by treating users with lower power coefficients as noise.

Consequently, in NOMA, multiple users can simultaneously use a time-frequency block to improve Spectral Efficiency at the expense of additional complexity. This additional com-

plexity, however, is still considered an acceptable trade-off to increase Spectral Efficiency.

c) Conventional Solutions and Issues: Although Non-orthogonal Multiple Access offers numerous features that could support next-generation systems, it also has some limitations that must be addressed to utilise its capabilities thoroughly. Compared to OMA, Non-orthogonal Multiple Access has greater computational complexity because each user must first decode another user's signal before encoding their own signal. To decompose the overall Non-orthogonal Multiple Access signal recovery into distributed low-complexity processes, in [330], a low-complexity receiver has been developed with iterative processing which consists of a linear Minimum Mean-Square Error multi-user detector and a bank of single-user message-passing decoders. Other works have followed similar approaches, such as [331] and [332]. They examine an iterative linear MMSE multi-user detection for both symmetric and asymmetric MIMO-NOMA, and an IRA code is designed for the receiver for asymmetric and symmetric Multiple-Input Multiple-Output-Non-orthogonal Multiple Access. The authors of [333] formulate a joint network stability and resource allocation optimisation problem to maximise the long-term network utility of Non-orthogonal Multiple Access S-IoT downlink system. The authors consider the condition of SIC decoding and propose a practical solution under the Karush–Kuhn–Tucker (KKT) conditions, and then solve the joint resource allocation problem by introducing an optimal solution using the particle swarm optimisation (PSO) algorithm.

In NOMA, the base station should be notified of the channel gains of all users, but doing so will require a significant amount of channel state information (CSI) feedback overhead. To reduce it, [334] propose a scheme where the CSI is acquired only if the BS estimates a high probability of successful UE pairing. [335], [336] study the performance of NOMA with imperfect channel state information (CSI).

d) Proposed ML Solutions: The authors of [115] study the Deep Learning aided NOMA for resource allocation, power optimization, Channel Assignment and User-Pairing, Data rate Maximization, Signal Detection and Channel State Estimation. Using a Deep Reinforcement Learning algorithm, the authors of [119] have developed a dynamic power allocation strategy for NOMA-based S-IoT systems subject to QoS constraints to optimize power allocation factors for NOMA users. According to the authors of [116], a hybrid approach combining data-driven learning and model-based optimization can minimize transmission time in a NOMA-enabled satellite system by optimizing transmit power and terminal time slot scheduling. In [120], the authors minimize the average Age of Information of a NOMA downlink by formulating a Markov Decision Process problem. Furthermore, to overcome the curse of dimensionality and non-convexity, they propose a Deep Reinforcement Learning-assisted age-optimal power allocation algorithm.

In [118], Deep Learning was applied to satellite-IoT systems to determine the optimal decoding order for NOMA's SIC process based on an approximation of the implicit mapping between queue size, channel gain, and decoding order. Also,

to perform completion time minimization, [113] employ Deep Belief Network (DBN) to solve the joint downlink resource allocation non-convex problem using end-to-end learning. Additionally, [117] proposes the DPO combining learning and optimization to provide a feasible and near-optimal solution. Furthermore, to address the problems introduced by imperfect Channel State Information and channel estimation errors, the Support Vector Machine model in [114] showed promising performance by significantly improving user pairing.

3) Rate Splitting:

a) Motivation: Rate-splitting multiple access (RSMA) is an effective multiple access technique that includes SDMA and Non-orthogonal Multiple Access as special cases. One extreme case of RSMA is the linearly precoded SDMA scheme which relies on a transmit-side interference cancellation strategy which treats interference as noise. Another extreme case of RSMA is power domain NOMA which uses a receive-side interference cancellation strategy and fully decodes interference [337].

Furthermore, RSMA can benefit from both SDMA and NOMA, and outperforms both in terms of better spectral efficiency, energy efficiency and QoS enhancements with lower computational complexity [338]. The performance of NOMA is more effective with both underloaded and overloaded network loads, along with various channel strengths, perfect/imperfect CSI and channel directions. Thus, RSMA can be described as a computationally simpler bridge between SDMA and NOMA. Additionally, RSMA offers the advantage of controlling computational complexity and data rate by adjusting the ratio for splitting user signals into common and private portions.

b) Description: As part of RSMA, linearly precoded rate splitting (RS) occurs at the transmitter, and Successive Interference Cancellation occurs at the receiver. The result is that part of the interference can be decoded, and the remaining interference can be treated as noise. Hence, RSMA can partially decode interference and partially treat interference as noise. While SDMA and NOMA rely entirely on either extreme, fully decoding interference and fully treating interference as noise, RSMA acts as a combination.

Specifically, the RSMA scheme divides all users' messages into two independent streams: common and private. Then, the common signals are jointly encoded to form an independent data stream. It should also be noted that private signal interference is considered noise when encoding the common part. As a result, the common stream represents the part of the desired signal expected to be used by intended and non-intended users. On the other hand, the private signals are encoded according to linear precoding schemes in which each user encodes his own signal.

In order to decode the common stream, each receiver first decodes the common stream and treats the private streams as noise. Accordingly, the common stream is decoded and removed by Successive Interference Cancellation scheme. Later, each user decodes their private stream by treating all interference streams as noise. Thus in RSMA, a soft transition is realized between treating interference as noise and fully decoding interference. This is done by carefully adjusting the

message split and the power allocation between the common and private messages. Nevertheless, it is worth noting that although the RSMA has the potential to fundamentally alter the physical and Medium Access Control layers of wireless communication networks, it still faces numerous challenges.

c) *Conventional Solutions and Issues:* A comprehensive tutorial of RSMA can be found in [339] that discuss the myths associated with RSMA, along with answers to frequently asked questions about RSMA. Also, [340] discusses open issues of RSMA in Cognitive Radio Networks (CRN), in particular, how to handle mutual interference (functions of the power allocation, rate splitting parameters or the power allocation factors). Thus, both RSMA and CRN parameters (spectrum sharing) must be solved jointly, which leads to a non-convex problem. Furthermore, open issues such as spectrum sensing, physical layer security issues, along with the impact of RSMA on the physical layer and lower MAC layers have been discussed. The authors of [339] discuss the Physical Layer Security issues and state that if encryption is used in higher layers, there will be no privacy or security issues in RSMA.

More specifically to satellite systems, the authors of [341] discuss the Max-Min fairness (MMF) achieved by conventional linear precoding and RSMA in multigroup multicast with imperfect CSI and demonstrate the benefits of RSMA strategies in both underloaded and overloaded scenarios. Also, by combining a modified WMMSE approach with an alternating optimization algorithm, the authors in [342] solve the MMF optimization problem and propose an RSMA approach that is highly promising for multibeam satellite communication systems to handle inter-beam interference, taking into account practical challenges such as CSI uncertainty and practical per-feed constraints. A similar approach is followed in [343], but for Multigateway Multibeam Satellite Systems. Security Analysis was investigated in [344], and a beamforming scheme based on RSMA is proposed to suppress eavesdropping.

d) *Proposed ML Solutions:* With limited knowledge of channel information and uncertainty of the communication channel, optimizing power allocation in RSMA is very challenging. In the conventional approach, the sum rate maximization of RSMA is generally achieved by the WMMSE algorithm. However, due to high computational complexity, the authors of [121], [122] use a hybrid Machine Learning technique based on deep learning called Deep Unfolding with momentum accelerated Projection Gradient Descent algorithm and outperforms the original WMMSE algorithm in sum rate and speed. In [123], the power allocation problem to common and private streams is discussed. The paper proposes a highly-effective proximal policy optimization based solution, enabling the LEO satellite to learn an optimal power allocation strategy to maximize the sum rate with low computation complexity.

4) Constellation Routing:

a) *Motivation:* Satellites are usually grouped in a constellation to offer global or near-global coverage and service. Several constellations are already operating, such as O3b, Galileo, Global Positioning System (GPS) at the MEO orbit or Starlink and Iridium at the LEO orbit. Within each constellation, satellites can be connected with the neighbor satellites via inter-satellite links (ISLs) either inter-plane, i.e.,

the link between two satellites in a different orbit, or intra-plane, i.e., the link between two satellites in the same orbital plane. GEO system has a wide coverage and typically does not require complex routing over the satellite component. On the contrary, MEO and, especially, LEO satellites have a smaller coverage. This brings to a constellation with several satellites and many ISLs between them. As a consequence, efficient routing algorithms are of vital importance to offer global services. In addition, beyond 5G networks will provide reliable and ubiquitous connection while accommodating heterogeneous services, e.g. emergency services, mobile broadcasting & broadband, sensor-based networks, earth observation missions [345], [346]. To achieve this goal, beyond 5G satellite networks are expected to have a multi-layer (LEO-MEO-GEO) structure, that introduces several challenges in the intra-layer, making the constellation routing even more complex to handle. In the following sections, we will describe the above mentioned challenges, together with the proposed AI/ML techniques to overcome them.

b) *Description:* Routing defines a path on the substrate network, i.e., a sequence of nodes, to allow the traffic generated by a source node to reach the destination node(s). The routing algorithm is an optimization decision process based on several parameters, such as path length, resources of intermediate nodes, link available capacity and delay. The output of a routing algorithm is then translated into commands for each intermediate node. These commands are entries of the forwarding table of each device. Thanks to the forwarding table, the device knows how to forward the arrived traffic to the next hop, as decided by the routing algorithm. The routing algorithm is computed with a fixed transport network once the intermediate nodes receive the commands and traffic flows. However, when the substrate network has a highly dynamic structure, e.g. the NGSO satellite constellations, especially for low altitude orbit ones, and due to the strict power/computational constraints on-board the network nodes, i.e., satellites, the routing has to face numerous challenges. Primarily, with the increase of the nodes and links in the network, the number of available paths and their dynamic behaviour over time increases considerably. This brings more complex routing algorithms to be computed. Secondly, the heterogeneity of on-board capabilities between the GEO-MEO-LEO satellites and the difference in the link latencies additionally increases the complexity of the algorithms since the number of constraints grows. In the following subsection, we present conventional algorithms to address satellite routing and the issues highlighted for the next-generation integrated satellite-terrestrial networks.

c) *Conventional Solutions and Issues:* First, a valuable routing algorithm should guarantee each application's required QoS, e.g. latency, packet loss, and data rate. To handle this together with seamless handovers, due to the mobility of the substrate network, a temporal subdivision into time intervals is applied such that the substrate network is considered static within each time interval. Then, traditionally, the most common approach to solve the routing is the shortest-path (SP) [347]. While SP was efficient with the first generation of MEO and Low Earth Orbit constellations, several issues were raised

with the advent of more complex and more extensive networks. Alternative conventional solutions such as the minimum hop count [348] has also been proposed for Mega-Constellations. However, due to the high mobility of the Space-Air-Ground Integrated Network nodes, i.e., satellites, flying devices and mobile users, and the high volumes of data which may easily congest the channel [349], providing seamless transitions remains challenging. Additionally, as highlighted in [126], the routing problem along the Space-Air-Ground Integrated Network not only deals with the link capacity constraints but also needs to consider the different caching and computing capabilities of satellites, which increases the overall computational complexity.

d) Proposed ML Solutions: Authors in [124] describe a Convolutional Neural Network based routing algorithm to optimize the Space-Air-Ground Integrated Network's performance via traffic patterns and the remaining buffer size of GEO and MEO satellites. They propose online training, distributed over multiple satellites, due to a high volume of data, and does not interrupt the throughput while training. Authors in [126] propose a Deep Q-learning approach to overcome the computational complexity of a joint optimization routing algorithm. The DQL algorithm is described with 1) the state space, function of user-LEO angle, networking, caching and computing states, 2) the action space and 3) the reward function, defined as the efficiency of unit resource. In [125], authors propose a DRL algorithm where the rational agent of the model learns following the replays and selects the route with the smallest Round Trip time. The routes chosen, although in principle crossing more nodes, and therefore longer, in the end, have better effectiveness and a lower load on the bottleneck router.

5) IoT Channel Access to Scheduling:

a) Motivation: IoT via satellite is envisioned as a viable solution to connect devices in remote places due to the ubiquitous nature of satellite coverage, and thus dramatically complements terrestrial IoT solutions. However, due to the massive number of terminals, using optimal resource access and allocation policies would result in a computational load incompatible with the processing constraint of IoT devices. On the other hand, using simple access techniques such as random access lead to under-utilization of the network resources. Hence, machine learning is an interesting solution to offer a trade-off between channel utilization and computation complexity.

b) Description: Due to the specific nature of IoT networks, access and resource allocation protocols for IoT satellite networks must satisfy several criteria regarding efficiency, delay, complexity, and scalability. A detailed list is available in [350, Sec. 6.3]. The authors of [16] present a survey on current technologies and standards in the field of satellite IoT technologies, as well as the challenges inherent to this type of communication. As presented in [7, pp. 92], access methods can be grouped in two categories: *fixed assignment* (FA) based, and *random access* (RA) based. The former ensures contention-free access but requires precise synchronization, while the latter has little synchronization constraints but has to deal with packet collision probabilities. In the following

paragraphs, we will separate FA and RA-based solutions with different constraints, challenges, and network and IoT device designs.

c) Conventional Solutions and Issues: Two leading well-known FA solutions are the Non-Terrestrial Network narrow band Internet of Things (NB-IoT) [351] and Fifth Generation New Radio standards. Both are part of the Fifth Generation cellular access technology developed by the 3GPP. The fact that those solutions are part of a cellular standard is at the same a strength and a weakness. On the one hand, it allows seamless integration of Non-Terrestrial Network into other cellular-based networks. On the other hand, those standards impose restrictions that do not necessarily mix well with the constraints found in NTN. One issue is the higher Round Trip time delay, especially in GEO links. In LEO systems, other common issues are differential doppler and high demand variability due to short orbital periods. These issues can conflict with the precise synchronization needs of Orthogonal Frequency-Division Multiple Access, and lead to more complex and expensive terminal design (e.g. by including GNSS receivers). Most of the existing literature focuses on finding solutions that solve those issues with as few modifications as possible to the existing standards.

RA solutions are well suited for satellite Machine-to-machine communications as they satisfy most of the criteria inherent to this type of communication. The authors of [350] provide a survey on RA solutions for satellite uplinks, focusing on Machine-to-machine communications. Originally based on the ALOHA protocol, RA methods were first used in satellite communications for capacity requests [352]. Successive improvements, such as in [353], have allowed new RA methods to achieve low ($< 10^{-3}$) Packet Loss Probability (PLR) even at high (≥ 1) network loads, enabling applications in massive machine-type communications (mMTC). However, the very high number of terminals, and thus access requests, coupled with the constraints of NTN networks (e.g. high RTT), brings down the performances of those solutions in satellite IoT networks.

d) Proposed ML Solutions: As for the conventional solutions, we will present the FA Machine Learning solutions first. The authors of [127] presented how classical DAMA with RA requests could be improved with dynamic channel allocation and Deep Reinforcement Learning, retaining the high efficiency of FA while adding the flexibility of RA. The optimization is done on the satellite side, which receives transmission requests from the terminals and schedules the data transmission up to a given number of time slots ahead. The bandwidth, the transmission completion time, and the number of scheduled transmissions for each time slot characterize the states. The reward function increases when the sum of transmission completion time decreases. The authors also present several techniques to speed up training sessions. The results showed that CA-DRL reduces by up to 87% the delay in request satisfaction compared to first-come-first-served and pure ALOHA policies. However, no comparison is offered against more advanced techniques like CRDSA. This work was expanded by the same authors in [82], of which a description can also be found in [354, Sec. VI.B.a] and [355, Sec. V.B.3].

The authors added multibeam, QoS satisfaction, and energy efficiency, extending the terminals' lifetime in remote IoT networks. The authors formulate a Markov Decision Process for allocating tasks, defined by their size and location, to channels and beams. They present how to use DRL to solve this problem. The action space is defined as all the possible channel allocations. To reduce its size, the allocation of a channel to a beam is done sequentially. The reward is a trade-off between two components: the first represents power management. In contrast, the second represents QoS satisfaction as a service blocking rate inside the time window. The authors also construct the users' requests into an image as the input of the considered NN, which can reduce the input size and accelerate the learning process. Results show a reduction in energy consumption of up to 67% compared to random and greedy algorithms and minimize the service blocking rate amongst all algorithms. In both the above works, the agent is placed on the satellite, and the algorithm is supposed to be run onboard.

Regarding RA Machine Learning solutions, the authors of [132] have presented a Deep Dyna-Q Learning solution for the random access control in IoT LEO networks (slotted ALOHA). IoT devices connect via random access opportunities, of which the product of preambles and physical random access channels available determines the number. The authors present an optimization problem to maximize the utilization of idle RAOs while keeping the transmission delay under a given threshold. An Enhanced access barring (EAB) mechanism is used in case of congestion. The states in the Markov Decision Process are defined as the number of devices connecting to each satellite. The action set is the number of available RAO (discrete set). The reward takes values in the set $\{1, 0, -1\}$, depending on if the utilization at the next step is respectively higher, equal, or lower than the current one. Probabilities of transition between states are expressed analytically based on the Poisson process for both contention-based and contention-free random access. The authors then present a solution to the optimization problem in the form of a Deep Dyna-Q Learning algorithm and a model-free Deep Q-learning algorithm for comparison. Both Q-learning algorithms present a 3.5-fold improvement in access efficiency over classical dynamic RAO allocation. The proposed Deep Dyna-Q Learning solution performs similarly to the model-free Deep Q-learning. It does not necessitate the feedback interactions between satellites and IoT devices needed in model-free solutions. The authors consider the agent to be placed on the satellite and consider onboard training without, however, offering an overview of the demanded computational power to run the proposed Reinforcement Learning algorithm. To improve the efficiency of satellite communication systems, the NOMA technique has been applied with Machine Learning in several works. In [120], the authors minimize the average Age of Information of a NOMA downlink by formulating a Markov Decision Process problem. Furthermore, to overcome the curse of dimensionality and non-convexity, they propose a Deep Reinforcement Learning-assisted age-optimal power allocation algorithm. The authors of [128] have applied a model-free Q-learning algorithm in the context of IoT satel-

lite terrestrial relays network (STRN) with Non-orthogonal Multiple Access. This work is a combination of [129] and [130]. In this scenario, the IoT devices communicate with a nearby terrestrial relay, which then transmits the aggregated data to satellites. The terrestrial relays are assumed to perform successive interference cancellation (SIC) based on the power imbalance between received signals. Each IoT device is a learning agent. The states are defined as the timeslots, while the actions are defined as the channel's choice. The reward comes from a single feedback bit transmitted by the relay to the device and is either one if the transmission is successful or minus one if it is not. The Q-learning algorithm replaces the random time slot and channel selection by learning based on the relay's feedback. Results show up to an 18% improvement in spectral efficiency over SA-NOMA. In this context, Yan et al. [119] studied the power allocation problem for NOMA-enabled SIoT using a single-agent DQN-based DRL scheme. In their system, the agent is the satellite, whose action space is discrete, corresponding to selecting the power allocation coefficient for each NOMA user. The state space is continuous, consisting of NOMA users' average SNR, link budget, and delay-QoS requirements. At the same time, the reward is discrete, which is a function of the effective capacity of each NOMA user. Experimental results demonstrate that their proposed DRL-based power allocation scheme can produce optimal/near-optimal actions, providing superior performance to the fixed power allocation strategies and OMA scheme. Sun et al. utilize the DL technique to optimize the Successive Interference Cancellation decoding order for NOMA downlink system in satellite IoT networks [118]. This research proposes a long-term joint power allocation and rate control scheme to improve the NOMA downlink rate. Then, the Lyapunov optimization framework is adopted to convert the original problem to a series of online optimization sub-problems. The power allocation depends on the SIC decoding order, which is further affected by the queue and channel states. Due to the continuous changes, the DNN model is adopted to map from the states of queues and channels to the SIC decoding order. Moreover, the DNN is trained supervised with the data obtained by traversing all possible choices. Literature [131] presents a resource allocation for IoRT with joint time-variant channel fading process and stochastically fluctuating solar infeed process. The solution is based on SARSA based actor-critic Reinforcement Learning and is evaluated in average amount of downloaded IoRT data by the satellite. The authors conclude that performances depend on battery capacity.

C. Upper Layers

1) Traffic/Congestion Prediction:

a) *Motivation:* Nowadays, most existing satellite network systems operate independently. Integrating different satellite constellations is considered a promising feature of next-generation satellite networks. The connection inter different satellite constellations could improve the performance of the satellite network and create the potential for new uses [356]. Multilayer satellite networks consisting of multiple satellite constellations with different orbital altitudes have been

studied as significant space network, which carries traffic to any point on the Earth [191].

Multilayer satellite networks formed by combined networks have several advantages, e.g. the wider coverage area served by the upper layers and the comparatively shorter propagation delay provided by the lower layers [357]. From the standpoint of fair and efficient utilization of network resources, the load of satellites can be balanced by broad coverage areas of satellites with higher altitudes. Overall, due to population imbalance based on different causes such as terrain shape and weather, traffic convergence to certain satellites usually occurs in constellations of mesh satellites with lower orbits. Therefore, intelligent routing control technology and smart resource management are necessary for multilayer satellite networks to coordinate the traffic load according to QoS requirements effectively.

Monitoring changes in the incoming traffic rate is the most general way to detect network congestion that occurs when rate measurements exceed a predefined or dynamically adjusted threshold. Balancing the load is a timely procedure that, however, might not be enough to prevent network congestion effectively. To overcome this challenge it is necessary to develop and employ congestion prediction mechanisms for load balancing by distributing the traffic amongst available network resources.

b) Description: Network traffic prediction plays an important role in many areas of satellite networking. Result outputs of congestion predictors could be used for network management, network design, short and long-term resource allocation, traffic (re)-routing, and anomaly detection. Two categories of prediction methods, based on long and short-term's periods, are typically considered. The former, long-term traffic prediction, estimates future capacity requirements, enabling more effective planning decisions. The latter, predictions within minutes, even seconds, is usually linked to dynamic resource allocation. It can be used to improve Quality of Service mechanisms, congestion control, and optimal resource management.

c) Conventional Solutions and Issues: The authors of [358] study the hybrid GEO/LEO satellite network and address the load balancing problem with QoS provision using the advantage of interlayer interconnection. They base the effectiveness of the traffic distribution scheme based on network congestion prediction using computer simulations.

In the proposed scheme in [358], two different states, i.e., normal and warning states, are defined to present the state of each LEO satellite. The normal and sign conditions imply that the corresponding LEO satellite has a high and low probability of congestion. In the normal state where no congestion event is detected, an LEO satellite attempts to catch a congestion precursor by simply monitoring the incoming traffic rate and not exhibiting any load balancing behavior. On the other hand, in the alert state where there is a threat of the cause of congestion, not only traffic monitoring is performed, and traffic distribution and information exchange for congestion prediction as it exists in congested or non-congested areas.

d) Proposed ML Solutions: The work presented in [359] proposed a link congestion prediction method using machine

learning, which can operate at the upper layer of the controller. The authors propose a software-defined network system model consisting of a data layer, a control layer, and an application layer. The state information of switches and links is collected to train and test the proposed machine learning algorithm to predict the related congestion. The prediction performance of four machine learning algorithms was compared; in this case, CNN was shown to obtain the highest pressure, reaching 98.3%.

Neural Network models are also used to predict network traffic, such as Convolutional Neural Network, Long Short-term Memory, and Full-Connected Deep Neural Network. These techniques mentioned above did not explicitly consider the topological information of the network. Therefore, the prediction performance on network traffic was low. Recently, Diffusion Convolution Recurrent Neural Network (DCRNN) was proposed to capture critical topological properties of the network, which are expected to influence the patterns followed by significantly the traffic when propagating through the network [360].

Authors of [361] introduce an hybrid architecture composed by stacking Convolutional Neural Network and Long Short-term Memory which produce predictions of the traffic for a particular beam in a two-hours window with a resolution of 5 minutes. In this work, the traffic data statistics provided by a european satellite operator within several months made possible to train extensively the implemented Neural Network to achieve accurate predictions. Despite it is not possible to detect outliers and peaks in traffic prediction, the results unveil predictions at trending level, proposing a mechanism to indicate the level of future congestion.

2) Integrated Satellite-Terrestrial Networks:

a) Motivation: Satellite systems were developed when reliable and secure communications were needed. They have evolved today to cover a wide range of services, from broadband to critical communication services [362]. Each service has its specificity in performance, architecture, or technology, which puts flexibility as a requirement for today's satellite networks. The same constraint is also applicable in the terrestrial networks, and therefore 5G networks and beyond put flexibility and versatility at their core definition. Since 3GPP Release 15 [363], as the satellite and terrestrial networks share common objectives, satellite networks are now part of 5G and are referenced as Non-Terrestrial Networks within the standard. Terrestrial networks expect to benefit from the wide coverage area of satellites, which does not require deploying any terrestrial infrastructure to extend their coverage. This integration should bring more flexibility and efficiency in network deployment and management. However, this integration is challenging as both networks have specificities and must converge technologically. To achieve this objective, the network slicing paradigm can be used [364]. A network operator can create end-to-end partitions and deploy services while guaranteeing their end-to-end Key Performance Indicators.

b) Description: A multitude of research projects has focused on the integration of satellite networks into terrestrial mobile networks. We can cite the VITAL project, which

TABLE IV
SATELLITE-TERRESTRIAL INTEGRATION MODES [365]

Integration type	Challenges
3GPP Access	Support of the NR waveform within the satellite
Trusted non-3GPP Access	Integration of the satellite in trusted non-3GPP Access mode within the standard
Untrusted non-3GPP Access	Support for 3GPP untrusted access mechanisms within the satellite network
Relay node with 3GPP Access	Support of the NR waveform within the satellite, adaptation of the relay mechanism within the satellite terminal
Relay node with Trusted non-3GPP Access	Integration of the satellite in trusted non-3GPP Access mode within the standard, adaptation of the relay mechanism within the satellite terminal
Relay node with Untrusted non-3GPP Access	Support of 3GPP untrusted access mechanisms within the satellite network, adaptation of relay mechanisms within the satellite terminal
Transport Network based on 3GPP System specification	Definition of a satellite system compatible with the 5G standard
Transport Network non-based on 3GPP System specification	Define an adaptation layer for the satellite network considered as a transport network

aimed to apply the concepts of Software Defined Networks and Network Function Virtualizations to the satellite network [366] in order to increase its flexibility and allow close integration with 4G networks. It was followed by the SaT5G project [367], which studied the integration of satellites in 5G networks. The table IV summarizes the different modes of satellite integration into 5G networks: direct access modes, indirect access modes as a relay node (using gateways to connect other, smaller User Equipments (UEs)), and integration as a Transport Networks for backhaul link services. Following these projects, Zhu et al. [24] listed all challenges regarding the 6G-satellite integration. Depending on the integration scheme, adaptations have to be achieved at various layers of the systems. Interconnection and integration should take place at the orchestration plane, control plane and data plane. This also implies comprehensively allocating and efficiently managing terrestrial and satellite components' resources. A complete integration scheme is direct access where Satellite Terminals embeds all the 5G protocol stack and connects directly to satellites also compliant with 5G Radio Access Network. This method is also the most complex, as satellite networks need to be fully 3GPP compliant. Then, depending on the trusted mode, the relay node can be an intermediate solution where the Satellite Terminal relays (with or without embedding the 5G protocols) all data from User Equipments to the Radio Access Network. The last and easiest mean to realize such integration is to consider the satellite network as a transport network for the 5G system and create backhaul links between 5G Radio Access Networks and Core Networks. Considerable research was conducted using conventional or Machine Learning solutions.

c) Conventional Solutions and Issues: Ahmed et al. [368] have worked on the integration of satellite and 5G using Software Defined Network/Network Function Virtualization and proposed a method for hybrid resource reservation across terrestrial and satellite Point of Presences using Mixed-Integer Linear Programming combined with an online algorithm to ensure a minimum end-to-end delay. Drif et al. [369] have analyzed the network slicing paradigm and applied it to satellite networks to enhance its flexibility and integrate it as a network slicing within 5G networks. They have listed all the challenges related to slicing in satellites, identified interconnection points between terrestrial and satellite networks and produced a novel

architecture to mutualize all satellite resources into a single physical infrastructure. They have implemented their work on a 5G and satellite network in [370] and demonstrate the end-to-end service and Quality of Service continuity using network slicing. Kodheli et al. have listed the challenges of using 5G direct access in Low Earth Orbit mega-constellation and proposed high-level solutions in [371] regarding the Random Access procedure. The authors pursued their research in [372], where they have detailed all the current limitations of using mobile network protocols in Geo-Stationary Orbit Non-Terrestrial Networks. They have proposed solutions to push these boundaries and validated them on a physical testbed using a channel emulator and a 4G modified stack of the OpenAirInterface software.

d) Proposed ML Solutions: Lei et al. [133] have presented AI-based solutions to manage network slices deployed over the integrated 5G-satellite networks. They have listed major issues of applying Artificial Intelligence in such networks and proposed solutions to two distinct dynamic scenarios: predictable network variation using Time-expanded Graph and unpredictable network variation (e.g., sudden change in mega-constellation topologies), which requires fast adaption mechanisms due to model drift. Solutions can be conventional learning (retraining a model), Transfer learning, Meta-learning, or Reinforcement Learning. Rodrigues et al. [135] proposed combining network slicing with machine learning to optimize links in combined 6G-satellite networks. Their ranking-based framework shows an increased user acceptance ratio compared to the traditional method. Similarly, Bisio et al. in [134] have developed a solution mixing queuing theory and Neural Network (NN) to minimize this time latencies in end-to-end 5G-satellite communications.

3) Edge Computing and Caching:

a) Motivation: Edge computing [373] proved to be the key technology to overcome challenges such as massive and heterogeneous data transmission coming from the Fifth Generation and beyond Fifth Generation applications which have undoubtedly made the cloud computing paradigm unable to manage a such large number of computation tasks, both for the network capacity and for the power need for the cloud computing infrastructures. As emerging for mobile terrestrial networks, edge computing gained interest as relevant technology for beyond Fifth Generation satellite networks [7], [374],

[375]. Edge computing concept is to perform data processing close to the source, which could be the key to enabling widespread adoption of Machine Learning for satellite applications. Satellites, due to their ubiquitous coverage and low latency, when considering LEO satellites, can strongly support terrestrial networks to apply the edge computing paradigm. At the same time, edge caching is one of the key technological enablers in beyond 5G (B5G) networks to enhance Quality of Experience and reduce transmission latency. The premise of edge caching is to prefetch popular contents at the edge caches close to end users, thus, when the users request the cached contents, they can be served immediately without transmission from the data server through the core networks.

b) Description: Edge computing was introduced to avoid computing and storage capabilities being concentrated in centralized locations, also known as the cloud computing paradigm, and most likely far, in terms of network topology, from the network user devices. In fact, on the opposite, the edge computing paradigm runs applications and services in hosting servers/hardware close to the user devices. Edge computing is an optimization problem to decide how/where to allocate the tasks to be executed. The larger the network and the traffic demand, the higher the complexity of the problem. Accordingly, ML relies on learning techniques to support and reduce this complexity increase and achieve good solutions in a reasonable time. Edge computing is typically used in Satellite Communication in two main solutions. Computation offloading [376], whether total or partial, is the technique to offload computational tasks from the cloud to the local Mobile Edge Computing servers. This is a flexible approach, depending on the Quality of Service application/user requirements, e.g. tolerated latency and power constraints, the computational tasks are offloaded to servers accordingly. For instance, delay-sensitive applications are expected to be partially offloaded to servers due to time constraints.

Content caching is the second common technique for edge computing over SATCOM [377]. The lack of spectrum availability has forced to find new solutions to minimize the traffic over information networks. Caching the contents to servers/stations close to the user, reduces the transmissions over the network and the overall latency experienced by each user. Edge caching in general comprises of two phases: cache placement phase and delivery phase. In the first phase, selected contents are determined and are prefetched on the caches, which usually happens during off-peak times. The second phase occurs during peak hours when users request contents. If the requested contents are available in the edge caches, they can be served immediately without being sent through the backhaul network. Two main issues in edge caching is i) how to determine the efficiently determine the popular contents, and ii) how to optimally allocate the edge caches in the network.

c) Conventional Solutions and Issues: Conventional caching is to exploit the long-term content popularity to determine that cached content. In general, based on the history of requested data, the most popular contents will be prefetched to the cache [378]. In [374], several techniques are proposed to decide the location to which tasks should be offloaded. Authors propose different scenarios where tasks are either offloaded

to close terrestrial servers, on-board of satellites or close to ground stations (gateways).

One of the issue/challenge, presented in [137], when implementing edge computing over SATCOM, is the high dynamics of the network. Due to fast and frequent variations, the channel features and devices availability is critically unstable and time-dependent. This might bring to delay or failures when communicating to the required servers. In addition, as the beyond 5G SATCOM networks are expected to be very heterogeneous, especially in terms of on-board available resources, the optimization design of the offload policies should be properly analyzed, considering the variety of available on-board capabilities/resources.

d) Proposed ML Solutions: The majority of proposed ML solution for edge computing in the literature are developed for terrestrial networks. However, recently few contributions have emerged in the area of integrated satellite-terrestrial networks. Authors in [126] propose a deep Q-learning approach to solve the joint optimization of routing and caching for Space-Air-Ground Integrated Network. ML-based edge computing is mainly proposed for IoT scenarios, because the highly dynamic and random computing task arrival makes the edge computing deployment challenging. Authors in [136] propose a Deep Reinforcement Learning approach to offload IoT tasks with the objective of reducing the average processing delay while considering the Unmanned Aerial Vehicle energy capacity (consumed by communication and computing). The Reinforcement Learning approach is used via a risk function which accounts for the violation of the Unmanned Aerial Vehicle energy availability. In fact, the main idea is to use Unmanned Aerial Vehicle to collect Internet of Things tasks, and run the Deep Reinforcement Learning to schedule them either locally on-board (UAV), in the closest base-station or to the LEO constellation. These offloading servers are complementary since UAV is very close to the IoT, the LEO constellation has a vast coverage and the terrestrial base-station has the highest computation capability. Clearly, the decision will be driven by the requirements from each task. A similar scenario is considered in [137], where authors work on the optimization of computational task offloading for IoT-based Space-Air-Ground Integrated Network networks. In this context, they propose a Reinforcement Learning approach to learn the timely network behaviors in order to determine the optimal offloading policy. The learning process is improved by the use of Deep Neural Networks, policy gradient and actor-critic methods. The final objective is to minimize a composite function of delay, energy consumption and server usage cost.

Centralized learning for caching is studied in [379]–[381]. The authors of [379] consider a SBS with limited cache capacity connected to a central controller via a limited capacity backhaul link. The time is divided into separate periods in which the combinatorial multi-armed bandit (CMAB) learning algorithm is used to learn the file popularity over the time and improve caching efficiency. The work of [380] considers centralized content popularity learning algorithms of a network comprised of multiple Wi-Fi or WiMAX access points which have a shared cache storage and serve multiple users. The authors of [381] consider multiple SBSs where each of them

is connected to the core network using a limited capacity backhaul link. Device-to-device interactions of each user are used to estimate the file popularity matrix. This matrix is used to apply the transfer learning on each SBS to optimally cache the content according to storage limit, estimated popularity, and backhaul capacity. Distributed caching is studied in [382], [383]. The work of [382] considers a network comprised of a central BS and multiple SBSs where each SBS serves multiple users. The authors use the coded caching for distributed file replacement in SBSs. The multi-armed bandit algorithm is used to learn the file popularity in time and reduce the traffic from the central BS to the SBSs by considering the users' connectivity to each SBS and pre-fetching the content. The works of [382], [384] consider only one learner in the network which can impose a huge computation load on the learning agent. This issue is addressed in [383] using a distributed multi-agent learning algorithm where the computation load is distributed and there is no need for the channel state information at the macrocell BS. However, authors in [382], [383] have not considered improving the learning convergence rate of the distributed caching by coordinating the caches through satellite connectivity. Furthermore, backhaul and user link capacities were not included in the optimization model. Theoretical analysis of time and samples to achieve a specific performance is studied in [385]. The work of [385] considers a network of multiple BSs and cache-enabled SBSs where users are spatially distributed according to Poisson point process. Transfer learning algorithm is used to estimate the file popularity at the BS in order to reduce the user waiting time using distributed caching. Proactive caching along with file popularity estimation has been studied in [386], [387]. The authors of [386] predict user demand based on Wi-Fi access, battery status, and download bandwidth analysis. This way, the predicted data are cached before the user demand and this reduces the network congestion. The work of [387] uses the content request delay to predict users' requests. A schedule is used to push the content into the receiver buffer and cache it to maximize the number of available user requests at the cache. The application of satellite communications due to high bandwidth and wide area coverage is investigated in feeding several network caches at the same time using broad/multi-cast [388], [389]. The work of [388] uses the satellite broad/multi-cast to feed the caches at the user side. The authors of [389] consider a network comprised of proxy servers with individual cache storages. If the requested content is not available in the cache, it is routed to the gateway for satellite transmission. Based on the number of requests, the file size, and the available satellite bandwidth the request will be admitted or dropped. The satellite multicasts the requested content and the global popularity. Each server uses the local and global file popularity to update the cache. While [389] proposes satellite multicast in cache feeding, it does not consider the backhaul cost, user link, cache storage limit, handling the dropped packets at the gateway, and the interaction between the macrocell and SBSs. Content popularity estimation, selection, and delivery are studied in [138], [390], [391]. The works of [390] studies the caching in a network where small cell networks with high capacity cache are connected to central scheduler via a limited

backhaul link and serve multiple users. The central scheduler learns the file popularities by tracking users request and places the right file in the cache of the right small cell network. The authors maximize the ratio of the satisfied requests to the total user request while considering the backhaul, cache, and user wireless link capacities. The authors in [138] consider a network comprised of a macrocell BS and multiple cache-enabled SBSs where each of them serves a group of users. The authors minimize the content delivery delay to the users by learning the content popularity at each SBS and updating the cache according to it. As a practical approach, [391] studies the edge caching in a cellular 5G wireless network comprised of cache-enabled SBSs. Statistical machine learning is used for file popularity estimation and file caching in off-peak hours. The cached content is assumed fixed during the usage hours. The work minimizes the average backhaul load by considering the file size, file bit rate, backhaul link capacity, wireless link capacity, cached content, and user quality of experience (QoE). The authors perform big data analysis on users' data and show that it can be modelled as a Zipf-like distribution [378].

D. Others

1) Integrated Sensing and Communication:

a) *Motivation:* Current networks are evolving in all spectrum bands: RF, mm-wave, and terahertz to solve the shortage of spectrum. Another promising solution is represented by multi-function systems simultaneously performing. This is the case for joint sensing and communications systems which converge spectrally through waveform design. Both radar and communications systems need wide bandwidth to provide a designated quality of service, thus resulting in competing interests in exploiting the spectrum. Hence, sharing spectral and/or hardware resources of communications and radar is imperative for efficient spectrum utilization. This has led to the emergence of Integrated Sensing and Communication systems which employ one of the following: common channel access, transmit/receive hardware units, processing, or waveforms [392].

b) *Description:* The electromagnetic spectrum in different bands (RF, mm-wave, and terahertz) is not nearly sufficient, as the number of emerging technologies with different functions, communications, radar, positioning, navigation, spectral situational awareness, and opportunistic spectrum access, is foreseen to rise in the next years.

Various technologies able to simultaneously performing two or more radio tasks, are currently under discussion to tackle this challenge. One promising solution is represented by Integrated Sensing and Communication systems [393], or Joint Communication and Sensing and JCAS, as can be found in literature, where communication and sensing are performed simultaneously in a dedicated or existing infrastructure. Remote sensing conventionally refers to satellite- or aircraft-based sensor technologies to detect and classify objects usually on Earth. Remote sensing can be split into *active* remote sensing (when a signal is emitted by a satellite or aircraft to the object, and its reflection is detected by the sensor) and *passive* remote sensing (when the sensor detects the reflection

of sunlight). Integrated Sensing and Communication aims to unify radar and communication systems through a combination of joint hardware, waveforms, joint signal design, and joint signal processing.

c) Conventional Solutions and Issues: Generally, the most tracked objects in satellite video used for RS are aeroplanes, ships, vehicles, and trains. However, in SVs, these objects, especially vehicles, are extremely small relative to the captured region, which raises a great challenge for general tracking tasks. Therefore, many excellent tracking methods [394], [395] have been developed to cope with these challenges. Early tracking methods are mainly based on movement features, and colour features, such as Kalman filter (KF) or particle filter [396]. These classic filters or statistical methods, predominantly based on mathematics and physics derivations, aim at achieving relatively fast and robust tracking. However, these trackers are limited by strict parameters or single-usage scenarios. In [397], a model has been developed and validated to make it functional so that elevation-independent data acquisition can occur with specific considerations for LEO satellites. Some issues in conventional RS have been identified in [398], such as increased spatial and temporal coverage and resolution of satellite observations or an increased information content and exploring the synergy of observations. Also, another issue present in conventional satellite RS is achieving continuity in consistent observations by satellites, and the long-term data record [398]. Increasing RS data's resolution in the spatial, spectral, and time dimensions is a possible solution. However, it also presents new problems with the high amount of data, such as insufficient upstream and downstream transmission bandwidth, the need for more adaptability of the remote sensing payload imaging method to the complex ground and feature changes, and time-efficiency problems with information services. Regarding Integrated Sensing and Communication, literature offers already several works focused on terrestrial and aerial networks and explore promising techniques as Multiple-Input Multiple-Output [399] or Unmanned Aerial Vehicle applications [400], [401]. However, the direct application of these techniques is not directly applicable to satellite systems due to the difference in wave propagation properties. In particular, two significant differences can be highlighted, i.e., the inevitably high propagation delay and large Doppler shifts due to the long distances between the LEO satellites and the targets as well as their mobility [402]. Given these substantial challenges, research of Integrated Sensing and Communication for Satellite Communication is still in its infancy stage. [403] proposes a physical layer design to sense the atmosphere in the terahertz band using communication signals from LEO satellites. The work in [402] performs Integrated Sensing and Communication in a massive Multiple-Input Multiple-Output Low Earth Orbit system, which can provide wide coverage for wireless Integrated Sensing and Communication. Due to the inaccuracy of the instantaneous Channel State Information, authors consider statistical Channel State Information.

d) Proposed ML Solutions: The above works show the promise of the potential usage of Integrated Sensing and Communication in conjunction with Low Earth Orbit satellites.

Integrated Sensing and Communication, however, is not only one of the pillars for a new air interface for Sixth Generation networks, but, if coupled with Machine Learning can serve as foundation of a world where everything is connected, sensed and intelligent. Toward this direction, [154] proposes a novel approach using an auto-encoder for data-driven Integrated Sensing and Communication. K-means-based clustering and DRL-based resource allocation (K-DRL) are proposed in [150]. These techniques extract pixel characteristic-based features from satellite images to improve localization accuracy. In addition, the proposed UK-medoids and DRL-based scheme (UKM-DRL) is compared with two other schemes: K-means-based clustering and DRL-based resource allocation (K-DRL) and UK-means-based clustering and DRL-based resource allocation (UK-DRL). In table III we have reported also some remote sensing articles, for the sake of completeness.

2) Cooperative Satellite Communications:

a) Motivation: Cooperative satellite communication is integral to satellite-based applications like remote sensing, surveillance, joint communication, joint scheduling, distributed beamforming, and sensing [404]. Specifically for LEO satellite networks, implementing cooperative communications becomes challenging due to their high mobility. The existing cooperative satellite-based communication requires a huge amount of communication among satellites for cooperative tasks. This creates problems in establishing joint beamforming, joint scheduling, and other cooperative satellite applications in real-time [405]. In the current scenario, the LEO satellite communication networks such as SpaceX, and OneWeb are gaining popularity because of their low latency, seamless connectivity and broad coverage [406]. However, the networking of these small satellites considers different factors for successful cooperative communication. One crucial factor is the distributed processing and synchronization of satellite networks for coordinated applications [407]. Distributed processing refers to the decentralization of the computing processors, which are physically separated. This distributed processing is application specific. Additionally, these distributed processors can have different structural architectures. It is essential to have efficient distributed processing for effective cooperative satellite-based communication.

b) Description: Cooperative satellite communication comprises a wide range of applications from distributed beamforming to joint scheduling [404] [405]. The existing LEO satellite networks comprise monolithic antennas with different antenna elements. Deployment of such satellites is not cost-effective. Therefore, the current generation of satellite networks with SpaceX, OneWeb, and Amazon Kuiper have aimed for smaller and lighter satellite constellation networks [406]. One of the prime applications focused on by such satellite constellations is distributed beamforming. A pictorial representation of the distributed beamforming performed by a satellite constellation is shown in Figure 6. Figure 6 shows two sets of satellite swarms with different structures. One satellite set is performing the distributed beamforming.

c) Conventional Solutions and Issues: The distributed processing in cooperative satellite communication exists in different architectures, namely a Star, a Link, Linear, Hybrid and

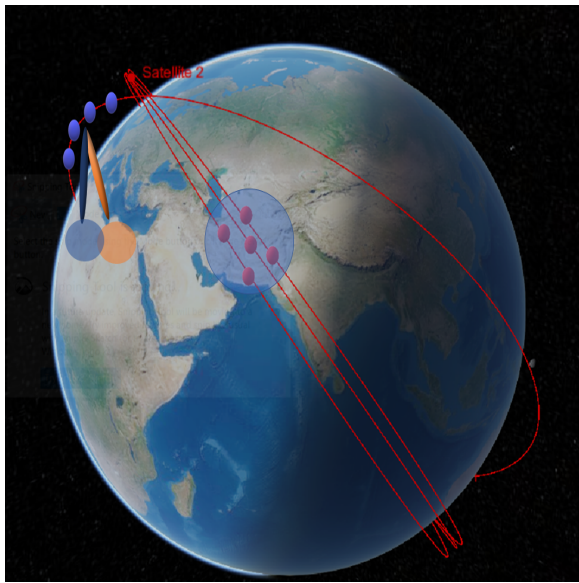


Fig. 6. Distributed beamforming using satellite constellation.

Layered architectures [408]. These existing architectures put in challenges for distributed processing in terms of power, volume, weight and size constraints of satellite platforms [409]. However, the main issue lies in the non-optimal utilization of the parameters. This allows employing model compression schemes via removing redundancy or adapting to a compact design with high accuracy using neural architecture search [409].

d) Proposed ML Solutions: The architectural optimization of the neural network is termed Neural Architecture Search [410]. The Neural Architecture Search is an efficient tool for optimizing hardware architecture for distributed processing in satellites on-board processing while keeping check of power, volume, weight, and size constraints [409]. It works on a reward and learning model to generate high-performance architecture. The Neural Architecture Search generated architectures are observed as faster, more accurate, less complex, and less expensive than conventional schemes like gradient descent, evolutionary algorithm, and reinforcement learning [411].

3) Image Processing:

a) Motivation: Satellite images are becoming an essential spatial data resource for scenarios such as dynamic target tracking [412], disaster monitoring [413], and maritime surveillance [414], where the satellite is a crucial element. Compared with terrestrial images, the quality of satellite images is affected by additional factors, such as atmospheric disturbance, over-the-horizon imaging, the relative motion between satellite and ground objects during observation, and scattering. In addition, to this is added the incorporation of video to SATCOM generates an enormous amount of data collected to meet the data real-time transmission with limited bandwidth. Thus, to improve the image compression ratio or reduce the spatial sampling resolution must be performed by the satellite system. Super Resolution technique is an important solution which maps low-resolution images to high-

resolution images with different scale factors. However, the above unpredictable degradation causes the reconstruction of satellite images to be quite different from terrestrial images. In addition, ground terminals' resolution varies, so it is important to reconstruct satellite images at arbitrary scales. However, due to the limitations of hardware equipment and transmission bandwidth, the images received on the ground have low resolution and weak texture. In addition, ground terminals have various resolutions and real-time playing requirements. Another problem in satellite images is the loss of high-frequency details caused by over-compression. Therefore, achieving arbitrary scale super-resolution of satellite images is essential for signal processing in SATCOM.

b) Description: Satellite Image Processing (SIP) consists of the images of the earth and satellites taken by artificial satellites. In broader terms, SIP is a kind of remote sensing which works on pixel resolutions to collect coherent information about the earth's surface. Depending on the application, the images can have object land cover, clouds, or asteroids. Firstly, the images are taken in digital form and later are processed by the computers to extract the information. SIP is applied by statistical methods to the digital images to identify the various discrete surfaces and analyse the pixel values. The key application of SIP is to detect responses to upcoming disasters or design sustainable infrastructures to monitor environmental conditions.

Four kinds of resolutions are associated with satellite images:

- 1) Spectral resolution measures the wavelength interval size and determines the number of wavelength intervals that the sensor measures.
- 2) Radiometric resolution is defined in bits size. It gives effective bit depth and records the imaging system's brightness levels.
- 3) Temporal resolution is defined as the time or days that pass between various imagery cloud periods.
- 4) Spatial resolution is defined as the pixel size of an image.

c) Conventional Solutions and Issues: The main application where SIP is applied is for Earth observation, which is categorized into four conventional techniques (1) Preprocessing, (2) transformation, (3) correction, and (4) classification. In pre-processing, some distortions produced on satellite images by imaging systems, sensors, and observing conditions are corrected before using the three following techniques to correct or minimize image distortions. In [157], a mixed high-order attention network is proposed, composed of a feature extraction network and a feature refinement network to exploit low-level feature and high-frequency detail, respectively.

d) Proposed ML Solutions: For satellite images, it is essential to reconstruct the fine spatial details for subsequent applications. Many research efforts [157]–[160] have been paid to several aspects. Some of these methods focus on improving CNN architecture. [158] proposed a cycle convolutional neural network (Cycle-CNN), which consists of two generative CNNs for down-sampling and SR, respectively, Cycle-CNN has great performance in re-constructing multi-spectral band images of real remote sensing satellites. [156] proposes a distinct residual dense network (RDN) as the feature learning module to extract

hierarchical satellite image features in LR space. Then, it is followed by an arbitrary upscale module to project LR features to arbitrarily scale enlarged SR images. Finally, the scheme adds an edge reinforcement module to re-cover the edge information and improve SR images with clearer textures. Particularly, in [155], proposed a technique for spatio-temporal fusion using deep DNNs in two phases with a large amount of RS data as the application background. As highlighted by the above works, the availability of large datasets is critical for Image processing in satellites and might represent one of the main challenges. We will present more details about the training dataset challenge in Satellite Communication in Section VI-A1.

4) *Ground Segment Dimensioning:*

a) *Motivation:* Parallel to the renewal in the space segment, we find that the ground segment (GS) must also cope with the immense number of changes made in SATCOM for the space segment. This implies new challenges not yet addressed for this element of the communications system. The ground segment is that part of a SATCOM system that employs a variety of node designs and network configurations to provide and manage services delivered, on the one hand, to end-users and, on the other hand, to satellites using feeder links. Service to the GS of an NGSO system depends largely on the nature of the link, which is shorter but more dynamic than its GEO counterpart.

b) *Description:* The GS dimensioning (GSD) includes determining the number of gateways (GW) needed, the antenna parameterization, the traffic model, and the switching strategies between backup GW. The location of the gateways is one key requirement and vital in NGSO constellations due to the high number of elements that make up the constellations. The location of these elements is strategic for properly functioning these constellations' new services. This strategy can be dynamic as a function of the weather conditions, considering services with greater security restrictions, such as government services and political conditions.

c) *Conventional Solutions and Issues:* [415] compares GS architectures for constellations using feeder links in Q/V-band against those using E-band. In addition, a method is developed to determine the locations of the minimum number of ground stations that maximize the system capacity while achieving desired QoS levels. The work in [416] analyzed the handover performance in a LEO-based Non-Terrestrial Network (LEO-NTN) via system-level simulations, focusing on ground-segment optimization by traditional methods. Two algorithms for dimensioning the optical ground station (OGS) network with MEO satellites are proposed in [417]. These algorithms aim to minimise the network's size while ensuring the target link availabilities due to cloud blockage. A joint strategy for GSD and routing network is proposed in [418] for satellite-terrestrial integration. CARAMUEL project [419] presents the GSD considering ground network with Quantum Key Distribution in GEO satellites. Regarding switching strategy, [420] proposes an optimal gateway selection in each moment using site diversity. In addition, to make the decision to locate these stations, we can make use of emerging technologies and techniques such as AI and ML. The application

of this field for the optimized design in the GS that allows selecting the position of the gateways based on criteria beyond the service requirements is still in its infancy.

d) *Proposed ML Solutions:* In [421], the authors present an overview of optimization techniques applicable to solving the GW location problem. On the other hand, a survey on optimization methods based on ML focused on LEO satellite networks is performed [422]. This survey deals with GS within future autonomous transportation applications using NGSO satellites. However, it is a preliminary study using ML GSD techniques for SATCOM applications. For GEO satellite, ML-based GSD is detailed in [423].

V. HARDWARE SOLUTIONS

AI and ML techniques are used to optimize time and resource-critical applications nowadays. Satellite communication is one of those critical fields. As highlighted in Section III-C, on-ground segment applications can be less restrictive, while onboard applications can be a real challenge from the point of view of power constraints, radiation, and fault-tolerant architectures. This section presents an analysis of the HW for possible AI and ML implementations for satellite communication.

A. Neuromorphic Processors

Mainstream microprocessors are unsuitable for deploying state-of-the-art deep neural networks that require many matrix and vector multiplication operations. This is mainly because on-chip memory, which holds the data for processing, is limited. In addition, operations involving large matrices require frequent access to off-chip memory and storage units, which consume a relatively large amount of power and take too long. Custom neuromorphic processors (NPs) have been built for this purpose that can solve the resulting von Neumann bottleneck. Toward this end, the hardware is specifically designed to allow tight integration of logic and memory units, and accelerator units are provided to execute frequent computational operations efficiently. Moreover, adopting spiking dynamics for the neurons also eliminates their need to perform expensive multiplication operations with high-precision real numbers represented in FP32 or another reduced-precision format [424]. The most advanced hardware implementations to date are

- TrueNorth from IBM [425]
- Loihi and Lohi2 from Intel [426]
- Zeroth from Qualcomm [427]
- SpiNNaker and SpiNNaker 2 from University of Manchester [428]
- BrainScaleS and BrainScaleS-2 from Heidelberg University [429]
- NeuroGrid and Braindrop from Stanford University [430], [431]
- DYNAP devices from SynSense [432]
- ODIN from Catholic University Louvain [433]

B. AI-capable Chipset/Accelerators

So far, complex onboard AI/ML applications can only be performed with very expensive custom-designed ASICs [434]. The increased performance requirements for onboard processing to support higher data rates and autonomy made the existing space CPU obsolete. New technologies, including non-qualified COTS devices, from other critical domains, are currently being explored [435]. Finding a device with adequate computer capacity, proper power consumption, and compliance with standards specifications for onboard and standalone applications is one of the most critical points for the implementation [3].

The commercial AI-capable embedded chipsets market is divided into embedded GPUs and dedicated devices for hardware co-processors or standalone embedded systems. Several devices have been launched to the market, led mainly by NVIDIA, AMD, Intel, and Qualcomm [436].

Table V resumes nowadays embedded COTS AI-capable chipsets from the main companies³:

- 1) Intel Corp.
 - Intel Movidius Myriad family [437]
- 2) NVIDIA Corp. [438]
 - NVIDIA Jetson Nano [439]
 - NVIDIA Jetson TX2 Family [440], [441]
 - NVIDIA Jetson Xavier NX Family [442]
 - NVIDIA Jetson AGX Xavier Family [443]
 - NVIDIA Jetson Orin NX Family [444]
 - NVIDIA Jetson AGX Orin Family [445]
 - NVIDIA Jetson Orin Nano Family [446], [447]
- 3) Qualcomm Technologies, Inc.
 - Qualcomm Cloud AI 100 Family [448]–[450]
- 4) Advanced Micro Devices, Inc.
 - AMD Instinct MI200 Family [451], [452]
 - XILINX Versal ACAP AI Edge Family [453]–[455]
 - XILINX Versal ACAP AI Core Family [456], [457]

AI-capable chipsets architecture includes a Central Processing Unit (CPU) and an on-chip accelerator, which define their capabilities and performance on AI and ML tasks. The amount of CPUs and cores depends on the manufacturer and the family, varying from one to two CPUs and two to 12 cores. Remarkably, the multi-core ARM Cortex-R5F2 on the Versal ACAP chips and the 12 cores on Jetson AGX Orin have high-end chip specifications. The first is suitable for embedded real-time and safety-critical systems. In contrast, the vast number of cores of the second allows the use of multi-threads in sequential parallelizable software [445], [457].

In the case of the on-chip accelerators, as their architecture differs from one developer to the other, a proper comparison takes into account the performance per operation, as for Table VI. Note that we have marked as *NR* wherever some data are not reported in literature and marked as *x* in case data are not detailed. As XILINX doesn't report the power consumption specifications for the Versal ACAP AI Core family, the analysis is done only for the Versal AI Core VC1902 device with the power estimation reported in [458].

The most popular AI on-chip accelerators are embedded GPUs (SHAVE, CUDA Cores, Streams Processors) and AI-capable cores (Tensor, AI Engines, and DSP Engines).

SHAVE cores were designed by Movidius primarily for physics acceleration in mobile applications. Low demand for expensive physics acceleration in smartphones has forced a re-focus on image and vision processing and machine vision processing. The union of several SHAVE cores creates a versatile parallel architecture that can adapt to other application fields such as AL and ML.

CUDA cores are NVIDIA's dedicated stream parallel processors initially released for graphics processing and 3D rendering on conventional GPUs and later ported to NVIDIA Jetson embedded Family. The architecture of a CUDA processor is composed of several CUDA cores as processing elements, making them a good option for AI/ML applications.

Tensor cores are AI-dedicated processing elements released by NVIDIA since the Volta GPU microarchitecture. They specialize in fused multiply-add (FMA) operations used extensively in neural network calculations to apply an extensive series of multiplications on weights and then add a bias. Tensor cores can operate on FP16, INT8, INT4, and INT1 data types executing 1024-bit FMA operations per clock per core (1024 FMA operations of INT1, 256 of INT4, 128 of INT8, and 64 of FP16) [459].

The XILINX AI Engine, a distinctive component of the Versal ACAP AI devices, is designed for intensive computing in various applications including but not limited to, machine learning and artificial Intelligence. Its architecture includes a vector processor capable of FP32 by FP16 multiply-and-accumulate (MAC) operations per clock cycle. It is integrated into an AI Engine tile with data memory and DMA engines. Each tile is connected, using AXI stream interfaces, by the four nearest tiles (except the tiles on the edges), forming a neural network and sharing the memory with the neighbours [458], [460], [461]. Versal ACAP chips, apart from the on-chip accelerators (AI and DSP Engines), include a programmable logic (PL) on-chip that can be used to design time-critical parts of the algorithm, exploiting the inherent parallelism of the hardware design.

The most supported data-types operations by the chipsets are integer bytes (*INT8*), half-precision floating-point (*FP16*), and single-precision floating-point (*FP32*). Instinct MI200 [451], [452] and Cloud AI 100 [448]–[450] families have the greatest operation rates in half-precision floating-point, but their form factor and power consumption make them not suitable for embedded systems applications. NVIDIA's new-generation AI processors family (Orin NX and AGX Orin [444], [445]) promises to achieve good operations per second (*OPs*) rate, as good as the Versal AI chip, increasing the power consumption compared to previous NVIDIA families.

It is essential to analyze the computer capacity per Watt to select the proper chipset for the application. Figure 7 resumes this information for the most common integer bytes operators and floating point. Cloud AI 100 family has better performance on byte (*INT8*) operations (between 2.8 and 5.33 TOPs/W), followed by the Orin NX Family (3.5 and 4 TOPs/W) and the incoming Orin Nano Family (2 and 4

³Data collected from available vendor's datasheets and published papers

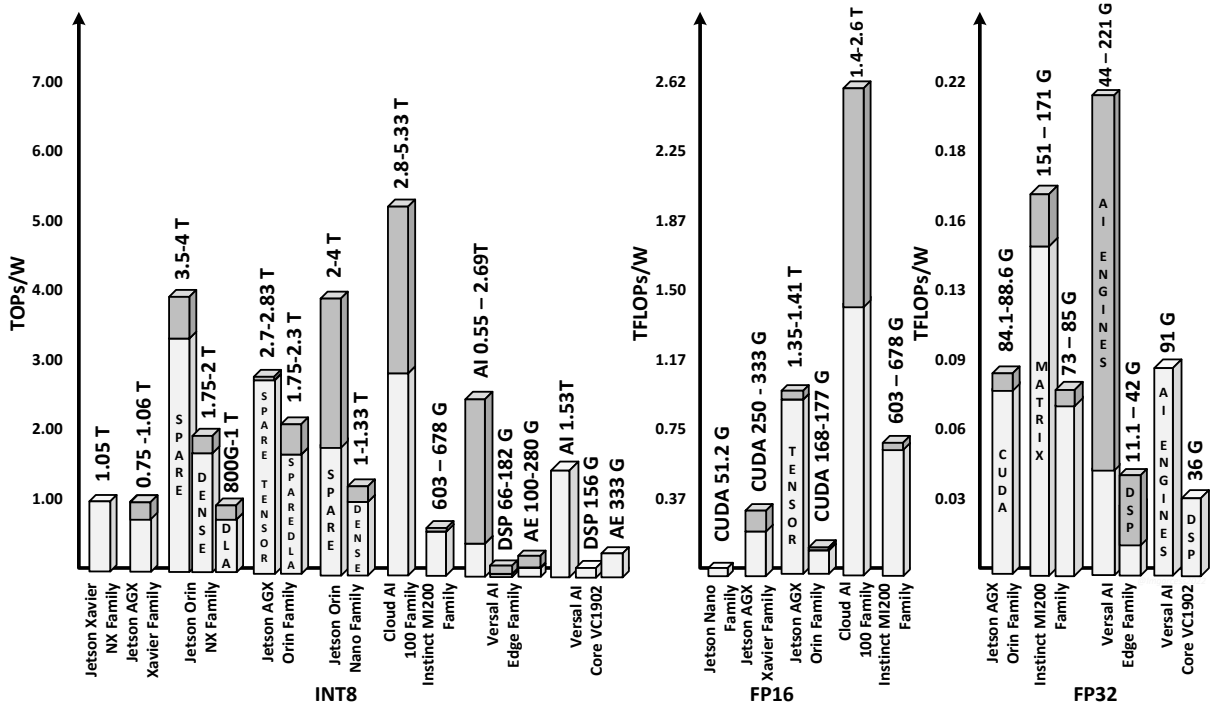


Fig. 7. AI Chipset/Accelerators off-the-shelf: Computer Capacity per operation

TOPs/W). However, Versal's AI Engines on Edge Family gets a wide range of performance (between 0.55 and 2.69 TOPs/W), similar to the AGX Orin Family (2.7 and 2.83 TOPs/W). On Versal's AI Engine on AI Core Family, although its reported computer capacity per operation is outstanding, the power consumption reported in [458] for the VC1902 chips is too high, especially for onboard processing on space missions. With half-precision floating point (FP16), Cloud AI 100 gets between 1.4 and 2.6 TFLOPs/W, followed by Orin AGX family (1.35 - 1.41 TFLOPs/W), while Versal AI's data is not reported for this kind of operations. For single-precision floating-point (FP32), Versal's AI Engines of Edge Family achieve the best performance (up to 221 GFLOPs/W), followed by the Instinct MI200 Family (between 151 and 171 GFLOPs/W on Matrix operations), while the Jetson AGX Family using CUDA cores (84.1 and 88.6 GFLOPs/W) reports slightly less performance than the Versal AI Core VC1902 AI Engines (up to 91 GFLOPs/W). This comparison proves that Versal AI Edge Family has a remarkable and wide range of performance per Watt, being a good solution for AI and ML applications for onboard Satellite Communications. Analyzing the Versal AI Core family for similar applications requires more accurate studies of chip performance.

Finally, Table VII details the chip dimension and availability. Proper chip size comparison takes much work to perform with the available datasheets. Some sellers, like Intel and XILINX, report the die size. Others, like NVIDIA, report the available PCB, including auxiliary chips and interfaces, and Qualcomm and AMD report the final product. Myriad Family reports the smallest size varies from $8 \times 9.5 \text{ mm}^2$ for Intel's Movidius Myriad 2, actually discontinued, to $14 \times 14 \text{ mm}^2$ for

the Myriad X. NVIDIA's family reports size from $45 \times 69.6 \text{ mm}^2$ to $87 \times 100 \text{ mm}^2$. NVIDIA's roadmap announced the new Orin Family, planned to be released between December 2022 and January 2023, depending on the chip availability. This family was recently updated with the Jetson Orin Nano, a device for edge applications with lower power consumption [438], [446], [447]. Qualcomm AI 100 and AMD Instinct MI200 families report more significant devices with PCIe interfaces, more suitable for systems co-processors and accelerators, not for standalone and on-board applications. On Versal's AI families, the chip size varies from $23 \times 23 \text{ mm}^2$ to $45 \times 45 \text{ mm}^2$. However, only one development board has been released, the VCK190 powered by the Versal AI Core VC1902, which is too big and power demanding for on-board applications ($190 \times 241.3 \text{ mm}^2$).

C. Hardware Analysis-Solutions Comparison

Although machine learning and artificial intelligence are emerging fields for computing in-the-edge, implementing those algorithms for space applications is challenging for the scientific community due to the duration of the missions, the radiation effects on the chips, the temperature span and the performance-to-power ratio. The use of COTS Processors for onboard space missions requires radiation-hardness-by-design (RHBD) or radiation-tolerance (RT) of the chip, fault-detection support and correction codes, single-point thermal load, long-term support and availability [462], [463].

Although none of the COTS AI-capable devices is considered radiation-hardened-by-design new studies are emerging for using COTS chipset on onboard AI applications. This is the case with the GPU for Space project (GPU4S) that uses

embedded GPUs for space workloads [464]. The authors focus on four premises for achieving the results: the parallelization capabilities of the software, the comparison with the most available COTS device in the market at this moment, and the use of a benchmarking tool created as one of the milestones of the project (GPU4S Bench) reported in [465], [466]. The test was performed for the ARM Mali G-72, NVIDIA Xavier NX, NVIDIA TX2, and AMD embedded Ryzen V1605B using complex software operations for space workloads applications. The NVIDIA Xavier reported better performance for the matrix multiplication benchmark than the rest of the platforms. Regarding energy efficiency, the best choice has been either the NVIDIA Xavier NX or the TX2.

As part of the same project, in [435], Rodriguez *et al.* present a case study for an onboard infrared detector algorithm implementation using an embedded GPU. The algorithm is based on image matrix operations, targeting an NVIDIA Jetson Xavier's embedded GPU, using the CUDA cores for the algorithm acceleration and the Carmel CPU for offload computation. The CPU implementation, for a standard $2k \times 2k$ pixels image, reports performances of 98 times compared with the same algorithm on the LEON2 processor and 15.5 times on the PowerPC 750, while the GPU implementation reports performances of 806 and 128 times, respectively. The authors confirm that embedded GPUs can be an option to consider in future space missions [435].

In [467], [468], the same authors continue the analysis of the existing space application domains and survey the COTS and soft-IP embedded GPU domain to assess which meets the required computational power and identify the challenges which need to be addressed for their adoption in space.

Steenari *et al.* present in [462] a survey of high-performance processors and FPGAs for onboard processing and ML applications. The study includes as test device types the single and multicore processors, DPS and many-cores processors, embedded GPUs and FPGAs, while the considered data focus on the peak performance and size, qualification status, radiation support and ML tools availability. The authors conclude a significant gap in processing capabilities between RHBD/RT and COTS devices, as space processors need to catch up to commercial devices in terms of process technology and performance. However, the available COTS processors and accelerators have not been provided with RT-qualified packages. They also affirm that using high-performance COTS processors could cover future high-performance requirements for the onboard mission. However, they could compromise mission lifetime, availability and reliability as the COTS devices usually need longer-time support.

In [469], Marques *et al.* analyze the feasibility of ML applications for onboard space weather detection on COTS devices. The neural network was trained to distinguish between background and radiation particles. The authors expose a set of HW devices and development tools for ML and AI. The HW tested includes FPGA devices like Zynq-7000 SoC, Zynq Ultrascale+, Kintex Ultrascale and Versal ACAP; processors like the onboard space qualified processor (GR740/Leon4) and generic ARM Cortex A53 available in the Zynq US+ MPSoC; and SBC like Unibap iX5 and Myriad X.

Comparing the implementation results for FPGA-based platforms, Versal ACAP AI Core shows the best performance for inference on Coronal Mass Ejection (CME) Detection and Particle Detection. In contrast, Kintex Ultrascale KU040 show the worst results [469]. In the case of microprocessors, although the results are not as good as previous platforms, the processing time of a non-optimized implementation on a single-core LEON4 is already sufficient for onboard detection with better latency than detection on the ground after downlinking. At the same time, the CNN model for the ARM Cortex A53 MPSoC needs 5.8 times more time than the FPGA implementation on the same platform. In the case of the SBC platforms, the best results are achieved using the embedded GPU (Unibap iX5 GPU) for CME Detection inference. At the same time, Myriad X gets three times better results in Particle detection Inference.

This analysis demonstrates that using dedicated HW architecture for ML and AI, like the AI Engines or embedded GPU, makes a difference in increasing the performance of implementing these types of applications.

In [460], [470], XILINX presents an application note for the implementation of massive MIMO beamforming algorithms for 5G New Radio. As the beamforming requires significative compute density and advanced high-speed connectivity to meet 5G throughput and low latency, they propose using AI Engines available on the Versal AI Core VC1902 to increase the performance per Watt of the implementation. The authors claim an increase of 2.14 times the performance per Watt compared with a similar implementation over an Intel Agilex FPGA.

In [458], XILINX published a solution brief for AI inference with a Versal AI Core VC1902 for vision and video processing. The exploit of the Versal AI architecture results in an adaptable accelerator that exceeds the performance, latency and power efficiency of traditional FPGAs and GPUs for AI/ML Workloads. The authors ensure the delivery of 2.7 times the performance per Watt over competing with Intel Agilex FPGA, using 87 Watts of power, estimated using the XILINX Power Estimator (XPE) Tool. Even when this implementation proves this device's capability to accomplish highly complex operations on ML and AI applications, the power consumption is too high for onboard satellite applications.

Working on onboard device requirements, AMD announced the release of the first Space-Grade Versal Adaptive SoCs enabling onboard AI processing in space [434], [469], [471]. The actual information about this device pointed out that the XQR Versal AI Core XQRVC1902, to be available in early 2023, will be based on the XILINX Versal AI Core VC1902, detailed in Section V-B, including Class B qualification, radiation-tolerance and a $45 \times 45 \text{ mm}^2$ packing [471]. There is no information yet about the possible power consumption, its predecessor's principal inconvenience, and the chip's computer capacity.

VI. CHALLENGES AND OPEN DISCUSSION

While the literature covered in this survey provides an interesting overview of the main technical challenges for

specific use-cases, there are still a number of points that need to be further investigated to unleash the potential of ML into the satcom industry.

The main point of discussion resides on *where* the ML is deployed. If we opt for on-board deployment, there are many requirements in terms of power consumption (which depends on the tasks to be executed), the mass incurred by this additional block and the additional interfaces to interact with this block. So far, most flying satellites operate in a simple relaying-mode, with no data processing on-board except the frequency conversion and amplification. If the ML chip is deployed on-board, the specific input data and the required storage may add constraints to the payload design. At this point, we could not identify any previous work where a trade-off in terms of ML potential performance benefit versus added cost and complexity on the satellite device. To be noted that most of the HW components in a payload are duplicated to add redundancy to the satellite in case of failure. The latter should probably apply to ML-components as well if the play a key role in the operations of the satellite.

When the ML deployment is done on-ground, the requirements can be relaxed a bit in the sense that the consumed power and mass are not as critical as for a flying satellite. Similarly, the access to different data indicators on-ground is much affordable.

In the below sub-sections, we have aimed to separately address different challenges that we believe may act as leading reasearch directions in future ML for SatCom works.

A. Implementation Challenges

Implementation challenges in applying ML in the satellite communications industry are numerous and complex. One of the main challenges is determining the best place to deploy ML: should it be deployed on the ground or onboard the satellite?

If on-board deployment is chosen, limitations in terms of power consumption, mass, and the need for additional interfaces to interface with the ML block must be considered. In addition, data storage and specific data input may add constraints to the payload design. In addition, implementing ML components may require redundancy, further increasing satellite complexity and development costs [3].

On the other hand, if the ML is deployed on the ground, power and mass consumption requirements may be somewhat relaxed. Access to different data indicators on the ground is also more affordable. However, the challenges of data transmission back and forth between the satellite and the ground control center must be taken into account.

There are challenges in selecting the most appropriate ML algorithms for each specific task, the availability of training data, the extensive training length, validating and verifying the model to ensure its reliability, and implementing adequate security systems to protect the ML and the data it handles.

Implementation challenges also include the hardware required for Machine Learning implementation in the satellite communications industry.

On the one hand, high-performance hardware is required to run complex ML algorithms. This may include graphics

processing units (GPUs) or tensor processing units (TPUs) that are specifically used for Machine Learning computation acceleration. It is important to consider the radiation tolerance of the hardware used. SatCom systems operate in extremely hostile environments, including cosmic radiation, solar radiation, and solar particle radiation events. Therefore, hardware capable of withstanding these harsh environments and ensuring the long-term reliability and performance of the system is required.

Another challenge is the integration of the ML hardware into the satellite payload. The mass, power consumption and complexity of the hardware must be carefully considered to ensure that the complete system is robust enough for use in space.

In that regard, it is important to keep in mind that hardware technology continues to evolve rapidly and that hardware component costs can be significant for the development and implementation of ML-based NTN systems. Therefore, careful evaluation of long-term costs and benefits is required to ensure the economic viability of ML-based systems.

1) *Training datasets Availability*: The availability of real-world datasets is one of the most crucial preconditions for evaluating any proposed Machine Learning based method. So far, the cost and access difficulty of highly representative datasets has hindered the development in Satellite Communication of Machine Learning techniques. Indeed, while some learning approaches can be trained with small simulated datasets, advanced techniques such as Deep Learning and Deep Reinforcement Learning requires a large dataset to converge and reach the desired performance. It is time consuming and costly to generate large dataset for Satellite Communication using simulators. In addition, even assuming large datasets can be produced via simulations, their accuracy remains a key open discussion point. This review paper has shown that the development of the Machine Learning research does not go hand in hand with the development and availability of training datasets. Most of the datasets available opensource are datasets to train Machine Learning models on satellite imagery [472]–[474] and it is pretty rare to find open high-resolution datasets for the other use-cases presented in Section IV. This makes it a tremendous challenge to compute the accuracy and scalability of the proposed Machine Learning models in the real world.

2) *Data Dimensionality Reduction*: Using large data dimensions during the learning and testing increases computational complexity and slows the decision-making process. A promising solution to this problem is mapping the data from high dimension space to a lower one while preserving the most relevant and important features. Specifically, data dimensionality reduction is the process of scaling down a set of data with high dimensions into data with lower dimensions, ensuring that it concisely conveys similar information without affecting the original information. Data dimensionality reduction removes irrelevant, noisy, and redundant data to obtain acceptable accuracy. In satellite communication systems, high-dimensional datasets are naturally aggregated from heterogeneous data sources and massive data streams, which bring about many processing challenges. Thereby, it is essential to find optimal techniques or solutions for dimensionality reduction while catering for the effect of reducing high-dimensional data with-

out compromising the data value. Selecting an appropriate data dimension reduction technique will expedite the processing time and reduce the efforts required to extract insightful information.

The data representation space size can be reduced via feature extraction or selection techniques. The feature extraction transforms the initial data attributes into a new one formed by the linear or non-linear combination of the initial attributes. Contrarily, the feature selection method identifies the most relevant attributes according to some given criteria. In this context, different dimensionality reduction techniques are available in the literature to eliminate irrelevant and redundant features, and the work in [475] has reviewed the existing dimensionality reduction techniques and their suitability for different types of data and application areas with highlighting the issues that may affect the accuracy and relevance of the obtained results. Furthermore, an unsupervised deep-learning framework named local deep-feature alignment (LDFA) for dimension reduction has been developed in [476] to extract the essential features. In addition, two-dimensionality reduction techniques, Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) are investigated in [477] on some popular Machine Learning algorithms.

In this direction, the work in [478] has proposed using the Random Forest imputation method to extract useful information and reduce the search space to facilitate data exploration. Moreover, a combination of convolutional deep belief network (CDBN) with autoencoding feature algorithm is applied for data dimensionality reduction and layer-wise learning [479]. Specifically, CDBN is a hierarchical generative model that employs probabilistic max-pooling to extract high-level features by replacing the restricted Boltzmann machine (RBM) in the deep belief networks used for the individual layers by convolutional RBMs. Thus, this technique has been proved as an efficient algorithm for dimensionality reduction, collaborative filtering, and faster feature learning [480]. Likewise, Auto-Encoder and PCA approaches are used for dimensionality reduction in [481] for designing intrusion detection systems with machine learning. Reference [482] has considered multivariate statistical methods like independent component analysis (ICA) and minimum noise fraction (MNF) as dimensionality reduction techniques for processing satellite data.

3) *On-board deployment and Radiation Tolerance:* The deployment of ML onboard satellites faces a significant challenge in terms of radiation tolerance. Satellites are exposed to radiation from space that can cause electronic component failures. Radiation can affect the performance and lifetime of the hardware; therefore, it is necessary to design hardware components resistant to radiation [483].

ML onboard satellites are especially susceptible to radiation effects because the electronic components used for ML implementation are susceptible to radiation effects. Processing input data and implementing ML models require a large amount of computation, often performed by custom integrated circuits. Radiation exposure can cause errors in these circuits, resulting in hardware and software failures [483].

Therefore, it is essential to design radiation-tolerant hard-

ware for satellite-borne ML implementation. This may include the use of integrated circuits specific to aerospace applications, as well as design techniques to reduce radiation susceptibility. In addition, it is important to perform rigorous testing of hardware components to ensure that they are radiation resistant and capable of operating in the space environment

B. AI Accelerators

Successful deployment of machine learning algorithms in SatCom systems presents unique challenges. One of these is limited computational resources in space environments, where efficient and fast data processing is needed. AI acceleration technology, which uses specialized hardware to perform specific machine learning computations, has been developed to address this challenge [484].

AI accelerators can optimize the data processing and energy efficiency of SatCom systems, enabling greater data processing and transmission capacity on satellites. Accelerators can also improve the accuracy and speed of machine learning models, enabling better real-time decision-making [484].

However, implementing AI accelerators in SatCom systems also presents challenges despite the potential benefits. Integrating specialized hardware into existing systems can be costly and require significant changes to satellite infrastructure. In addition, careful design and rigorous validation are required to ensure the reliability and safety of SatCom systems using these advanced technologies.

1) *Neuromorphic Computing:* In the future, deploying ML for satellite communications will set new challenges for computer processors, which will have to support large workloads as efficiently as possible in harsh environmental conditions. Legacy computing technologies, including CPUs and GPUs, which are based on a semiconductor architecture, still may need to be suited for some ML workloads, particularly for unstructured time-sensitive signal processing. Besides, legacy technologies are expected to hit a "digital wall" in 2025, potentially forcing a paradigm shift in terms of computing technologies [485].

From this context, bio-inspired neuromorphic processors are a promising alternative to work as efficient coprocessors in ML applications where temporal signals are involved and/or require continuous adaptation in SATCOM systems. As opposed to conventional processors that operate in batch mode, i.e., collecting many samples before processing them, NPs can process in streaming mode. Owing to their energy efficiency and continuous on-board adaptability, NPs represent an excellent opportunity to open up the potential benefits of AI solutions for SATCOM systems. Moreover, NP implementation can take advantage of non-volatile memory devices based on technologies beyond CMOS [424].

C. Future AI Frontiers/Visions for Satellite Communication

1) *Federated Learning:* Satellite Communication's deployment of systems globally has led to an increasing demand for ML solutions operating in distributed systems. In this regard, the federated ML approach has emerged as a promising

solution for deploying machine learning models in global Internet of Things systems such as SatCom [486].

The federated ML approach refers to training ML models on multiple devices, each of which performs a part of the training process on its local data. Instead of sending data to a centralized server for training, the data is maintained on local devices, and only model updates are sent to the server. This collaborative learning approach can provide a scalable, decentralized solution for training models on large data sets without requiring data transfer between devices.

However, the federated ML approach also presents some technical challenges. First, local devices and the central server must communicate reliably to ensure data integrity and model updates. In addition, federated ML models need to be able to handle heterogeneous data and devices with different processing capabilities.

Another major challenge is ensuring data privacy and security. Because data is held on local devices, robust security and privacy measures must be implemented to protect user and organizational data.

2) *ML and Blockchain*: Security and resilience are features claimed by the new constellations and satellite systems, especially for earth observation applications or Government communications such as those provided by the GOVSATCOM project. The typical long distances in SATCOM lead to significant communication delays among the network nodes, compromising the data's integrity and reliability. Establishing prioritized, secure, and efficient command and data communication is essential for space and ground systems. Using Blockchain assures a secure automated network infrastructure for many applications in SATCOM, including autonomous satellite constellation control through satellite-to-satellite communication, prioritized and stable data and command communication between space and ground stations, among others. Similarly, Blockchain networks can benefit from SATCOM thanks to their broad geographic reach and broadcast capabilities. The Blockchain-based network nodes participate by receiving and sending data across the distributed network architecture. The data maintain a unique structure transferred by transactions [487]. These data can be anything, such as personally identifiable data, satellite command, or key/value paired asset data stored on blocks and linked together in a chain. Any alteration in a transaction detects an introduction or threat in the communication, directly invalidating the transmission. The Blockchain application to benefit SATCOM has been proposed for different use cases such as interference detection in [488], spectrum management in [489], satellite manufacturing supply chain in [490], and Space Situational Awareness in [491]. Different contributions of Blockchain in the space industry are identified in [492]. Some of them are (1) identity authentication and exchanged data integrity for LEO satellites which include its dynamic topology and regular link switching, and (2) reducing complexity across a range of business, operational, and security applications using smart contracts by BC networks. In [493], from the point of view of communication, a solution to reduce end-to-end latencies in SATCOM is proposed based on a BC-based reputation framework and routing protocol. ML in blockchain applications

applied to SATCOM is currently a crucial challenge where only a few preliminary works are presented in the literature. Notably, Space-Air-Ground Integrated Networks (Space-Air-Ground Integrated Network) project in [494] presented a federated reinforcement learning approach to trusted traffic offloading using BC. Conventional previous proposals are subject to increasing their benefits by adding ML techniques.

3) *Tiny Machine Learning for Satellite*: TinyML deploys Machine Learning algorithms onto low-cost, low-power and limited resource devices, storing Neural Network models directly within memory and directly running inference on onboard sensors' output. TinyML aims to create algorithms and software capable of performing efficiently on intelligent or limited resources edge devices. This approach avoids communicating with the cloud to transmit data for external processing.

TinyML aims to improve the efficiency of Deep Learning Artificial Intelligence systems by requiring far less computation power, a less costly hardware platform and improving energy efficiency. [495] shows that TinyML is capable of using a fraction of the compute resources needed for traditional ML systems. Finally, while the heterogeneity and limited resources of MCU devices present new challenges for on-device training, model updating, and deployment, recent research and the development of ML frameworks such as TensorFlow Lite for Microcontrollers have increased the accessibility of TinyML. Due to its nature, TinyML is one of the fastest growing areas of Deep Learning and might become a disruptive application in satellite communication.

4) *Quantum Computing and AI*: In this survey we have shown (Section IV-A11) the importance of quantum communication for Satellite Communication in its role of ultra-secure communication. Beyond this, quantum computing holds a great potential for AI developers looking to create more advanced AI systems that can solve problems faster and more efficiently. Specifically, quantum computing can make AI training process faster and more accurate as it allows utilizing larger datasets to train AI models to be more accurate and better at decision-making. Although quantum computing is still in its early development stages, there are several imminent benefits AI can gain from quantum computers, including but not limited to:

- Increased computational power: Quantum computers have the capability of solving complex optimization problems exponentially faster than classical computers. This could open up new possibilities for AI development such as enabling more complex machine learning algorithms to be trained in less time. For instance, Quantum computers can help AI developers to optimize complex functions more efficiently by leveraging quantum annealing and other optimization techniques that are not available in classical computing [496].
- Improved algorithm development: Quantum computing can help AI developers improve existing algorithms or develop new ones that can take advantage of the unique capabilities of quantum computing. For example, quantum machine learning algorithms can leverage quantum entanglement and superposition to achieve better accuracy and faster training.

- Enhanced data analysis: Quantum computing can help AI developers to analyze large datasets more efficiently and identify patterns that are difficult to detect using classical computing. This can lead to improved data-driven decision-making and insights.
- Quantum computers can provide a means of safeguarding the confidential data used by AI systems from hacking and other forms of cybercrime. Additionally, the parallel processing capabilities of quantum computing can be used to develop countermeasures against cyberattacks.

Furthermore, a major challenge in the operation of the CubeSats and small satellites within NTN in lower altitudes is the rather low information processing capabilities of the onboard processors. Consequently, developing and utilizing AI models in these entities can be beyond the capability of a single satellite processor. Alternatively, quantum technologies along with space-based quantum clouds can be employed in such scenarios in order to offload the computational burden from small satellites. Thus, a space quantum network can be structured and interconnected via FSO links, which will benefit from the high computational capacity of the quantum servers with certainly enhanced security performance to efficiently use AI and ML techniques [497].

5) *AI and Optimization Methods*: This paper has shown that Machine Learning and Artificial Intelligence have tremendous potential in future Satellite Communication networks. Thus, while it is clear that the next generation of satellites will rely on Artificial Intelligence, it is unclear how Artificial Intelligence should be integrated into the architecture of satellite networks. Most of the data-driven approaches surveyed in this paper (see table III) are based on Neural Network and Deep Reinforcement Learning techniques that require a large amount of data and are complex to process. Therefore, it is convenient to investigate that to mitigate the computational complexity, latency, and power consumption.

Hybrid Machine Learning models allow incorporating expert knowledge in the learning approach. The external knowledge can be incorporated as an initialization procedure [498] or leveraged in the learning approach, as in deep unfolding [499]. In the former, the training set for the Deep Neural Network can be generated offline, alleviating the satellite of the burden of dealing with the high complexity of the model. The model can be thus optimized online with few empirical data. The intuition behind this approach is that, despite generally not being perfectly accurate, the initial Deep Neural Network contains some critical feature of the satellite network, and less data is needed to reach the desired accuracy compared to the case of no-training being performed. In the latter, Neural Networks impact the complexity and computational requirement of the Machine Learning model on the satellite. For instance, deep unfolding aims to map an iterative algorithm into a learnable Neural Network and to determine the number of nodes by the number of iterations of the algorithm [293]. Deep unfolding offers the possibility of initializing a Neural Network and refining it using empirical data. The first attempts to utilize Deep Unfolding for Satellite Communication applications have been mentioned in the Rate Splitting Use Case with the mentioned works [121], [122]. To reduce the complexity of the

power allocation and timeslot-terminal assignment in satellite NOMA systems, we have shown in this paper that the work in [116] makes use of a hybrid approach to low the complexity of the problem without, however, proposing an end to end Neural Network that may not be feasible.

While an extensive description of hybrid ML models is beyond the scope of this paper, we believe incorporating expert knowledge in Machine Learning models will receive increasing attention in the next future to overcome the conventional NN complexity and the relative challenges presented in Section VI-A.

D. Security Aspects

We have shown in this survey how big data is the core of many Satellite Communication research activities. In fact, Image Processing and Remote Sensing Artificial Intelligence developments rely on the data collected by satellites and drones and use it to train models for different outputs and use-cases. For this reason, Artificial Intelligence/Machine Learning tools used to handle and analyze diverse sources of data are closely linked with several ethical issues, not only the well known issue of privacy. Together with it, explainability, bias also ethical issues are becoming increasingly important as a result of data collection, the approaches used for analysis, the manner and purpose for which the results from such analyzed are used [500].

Security aspects are crucial in the satellite communications industry and the use of ML introduces new challenges. One of the biggest risks is the potential manipulation of training data and exploitation of vulnerabilities in the trained ML model, which could adversely affect the quality of transmitted data and the integrity of the SatCom system. In addition, the use of real-time ML systems also introduces the possibility of real-time attacks, such as malicious data injection into the ML model.

An additional major challenge is a need to ensure the privacy and security of the data that is used to train the ML models. In particular, the collected and processed data may be privacy sensitive, such as location or enterprise customer data. Appropriate measures must be taken to protect this data and ensure it is only used for legitimate purposes.

The security of the hardware and software used to implement the ML models is also an important consideration. SatCom systems must be designed and built with security in mind, including protection against physical and logical attacks. The exploitation of vulnerabilities in the hardware or software used to implement the ML could have serious consequences for the integrity of the SatCom system.

E. Cost Optimization

Cost optimization is a critical factor in the satellite communications industry, and the deployment of artificial intelligence technologies is no exception. When considering the deployment of ML systems, the costs of the necessary components and the resources required to train and maintain the ML models must be considered. In addition, the cost of putting

a satellite into orbit must also be considered, and cost-benefit optimization of ML systems must be sought.

The cost of deploying a Gbps in orbit can be extremely high [501]. The weight of the satellite components and associated power consumption can also be costly regarding satellite payload capacity and battery life. Therefore, more efficient and cost-effective design solutions, such as using low-power data processing technologies and optimizing ML algorithms to minimize the resources required, should be explored.

Cost optimization must also consider the scalability of ML systems. As more and more satellites are deployed, scalable and cost-effective solutions must be found to train and maintain ML models on all satellites. This may include using distributed computing technologies and implementing automation solutions to reduce maintenance and training costs.

REFERENCES

- [1] G. Geraci, D. Lopez-Perez, M. Benzaghta, and S. Chatzinotas, "Integrating terrestrial and non-terrestrial networks: 3d opportunities and challenges," *IEEE Communications Magazine*, 2022.
- [2] M. Á. Vázquez, P. Henarejos, A. I. Pérez-Neira, E. Grechi, A. Voight, J. C. Gil, I. Pappalardo, F. Di Credico, and R. M. Lancellotti, "On the Use of AI For Satellite Communications," *arXiv preprint arXiv:2007.10110*, 2020.
- [3] F. Ortiz, V. Monzon Baeza, L. M. Garces-Socarras, J. A. Vázquez-Peralvo, J. L. Gonzalez, G. Fontanesi, E. Lagunas, J. Querol, and S. Chatzinotas, "Onboard processing in satellite communications using ai accelerators," *Aerospace*, vol. 10, no. 2, p. 101, 2023.
- [4] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine Learning Paradigms for Next-generation Wireless Networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, 2016.
- [5] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, and X. Costa-Pérez, "A Machine Learning Approach to 5G Infrastructure Market Optimization," *IEEE Trans. Mob. Comput.*, vol. 19, no. 3, pp. 498–512, 2019.
- [6] M. M. Azari, S. Solanki, S. Chatzinotas, O. Kodheli, H. Sallouha, A. Colpaert, J. F. M. Montoya, S. Pollin, A. Haqiqatnejad, A. Mostaani *et al.*, "Evolution of Non-terrestrial Networks from 5G to 6G: A survey," *IEEE Commun. Surveys Tuts.*, 2022.
- [7] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff *et al.*, "Satellite Communications in the New Space Era: A Survey and Future Challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 70–109, 2020.
- [8] T. Hoeser and C. Kuenzer, "Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends," *Remote Sensing*, vol. 12, no. 10, p. 1667, 2020.
- [9] S. Walsh, D. Murphy, M. Doyle, J. Reilly, J. Thompson, R. Dunwoody, J. Erkal, G. Finneran, G. Fontanesi, J. Mangan *et al.*, "Development of the eirsat-1 cubesat through functional verification of the engineering qualification model," *Aerospace*, vol. 8, no. 9, p. 254, 2021.
- [10] N. Saeed, A. Elzanaty, H. Almorad, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "Cubesat communications: Recent advances and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1839–1862, 2020.
- [11] P. K. Chowdhury, M. Atiquzzaman, and W. Ivancic, "Handover Schemes in Satellite Networks: State-of-the-art and Future Research Directions," *IEEE Communications Surveys & Tutorials*, vol. 8, no. 4, pp. 2–14, 2006.
- [12] H. Al-Hraishawi, H. Chougrani, S. Kisseleff, E. Lagunas, and S. Chatzinotas, "A Survey on Non-geostationary Satellite Systems: The Communication Perspective," *IEEE Commun. Surveys Tuts.*, 2022.
- [13] L. M. Marrero, J. C. Merlano-Duncan, J. Querol, S. Kumar, J. Krivochiza, S. K. Sharma, S. Chatzinotas, A. Camps, and B. Ottersten, "Architectures and Synchronization Techniques for Distributed Satellite Systems: A Survey," *IEEE Access*, vol. 10, pp. 45 375–45 409, 2022.
- [14] P.-D. Arapoglou, K. Liolis, M. Bertinelli, A. Panagopoulos, P. Cottis, and R. De Gaudenzi, "Mimo over Satellite: A Review," *IEEE communications surveys & tutorials*, vol. 13, no. 1, pp. 27–51, 2010.
- [15] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite Communications Supporting Internet of Remote Things," *IEEE Internet Things J.*, vol. 3, no. 1, pp. 113–123, 2015.
- [16] M. Centenaro, C. E. Costa, F. Granelli, C. Sacchi, and L. Vangelista, "A survey on technologies, standards and open challenges in satellite IoT," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1693–1720, 2021.
- [17] B. A. Homssi, K. Dakic, K. Wang, T. Alpcan, B. Allen, S. Kandeeppan, A. Al-Hourani, and W. Saad, "Artificial Intelligence Techniques for Next-Generation Mega Satellite Networks," *arXiv preprint arXiv:2207.00414*, 2022.
- [18] F. Fourati and M.-S. Alouini, "Artificial Intelligence for Satellite Communication: A Review," *Intelligent and Converged Networks*, vol. 2, no. 3, pp. 213–243, 2021.
- [19] V. Kothari, E. Liberis, and N. D. Lane, "The final frontier: Deep learning in space," in *Proceedings of the 21st international workshop on mobile computing systems and applications*, 2020, pp. 45–49.
- [20] M. Lofqvist and J. Cano, "Accelerating deep learning applications in space," *arXiv preprint arXiv:2007.11089*, 2020.
- [21] A. Russo and G. Lax, "Using Artificial Intelligence for Space Challenges: A Survey," *Applied Sciences*, vol. 12, no. 10, p. 5106, 2022.
- [22] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to pareto-optimal wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1472–1514, 2020.
- [23] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [24] X. Zhu and C. Jiang, "Integrated satellite-terrestrial networks toward 6g: Architectures, applications, and challenges," vol. 9, no. 1, pp. 437–461, conference Name: IEEE Internet of Things Journal.
- [25] R. De Gaudenzi, M. Luise, and L. Sanguinetti, "The open challenge of integrating satellites into (beyond-) 5g cellular networks," *IEEE Network*, vol. 36, no. 2, pp. 168–174, 2022.
- [26] F. G. Ortiz-Gomez, L. Lei, E. Lagunas, R. Martinez, D. Tarchi, J. Querol, M. A. Salas-Natera, and S. Chatzinotas, "Machine Learning for Radio Resource Management in Multibeam GEO Satellite Systems," *Electronics*, vol. 11, no. 7, p. 992, 2022.
- [27] R. G. Smith and J. Eckroth, "Building ai applications: Yesterday, today, and tomorrow," *AI Magazine*, vol. 38, no. 1, pp. 6–22, Mar. 2017. [Online]. Available: <https://ojs.aaai.org/index.php/aimagazine/article/view/2709>
- [28] T. M. Mitchell and T. M. Mitchell, *Machine Learning*. McGraw-hill New York, 1997, vol. 1, no. 9.
- [29] Y. Zhao, Y. Li, X. Zhang, G. Geng, W. Zhang, and Y. Sun, "A survey of networking applications applying the software defined networking concept based on machine learning," *IEEE Access*, vol. 7, pp. 95 397–95 417, 2019.
- [30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [31] W. G. Hatcher and W. Yu, "A survey of deep learning: Platforms, applications and emerging research trends," *IEEE Access*, vol. 6, pp. 24 411–24 432, 2018.
- [32] R. Collobert and S. Bengio, "Links between perceptrons, mlps and svms," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 23.
- [33] V. Y. Kulkarni and P. K. Sinha, "Pruning of random forest classifiers: A survey and future directions," in *2012 International Conference on Data Science & Engineering (ICDSE)*. IEEE, 2012, pp. 64–68.
- [34] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications surveys & tutorials*, vol. 18, no. 2, pp. 1153–1176, 2015.
- [35] S. Na, L. Xumin, and G. Yong, "Research on k-means clustering algorithm: An improved k-means clustering algorithm," in *2010 Third International Symposium on intelligent information technology and security informatics*. Ieee, 2010, pp. 63–67.
- [36] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.
- [37] R. L. Stratonovich, "Conditional markov processes," in *Non-linear transformations of stochastic processes*. Elsevier, 1965, pp. 427–453.
- [38] P. K. Mallapragada, R. Jin, A. K. Jain, and Y. Liu, "Semiboost: Boosting for semi-supervised learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 11, pp. 2000–2014, 2008.

- [39] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1259–1270, 2017.
- [40] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, C. Wang, and Y. Liu, "A survey of machine learning techniques applied to software defined networking (sdn): Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393–430, 2018.
- [41] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, and F. L. Lewis, "Optimal and autonomous control using reinforcement learning: A survey," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 6, pp. 2042–2062, 2017.
- [42] M. Elsayed and M. Erol-Kantarci, "Ai-enabled future wireless networks: Challenges, opportunities, and open issues," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 70–77, 2019.
- [43] M. Qiao, H. Zhao, L. Zhou, C. Zhu, and S. Huang, "Topology-transparent scheduling based on reinforcement learning in self-organized wireless networks," *IEEE Access*, vol. 6, pp. 20221–20230, 2018.
- [44] V. Raj, I. Dias, T. Tholeti, and S. Kalyani, "Spectrum access in cognitive radio using a two-stage reinforcement learning approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 20–34, 2018.
- [45] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Reinforcement learning with network-assisted feedback for heterogeneous rat selection," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6062–6076, 2017.
- [46] S. Zhang and D. Zhu, "Towards artificial intelligence enabled 6g: State of the art, challenges, and opportunities," *Computer Networks*, vol. 183, p. 107556, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S138912862031207X>
- [47] T. M. Braun and W. R. Braun, *PROCESSING PAYLOAD AND FLEXIBLE PAYLOAD*, 2021, pp. 270–296.
- [48] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud ran for mobile networks—a technology overview," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 405–426, 2015.
- [49] T. Mortlock and Z. M. Kassas, "Assessing machine learning for leo satellite orbit determination in simultaneous tracking and navigation," in *2021 IEEE Aerospace Conference (50100)*, 2021, pp. 1–8.
- [50] M. Jalali, F. G. Ortiz-Gomez, E. Lagunas, S. Kisseleff, L. Emiliani, and S. Chatzinotas, "Radio regulation compliance of ngso constellations' interference towards gso ground stations," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2022, pp. 1425–1430.
- [51] S. Xu, X.-W. Wang, and M. Huang, "Software-defined next-generation satellite networks: Architecture, challenges, and solutions," *IEEE Access*, vol. 6, pp. 4027–4041, 2018.
- [52] J. Tronc, P. Angeletti, N. Song, M. Haardt, J. Arendt, and G. Gallinaro, "Overview and Comparison of On-ground and on-board Beamforming Techniques in Mobile Satellite Service Applications," *Int. J. Satell. Commun.*, vol. 32, no. 4, pp. 291–308, 2014.
- [53] G. Fontanesi, A. Zhu, M. Arvaneh, and H. Ahmadi, "A transfer learning approach for uav path design with connectivity outage constraint," *IEEE Internet of Things Journal*, 2022.
- [54] G. Mateo-Garcia, J. Veitch-Michaelis, L. Smith, S. V. Oprea, G. Schumann, Y. Gal, A. G. Baydin, and D. Backes, "Towards global flood mapping onboard low cost satellites with machine learning," *Scientific reports*, vol. 11, no. 1, pp. 1–12, 2021.
- [55] S. Dhar, J. Guo, J. Liu, S. Tripathi, U. Kurup, and M. Shah, "A Survey of On-Device Machine Learning: An Algorithms and Learning Theory Perspective," *ACM Trans. Internet of Things*, vol. 2, no. 3, pp. 1–49, 2021.
- [56] F. Mendoza, M. Minardi, S. Chatzinotas, L. Lei, and T. X. Vu, "An sdn based testbed for dynamic network slicing in satellite-terrestrial networks," in *2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*. IEEE, 2021, pp. 36–41.
- [57] M. Minardi, Y. Drif, T. X. Vu, I. Maity, C. Politis, and S. Chatzinotas, "SDN-based Testbed for Emerging Use Cases in Beyond 5G NTN-Terrestrial Networks," in *2nd International Workshop on Autonomous Network Management in 5G and Beyond Systems*, Miami, FL, USA, May 2023, pp. 1–5.
- [58] "The new space age: Ibm develops a unique, custom edge computing solution in space," <https://www.ibm.com/cloud/blog/ibm-develops-a-unique-custom-edge-computing-solution-in-space>.
- [59] "Stream b: Research for revolutionary technology advancement towards 6g - smart networks," <https://smart-networks.europa.eu/stream-b-research-for-revolutionary-technology-advancement-towards-6g/>, fecha de acceso: 20-03-2023.
- [60] "Objectives - atria," <https://www.atria-h2020.eu/objectives/>, fecha de acceso: 20-03-2023.
- [61] "Dynasat," <https://www.dynasat.eu/>, fecha de acceso: 20-03-2023.
- [62] K. Chen, W. Wang, X. Chen, and H. Yin, "Deep learning based antenna array fault detection," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*. IEEE, 2019, pp. 1–5.
- [63] N. Xu, C. G. Christodoulou, S. E. Barbin, and M. Martinez-Ramon, "Detecting failure of antenna array elements using machine learning optimization," in *2007 IEEE Antennas and Propagation Society International Symposium*, 2007, pp. 5753–5756.
- [64] R. Aysterán and F. Las-Heras, "Near field to far field transformation using neural networks and source reconstruction," *Journal of Electromagnetic Waves and Applications*, vol. 20, no. 15, pp. 2201–2213, 2006.
- [65] D. Vakula and N. V. Sarma, "Using neural networks for fault detection in planar antenna arrays," *Progress In Electromagnetics Research Letters*, vol. 14, pp. 21–30, 2010.
- [66] S. Bianco, P. Napoletano, A. Raimondi, M. Feo, G. Petraglia, and P. Vinetti, "AESA Adaptive Beamforming Using Deep Learning," in *2020 IEEE Radar Conference (RadarConf20)*. IEEE, 2020, pp. 1–6.
- [67] A. J. Singh and M. Jayakumar, "Machine learning based digital beamforming for line-of-sight optimization in satcom on the move technology," in *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2020, pp. 422–427.
- [68] I. Mallioras, Z. D. Zaharis, P. I. Lazaridis, and S. Pantelopoulou, "A Novel Realistic Approach of Adaptive Beamforming Based on Deep Neural Networks," *IEEE Trans. Antennas Propag.*, 2022.
- [69] R. Kumar and S. Arnon, "Dnn beamforming for leo satellite communication at sub-thz bands," *Electronics*, vol. 11, no. 23, p. 3937, 2022.
- [70] W. Li, X. Huang, and H. Leung, "Performance evaluation of digital beamforming strategies for satellite communications," *IEEE Transactions on Aerospace and Electronic systems*, vol. 40, no. 1, pp. 12–26, 2004.
- [71] L. Lei, E. Lagunas, Y. Yuan, M. G. Kibria, S. Chatzinotas, and B. Ottersten, "Deep learning for beam hopping in multibeam satellite systems," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. IEEE, 2020, pp. 1–5.
- [72] —, "Beam illumination pattern design in satellite networks: Learning and optimization for efficient beam hopping," *IEEE Access*, vol. 8, pp. 136 655–136 667, 2020.
- [73] X. Hu, Y. Zhang, X. Liao, Z. Liu, W. Wang, and F. M. Ghannouchi, "Dynamic beam hopping method based on multi-objective deep reinforcement learning for next generation satellite broadband systems," *IEEE Transactions on Broadcasting*, vol. 66, no. 3, pp. 630–646, 2020.
- [74] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Dynamic beam pattern and bandwidth allocation based on multi-agent deep reinforcement learning for beam hopping satellite systems," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3917–3930, 2022.
- [75] F. de Guadalupe Ortíz Gómez, R. M. Rodríguez-Orsorio, M. A. S. Natera, and S. L. Ayala, "On the use machine learning for flexible payload management in vhts systems," in *Proceedings of 70th International Astronautical Congress 2019*, 2019, pp. 1–6. [Online]. Available: <https://oa.upm.es/64705/>
- [76] F. de Guadalupe Ortíz Gómez, R. M. Rodríguez-Orsorio, M. A. S. Natera, S. L. Ayala, D. Tarchi, and A. Vanelli-Coralli, "On the use of neural networks for flexible payload management in vhts systems," in *25th Ka Broadband Commun. Conf.*, 2019, pp. 1–10.
- [77] P. Zhang, X. Wang, Z. Ma, S. Liu, and J. Song, "An online power allocation algorithm based on deep reinforcement learning in multibeam satellite systems," *International Journal of Satellite Communications and Networking*, vol. 38, no. 5, pp. 450–461, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1352>
- [78] J. J. G. Luis, M. Guerster, I. del Portillo, E. Crawley, and B. Cameron, "Deep reinforcement learning for continuous power allocation in flexible high throughput satellites," in *2019 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAaw)*, 2019, pp. 1–4.
- [79] J. J. G. Luis, N. Pachler, M. Guerster, I. del Portillo, E. Crawley, and B. Cameron, "Artificial intelligence algorithms for power allocation in high throughput satellites: A comparison," in *2020 IEEE Aerospace Conference*, 2020, pp. 1–15.
- [80] T. S. Abdu, S. Kisseleff, L. Lei, E. Lagunas, J. Grotz, and S. Chatzinotas, "A deep learning based acceleration of complex satellite resource

- management problem,” in *2022 30th European Signal Processing Conference (EUSIPCO)*, 2022, pp. 1092–1096.
- [81] S. Liu, X. Hu, and W. Wang, “Deep reinforcement learning based dynamic channel allocation algorithm in multibeam satellite systems,” *IEEE Access*, vol. 6, pp. 15 733–15 742, 2018.
- [82] B. Zhao, J. Liu, Z. Wei, and I. You, “A deep reinforcement learning based approach for energy-efficient channel allocation in satellite internet of things,” *IEEE Access*, vol. 8, pp. 62 197–62 206, 2020.
- [83] Y. He, B. Sheng, H. Yin, D. Yan, and Y. Zhang, “Multi-objective deep reinforcement learning based time-frequency resource allocation for multi-beam satellite communications,” *China Communications*, vol. 19, no. 1, pp. 77–91, 2022.
- [84] F. G. Ortiz-Gomez, D. Tarchi, R. Martínez, A. Vanelli-Coralli, M. A. Salas-Natera, and S. Landeros-Ayala, “Convolutional neural networks for flexible payload management in vhts systems,” *IEEE Systems Journal*, vol. 15, no. 3, pp. 4675–4686, 2021.
- [85] —, “Cooperative multi-agent deep reinforcement learning for resource management in full flexible vhts systems,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 335–349, 2022.
- [86] J. Ebert, W. Bailer, J. Flavio, K. Plimon, and M. Winter, “A method for acm on q/v-band satellite links based on artificial intelligence,” in *2020 10th Advanced Satellite Multimedia Systems Conference and the 16th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, 2020, pp. 1–5.
- [87] A. Tsakmalis, S. Chatzinotas, and B. Ottersten, “Interference constraint active learning with uncertain feedback for cognitive radio networks,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4654–4668, 2017.
- [88] L. Pellaco, V. Saxena, M. Bengtsson, and J. Jaldén, “Wireless link adaptation with outdated csi — a hybrid data-driven and model-based approach,” in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.
- [89] D. Lee, Y. G. Sun, I. Sim, J.-H. Kim, Y. Shin, D. I. Kim, and J. Y. Kim, “Neural episodic control-based adaptive modulation and coding scheme for inter-satellite communication link,” *IEEE Access*, vol. 9, pp. 159 175–159 186, 2021.
- [90] X. Wang, H. Li, and Q. Wu, “Optimizing adaptive coding and modulation for satellite network with ml-based csi prediction,” in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, 2019, pp. 1–6.
- [91] P. Henarejos, M. A. Vazquez, and A. I. Perez-Neira, “Deep learning for experimental hybrid terrestrial and satellite interference management,” in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019, pp. 1–5.
- [92] M. Kulin, T. Kazaz, I. Moerman, and E. De Poorter, “End-to-End Learning From Spectrum Data: A Deep Learning Approach for Wireless Signal Identification in Spectrum Monitoring Applications,” *IEEE Access*, vol. 6, pp. 18 484–18 501, 2018.
- [93] W. Qin and F. Dovis, “Situational awareness of chirp jamming threats to gnss based on supervised machine learning,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 3, pp. 1707–1720, 2022.
- [94] Y. Wang, C. Zhang, Q. Peng, and Z. Wang, “Learning to detect frame synchronization,” in *Neural Information Processing*, M. Lee, A. Hirose, Z.-G. Hou, and R. M. Kil, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 570–578.
- [95] J. Wang, W. Tu, L. C. K. Hui, S. M. Yiu, and E. K. Wang, “Detecting time synchronization attacks in cyber-physical systems with machine learning techniques,” in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017, pp. 2246–2251.
- [96] S. Lee, S. Kim, M. Seo, and D. Har, “Synchronization of frequency hopping by lstm network for satellite communication system,” *IEEE Communications Letters*, vol. 23, no. 11, pp. 2054–2058, 2019.
- [97] D. Zibar, L. H. H. de Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, J. Estaran, H. Haisch, N. G. Gonzalez, J. C. R. F. de Oliveira, and I. T. Monroy, “Application of machine learning techniques for amplitude and phase noise characterization,” *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1333–1343, 2015.
- [98] H. Wu, Z. Sun, and X. Zhou, “Deep learning-based frame and timing synchronization for end-to-end communications,” *Journal of Physics: Conference Series*, vol. 1169, pp. 012–060, Feb 2019.
- [99] V.-P. Bui, T. Van Chien, E. Lagunas, J. Grotz, S. Chatzinotas, and B. Ottersten, “Robust congestion control for demand-based optimization in precoded multi-beam high throughput satellite communications,” *IEEE Transactions on Communications*, vol. 70, no. 10, pp. 6918–6937, 2022.
- [100] L. Bai, Q. Xu, S. Wu, S. Ventouras, and G. Goussetis, “A Novel Atmosphere-Informed Data-Driven Predictive Channel Modeling for B5G/6G Satellite-Terrestrial Wireless Communication Systems at Q-Band,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14 225–14 237, 2020.
- [101] L. Bai, C.-X. Wang, Q. Xu, S. Ventouras, and G. Goussetis, “Prediction of Channel Excess Attenuation for Satellite Communication Systems at Q-Band Using Artificial Neural Network,” *IEEE Antennas and Wireless Propagation Letters*, vol. 18, no. 11, pp. 2235–2239, 2019.
- [102] A. Cornejo, S. Landeros-Ayala, J. M. Matias, F. Ortiz-Gomez, R. Martinez, and M. Salas-Natera, “Method of Rain Attenuation Prediction Based on Long-Short Term Memory Network,” *Neural Processing Letters*, vol. 54, no. 4, pp. 2959–2995, 2022.
- [103] A. Ferdowski and D. Whitefield, “Deep Learning for Rain Fade Prediction in Satellite Communications,” in *2021 IEEE Globecom Workshops (GC Wkshps)*, 2021, pp. 1–6.
- [104] L. Sun, P. Chu, and R. Zhu, “Navigation signal radio frequency channel modeling and predistortion technology based on artificial intelligence technology and neural network,” *Mobile Information Systems*, vol. 2022, 2022.
- [105] P. Henarejos and M. Ángel Vázquez, “Decoding 5g-nr communications via deep learning,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 3782–3786.
- [106] K. Niu, J. Dai, K. Tan, and J. Gao, “Deep learning methods for channel decoding: A brief tutorial,” in *2021 IEEE/CIC International Conference on Communications in China (ICCC)*, 2021, pp. 144–149.
- [107] E. Nachmani, Y. Be’ery, and D. Burshtein, “Learning to decode linear codes using deep learning,” in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2016, pp. 341–346.
- [108] N. Agrawal, “Machine intelligence in decoding of forward error correction codes.” Ph.D. dissertation, 10 2017.
- [109] L. Lugosch, *Learning Algorithms for Error Correction*, ser. McGill theses. McGill University Libraries, 2018. [Online]. Available: <https://books.google.es/books?id=EyANugEACAAJ>
- [110] A. Haroon, F. Hussain, and M. R. Bajwa, “Decoding of error correcting codes using neural networks,” 2013.
- [111] A. Guidotti and A. Vanelli-Coralli, “Clustering strategies for multicast precoding in multibeam satellite systems,” *International Journal of Satellite Communications and Networking*, vol. 38, no. 2, pp. 85–104, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1312>
- [112] F. Ortiz, E. Lagunas, and S. Chatzinotas, “Unsupervised learning for user scheduling in multibeam precoded geo satellite systems,” in *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2022, pp. 190–195.
- [113] J. Luo, J. Tang, D. K. C. So, G. Chen, K. Cumanan, and J. A. Chambers, “A deep learning-based approach to power minimization in multi-carrier noma with swipt,” *IEEE Access*, vol. 7, pp. 17 450–17 460, 2019.
- [114] X. Yan, K. An, C.-X. Wang, W.-P. Zhu, Y. Li, and Z. Feng, “Genetic algorithm optimized support vector machine in noma-based satellite networks with imperfect csi,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8817–8821.
- [115] V. Andiappan and V. Ponnusamy, “Deep learning enhanced noma system: A survey on future scope and challenges,” *Wireless Personal Communications*, vol. 123, no. 1, pp. 839–877, Mar 2022. [Online]. Available: <https://doi.org/10.1007/s11277-021-09160-1>
- [116] A. Wang, L. Lei, E. Lagunas, S. Chatzinotas, and B. Ottersten, “Dual-dnn assisted optimization for efficient resource scheduling in noma-enabled satellite systems,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.
- [117] —, “Completion time minimization in noma systems: Learning for combinatorial optimization,” *IEEE Networking Letters*, vol. 3, no. 1, pp. 15–18, 2021.
- [118] Y. Sun, Y. Wang, J. Jiao, S. Wu, and Q. Zhang, “Deep learning-based long-term power allocation scheme for NOMA downlink system in S-LoT,” *IEEE Access*, vol. 7, pp. 86 288–86 296, 2019.
- [119] X. Yan, K. An, Q. Zhang, G. Zheng, S. Chatzinotas, and J. Han, “Delay constrained resource allocation for NOMA enabled satellite internet of things with deep reinforcement learning,” *IEEE Internet of Things Journal*, 2020.
- [120] Q. Liu, J. Jiao, S. Wu, R. Lu, and Q. Zhang, “Deep reinforcement learning-assisted noma age-optimal power allocation for s-iot network,”

- in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 1823–1828.
- [121] Q. Zhang and L. Zhu, “A deep learning approach for downlink sum rate maximization in satellite-terrestrial integrated network,” in *2022 International Symposium on Networks, Computers and Communications (ISNCC)*, 2022, pp. 1–5.
- [122] Q. Zhang, L. Zhu, S. Jiang, and X. Tang, “Deep unfolding for cooperative rate splitting multiple access in hybrid satellite terrestrial networks,” *China Communications*, vol. 19, no. 7, pp. 100–109, 2022.
- [123] J. Huang, Y. Yang, L. Yin, D. He, and Q. Yan, “Deep reinforcement learning-based power allocation for rate-splitting multiple access in 6g leo satellite communication system,” *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2185–2189, 2022.
- [124] N. Kato, Z. M. Fadlullah, F. Tang, B. Mao, S. Tani, A. Okamura, and J. Liu, “Optimizing space-air-ground integrated networks by artificial intelligence,” *IEEE Wireless Communications*, vol. 26, no. 4, pp. 140–147, 2019.
- [125] A. Cigliano and F. Zampognaro, “A machine learning approach for routing in satellite mega-constellations,” in *2020 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, 2020, pp. 1–6.
- [126] C. Qiu, H. Yao, F. R. Yu, F. Xu, and C. Zhao, “Deep q-learning aided networking, caching, and computing resources allocation in software-defined satellite-terrestrial networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5871–5883, 2019.
- [127] J. Liu, B. Zhao, Q. Xin, and H. Liu, “Dynamic channel allocation for satellite internet of things via deep reinforcement learning,” in *2020 International Conference on Information Networking (ICOIN)*, 2020, pp. 465–470.
- [128] D. A. Tubiana, J. Farhat, G. Brante, and R. D. Souza, “Q-learning NOMA random access for IoT-satellite terrestrial relay networks,” *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1619–1623, 2022.
- [129] B. Zhao, G. Ren, X. Dong, and H. Zhang, “Distributed q-learning based joint relay selection and access control scheme for iot-oriented satellite terrestrial relay networks,” *IEEE Communications Letters*, vol. 25, no. 6, pp. 1901–1905, 2021.
- [130] M. V. da Silva, R. D. Souza, A. Alves, and T. Abrão, “A noma-based q-learning random access method for machine type communications,” *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1720–1724, 2020.
- [131] D. Zhou, M. Sheng, Y. Wang, J. Li, and Z. Han, “Machine learning-based resource allocation in satellite networks supporting internet of remote things,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6606–6621, 2021.
- [132] X. Liu, H. Zhang, K. Long, A. Nallanathan, and V. C. M. Leung, “Deep dyna-reinforcement learning based on random access control in LEO satellite IoT networks,” *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 14 818–14 828, 2022.
- [133] L. Lei, Y. Yuan, T. X. Lu, S. Chatzinotas, M. Minardi, and J. F. M. Montoya, “Dynamic-adaptive ai solutions for network slicing management in satellite-integrated b5g systems,” *IEEE Network*, vol. 35, no. 6, pp. 91–97, 2021.
- [134] I. Bisio, F. Lavagetto, G. Verardo, and T. de Cola, “Network slicing optimization for integrated 5g-satellite networks,” in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, ISSN: 2576-6813.
- [135] T. K. Rodrigues and N. Kato, “Network slicing with centralized and distributed reinforcement learning for combined satellite/ground networks in a 6g environment,” vol. 29, no. 1, pp. 104–110, conference Name: IEEE Wireless Communications.
- [136] C. Zhou, W. Wu, H. He, P. Yang, F. Lyu, N. Cheng, and X. Shen, “Deep reinforcement learning for delay-oriented iot task scheduling in sagin,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 911–925, 2021.
- [137] N. Cheng, F. Lyu, W. Quan, C. Zhou, H. He, W. Shi, and X. Shen, “Space/aerial-assisted computing offloading for iot applications: A learning-based approach,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 5, pp. 1117–1129, 2019.
- [138] M. S. Elbamby, C. Perfecto, C.-F. Liu, J. Park, S. Samarakoon, X. Chen, and M. Bennis, “Wireless edge computing with latency and reliability guarantees,” *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1717–1737, 2019.
- [139] S. Semanjski, I. Semanjski, W. De Wilde, and A. Muls, “Use of supervised machine learning for GNSS signal spoofing detection with validation on real-world meaconing and spoofing data—part i,” *Sensors*, vol. 20, no. 4, Feb. 2020.
- [140] S. Semanjski, A. Muls, I. Semanjski, and W. De Wilde, “Use and validation of supervised machine learning approach for detection of GNSS signal spoofing,” in *2019 International Conference on Localization and GNSS (ICL-GNSS)*, Jun. 2019, pp. 1–6.
- [141] F. Gallardo and A. P. Yuste, “SCER spoofing attacks on the Galileo open service and machine learning techniques for end-user protection,” *IEEE Access*, vol. 8, pp. 85 515–85 532, 2020.
- [142] Q. Wu, C. Feres, D. Kuzmenko, D. Zhi, Z. Yu, X. Liu, and X. ‘Leo’ Liu, “Deep learning based RF fingerprinting for device identification and wireless security,” *Electron. Lett.*, vol. 54, no. 24, pp. 1405–1407, Nov. 2018.
- [143] N. Soltanieh, Y. Norouzi, Y. Yang, and N. C. Karmakar, “A review of radio frequency fingerprinting techniques,” *IEEE J. Radio Freq. Identification*, vol. 4, no. 3, pp. 222–233, Jan. 2020.
- [144] T. Jian, B. C. Rendon, E. Ojuba, N. Soltani, Z. Wang, K. Sankhe, A. Gritsenko, J. Dy, K. Chowdhury, and S. Ioannidis, “Deep learning for rf fingerprinting: A massive experimental study,” *IEEE Internet Things Mag.*, vol. 3, no. 1, pp. 50–57, Mar. 2020.
- [145] S. Qiu, K. Sava, and W. Guo, “Robust satellite antenna fingerprinting under degradation using recurrent neural network,” *Mod. Phys. Lett. B*, vol. 36, no. 12, p. 2250043, Apr. 2022.
- [146] H.-M. Chin, N. Jain, D. Zibar, U. L. Andersen, and T. Gehring, “Machine learning aided carrier recovery in continuous-variable quantum key distribution,” *npj Quantum Inf.*, vol. 7, no. 1, p. 20, Feb 2021.
- [147] Z.-A. Ren, Y.-P. Chen, J.-Y. Liu, H.-J. Ding, and Q. Wang, “Implementation of machine learning in quantum key distributions,” *IEEE Commun. Lett.*, vol. 25, no. 3, pp. 940–944, Mar. 2021.
- [148] J.-Y. Liu, H.-J. Ding, C.-M. Zhang, S.-P. Xie, and Q. Wang, “Practical phase-modulation stabilization in quantum key distribution via machine learning,” *Phys. Rev. Applied*, vol. 12, p. 014059, Jul. 2019.
- [149] W. Wang and H.-K. Lo, “Machine learning for optimal parameter prediction in quantum key distribution,” *Phys. Rev. A*, vol. 100, p. 062334, Dec. 2019.
- [150] Y. Yao, H. Zhou, and M. Erol-Kantarci, “Joint sensing and communications for deep reinforcement learning-based beam management in 6g,” in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 5019–5024.
- [151] L. Mou and X. X. Zhu, “Spatiotemporal scene interpretation of space videos via deep neural network and tracklet analysis,” in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 1823–1826.
- [152] B. Zhang, Y. Wu, B. Zhao, J. Chanussot, D. Hong, J. Yao, and L. Gao, “Progress and challenges in intelligent remote sensing satellite systems,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1814–1822, 2022.
- [153] Y. Cheng and Q. L. Huang, “Study on the data processing technique in high resolution remote sensing satellite,” in *Advances in Manufacturing Technology*, ser. Applied Mechanics and Materials, vol. 220. Trans Tech Publications Ltd, 11 2012, pp. 2079–2082.
- [154] J. M. Mateos-Ramos, J. Song, Y. Wu, C. Häger, M. F. Keskin, V. Jainanarayana, and H. Wymeersch, “End-to-end learning for integrated sensing and communication,” in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 1942–1947.
- [155] A. N. Rahman, V. Tripathiy, A. D. Gupta, B. Paul, M. T. Kurian, and V. P. Vijayan, “Satellite image fusion for obtaining high resolution images using deep neural network,” in *2022 International Conference on Intelligent Innovations in Engineering and Technology (ICIET)*, 2022, pp. 301–306.
- [156] J. Fang, J. Xiao, X. Wang, D. Chen, and R. Hu, “Arbitrary scale super resolution network for satellite imagery,” *China Communications*, vol. 19, no. 8, pp. 234–246, 2022.
- [157] D. Zhang, J. Shao, X. Li, and H. T. Shen, “Remote sensing image super-resolution via mixed high-order attention network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 6, pp. 5183–5196, 2021.
- [158] H. Zhang, P. Wang, and Z. Jiang, “Nonpairwise-trained cycle convolutional neural network for single remote sensing image super-resolution,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4250–4261, 2021.
- [159] J. Sublime and E. Kalinicheva, “Automatic post-disaster damage mapping using deep-learning techniques for change detection: Case study of the tohoku tsunami,” *Remote Sensing*, vol. 11, p. 1123, 05 2019.
- [160] S. Lei, Z. Shi, and Z. Zou, “Coupled adversarial training for remote sensing image super-resolution,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3633–3643, 2020.
- [161] H. Azodi, U. Siart, and T. F. Eibert, “A fast 3-d deterministic ray tracing coverage simulator including creeping rays based on geometry

- voxelization technique," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 1, pp. 210–220, 2014.
- [162] Z. Lou and J.-M. Jin, "Modeling and simulation of broad-band antennas using the time-domain finite element method," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 12, pp. 4099–4110, 2005.
- [163] P. A. Tirkas and C. A. Balanis, "Finite-difference time-domain method for antenna radiation," *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 3, pp. 334–340, 1992.
- [164] J. Johnson and Y. Rahmat-Samii, "Genetic algorithms and method of moments (ga/mom) for the design of integrated antennas," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 10, pp. 1606–1614, 1999.
- [165] R. R. Alavi, "Near-field to far-field transformation and fault detection using adaptive sampling and machine learning in source reconstruction method," 2019.
- [166] B.-K. Yeo and Y. Lu, "Array failure correction with a genetic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 5, pp. 823–828, 1999.
- [167] D. R. Prado, J. A. Lopez-Fernandez, M. Arrebola, and G. Goussetis, "Support vector regression to accelerate design and crosspolar optimization of shaped-beam reflectarray antennas for space applications," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 3, pp. 1659–1668, 2018.
- [168] H. M. E. Misilmani and T. Naous, "Machine learning in antenna design: An overview on machine learning concept and algorithms," in *2019 International Conference on High Performance Computing & Simulation (HPCS)*, 2019, pp. 600–607.
- [169] C. A. Balanis, *Antenna theory: analysis and design*. John Wiley & sons, 2015.
- [170] R. L. Haupt, *Antenna Arrays: A computational approach*. Wiley-IEEE Press, 2010.
- [171] Vázquez-Peralvo, J. Querol, F. Ortiz, J. L. Rios, E. Lagunas, V. Monzon Baeza, G. Fontanesi, L. M. Garces-Socorras, J. C. Merlano Duncan, and S. Chatzinotas, "Flexible beamforming for direct radiating arrays in satellite communications," *submitted to IEEE Access*, 2022.
- [172] K.-B. Yu, "Adaptive beamforming for satellite communication with selective earth coverage and jammer nulling capability," *IEEE Transactions on Signal Processing*, vol. 44, no. 12, pp. 3162–3166, 1996.
- [173] G. Kautz, "Phase-only shaped beam synthesis via technique of approximated beam addition," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 5, pp. 887–894, 1999.
- [174] H. Schippers, J. Verpoorte, P. Jorna, A. Hulzinga, A. Meijerink, C. Roeloffzen, R. Heideman, A. Leinse, and M. Wintels, "Conformal phased array with beam forming for airborne satellite communication," in *2008 International ITG Workshop on Smart Antennas*, 2008, pp. 343–350.
- [175] K. Sherman, "Phased array shaped multi-beam optimization for leo satellite communications using a genetic algorithm," in *Proceedings 2000 IEEE International Conference on Phased Array Systems and Technology (Cat. No.00TH8510)*, 2000, pp. 501–504.
- [176] T.-B. Chen, Y.-L. Dong, Y.-C. Jiao, and F. Zhang, "Synthesis of circular antenna array using crossed particle swarm optimization algorithm," *Journal of Electromagnetic Waves and Applications*, vol. 20, no. 13, pp. 1785–1795, 2006.
- [177] D. Boeringer and D. Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 3, pp. 771–779, 2004.
- [178] R. Webster, "A generalized hamming window," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 3, pp. 269–270, 1978.
- [179] T. Taylor, "Design of line-source antennas for narrow beamwidth and low side lobes," *Transactions of the IRE Professional Group on Antennas and Propagation*, vol. 3, no. 1, pp. 16–28, 1955.
- [180] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering — part i," *The Bell System Technical Journal*, vol. 37, no. 1, pp. 185–282, 1958.
- [181] C. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beam width and side-lobe level," *Proceedings of the IRE*, vol. 34, no. 6, pp. 335–348, 1946.
- [182] J. Anzalchi, A. Couchman, P. Gabellini, G. Gallinaro, L. D'Agristina, N. Alagha, and P. Angeletti, "Beam hopping in multi-beam broadband satellite systems: System simulation and performance comparison with non-hopped systems," in *2010 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop*, 2010, pp. 248–255.
- [183] Q. Zhao, Y. Hu, Z. Pang, and D. Ren, "Beam hopping for leo satellite: challenges and opportunities," in *2022 International Conference on Culture-Oriented Science and Technology (CoST)*, 2022, pp. 319–324.
- [184] P. Angeletti, D. F. Prim, and R. Rinaldo, *Beam Hopping in Multi-Beam Broadband Satellite Systems: System Performance and Payload Architecture Analysis*. [Online]. Available: <https://arc.aiaa.org/doi/abs/10.2514/6.2006-5376>
- [185] J. Lei and M. A. Vazquez-Castro, "Multibeam satellite frequency/time duality study and capacity optimization," *Journal of Communications and Networks*, vol. 13, no. 5, pp. 472–480, 2011.
- [186] R. Alegre-Godoy, N. Alagha, and M. A. Vázquez-Castro, "Offered capacity optimization mechanisms for multi-beam satellite systems," in *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 3180–3184.
- [187] L. Chen, V. N. Ha, E. Lagunas, L. Wu, S. Chatzinotas, and B. Ottersten, "The next generation of beam hopping satellite systems: Dynamic beam illumination with selective precoding," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.
- [188] V. N. Ha, T. T. Nguyen, E. Lagunas, J. C. M. Duncan, and S. Chatzinotas, "Geo payload power minimization: Joint precoding and beam hopping design," *arXiv preprint arXiv:2208.10474*, 2022.
- [189] J. Tang, D. Bian, G. Li, J. Hu, and J. Cheng, "Optimization method of dynamic beam position for leo beam-hopping satellite communication systems," *IEEE Access*, vol. 9, pp. 57 578–57 588, 2021.
- [190] Z. Lin, Z. Ni, L. Kuang, C. Jiang, and Z. Huang, "Multi-satellite beam hopping based on load balancing and interference avoidance for ngso satellite communication systems," *IEEE Transactions on Communications*, pp. 1–1, 2022.
- [191] M. Á. Vázquez, P. Henarejos, I. Pappalardo, E. Grechi, J. Fort, J. C. Gil, and R. M. Lancellotti, "Machine Learning for Satellite Communications Operations," *IEEE Commun. Mag.*, vol. 59, no. 2, pp. 22–27, 2021.
- [192] E. Godino, L. Escolar, and A. P. Honold, *Flexible Payload Operations of Satellite Communication Systems*. [Online]. Available: <https://arc.aiaa.org/doi/abs/10.2514/6.2018-2653>
- [193] F. d. G. Ortiz Gómez, R. Martínez Rodríguez-Orsorio, M. A. Salas Natera, S. Landeros Ayala, D. Tarchi, and A. Vanelli Coralli, "On the use of neural networks for flexible payload management in vhts systems," 2019.
- [194] F. G. Ortiz-Gómez, D. Tarchi, R. Martinez, A. Vanelli-Coralli, M. A. Salas-Natera, and S. Landeros-Ayala, "Supervised machine learning for power and bandwidth management in very high throughput satellite systems," *International Journal of Satellite Communications and Networking*, vol. 40, no. 6, pp. 392–407, 2022.
- [195] F. G. Ortiz-Gomez, D. Tarchi, R. M. Rodriguez-Orsorio, A. Vanelli-Coralli, M. A. Salas-Natera, and S. Landeros-Ayala, "Supervised machine learning for power and bandwidth management in vhts systems," in *2020 10th Advanced Satellite Multimedia Systems Conference and the 16th Signal Processing for Space Communications Workshop (ASMS/SPSC)*. IEEE, 2020, pp. 1–7.
- [196] K. An, T. Liang, X. Yan, Y. Li, and X. Qiao, "Power allocation in land mobile satellite systems: An energy-efficient perspective," *IEEE Communications Letters*, vol. 22, no. 7, pp. 1374–1377, 2018.
- [197] T. Qi and Y. Wang, "Energy-efficient power allocation over multibeam satellite downlinks with imperfect csi," in *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*, 2015, pp. 1–5.
- [198] C. N. Efrem and A. D. Panagopoulos, "Dynamic energy-efficient power allocation in multibeam satellite systems," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 228–231, 2020.
- [199] J. Choi and V. Chan, "Optimum power and beam allocation based on traffic demands and channel conditions over satellite downlinks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2983–2993, 2005.
- [200] F. R. Durand and T. Abrão, "Power allocation in multibeam satellites based on particle swarm optimization," *AEU - International Journal of Electronics and Communications*, vol. 78, pp. 124–133, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1434841116303739>
- [201] A. I. Aravanis, B. Shankar M. R., P.-D. Arapoglou, G. Danoy, P. G. Cottis, and B. Ottersten, "Power allocation in multibeam satellite systems: A two-stage multi-objective optimization," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3171–3182, 2015.
- [202] S. Liu, Y. Fan, Y. Hu, D. Wang, L. Liu, and L. Gao, "Ag-dpa: Assignment game-based dynamic power allocation in multibeam satellite systems," *International Journal of Satellite Communications*

- and Networking, vol. 38, no. 1, pp. 74–83, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1310>
- [203] J. Lei and M. A. Vázquez-Castro, “Joint power and carrier allocation for the multibeam satellite downlink with individual sinr constraints,” in *2010 IEEE International Conference on Communications*, 2010, pp. 1–5.
 - [204] T. S. Abdu, S. Kisseleff, E. Lagunas, and S. Chatzinotas, “A low-complexity resource optimization technique for high throughput satellite,” in *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, 2021, pp. 1–5.
 - [205] —, “Power and bandwidth minimization for demand-aware geo satellite systems,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.
 - [206] —, “Flexible resource optimization for geo multibeam satellite communication system,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7888–7902, 2021.
 - [207] T. Ramírez, C. Mosquera, and N. Alagha, “Flexible user mapping for radio resource assignment in advanced satellite payloads,” *IEEE Transactions on Broadcasting*, vol. 68, no. 3, pp. 723–739, 2022.
 - [208] G. Cocco, T. de Cola, M. Angelone, Z. Katona, and S. Erl, “Radio resource management optimization of flexible satellite payloads for dvb-s2 systems,” *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 266–280, 2018.
 - [209] A. Paris, I. Del Portillo, B. Cameron, and E. Crawley, “A genetic algorithm for joint power and bandwidth allocation in multibeam satellite systems,” in *2019 IEEE Aerospace Conference*, 2019, pp. 1–15.
 - [210] S. Kisseleff, E. Lagunas, T. S. Abdu, S. Chatzinotas, and B. Ottersten, “Radio resource management techniques for multibeam satellite systems,” *IEEE Communications Letters*, vol. 25, no. 8, pp. 2448–2452, 2021.
 - [211] A. Destounis and A. D. Panagopoulos, “Dynamic power allocation for broadband multi-beam satellite communication networks,” *IEEE Commun. Lett.*, vol. 15, no. 4, pp. 380–382, 2011.
 - [212] H. Fenech, S. Amos, A. Tomatis, and V. Soumolphakdy, “High throughput satellite systems: An analytical approach,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 192–202, 2015.
 - [213] Y. Kawamoto, T. Kamei, M. Takahashi, N. Kato, A. Miura, and M. Toyoshima, “Flexible resource allocation with inter-beam interference in satellite communication systems with a digital channelizer,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 2934–2945, 2020.
 - [214] U. Park, H. W. Kim, D. S. Oh, and B. J. Ku, “Flexible bandwidth allocation scheme based on traffic demands and channel conditions for multi-beam satellite systems,” in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, 2012, pp. 1–5.
 - [215] H. Wang, A. Liu, X. Pan, and L. Jia, “Optimal bandwidth allocation for multi-spot-beam satellite communication systems,” in *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, 2013, pp. 2794–2798.
 - [216] N. Pachler, J. J. G. Luis, M. Guerster, E. Crawley, and B. Cameron, “Allocating power and bandwidth in multibeam satellite systems using particle swarm optimization,” in *2020 IEEE Aerospace Conference*, 2020, pp. 1–11.
 - [217] B. S. Mysore, E. Lagunas, S. Chatzinotas, and B. Ottersten, “Precoding for satellite communications: Why, how and what next?” *IEEE Communications Letters*, vol. 25, no. 8, pp. 2453–2457, 2021.
 - [218] A. I. Pérez-Neira, M. A. Vázquez, M. B. Shankar, S. Maleki, and S. Chatzinotas, “Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios,” *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 112–131, 2019.
 - [219] Z. Li, Z. Xie, and X. Liang, “Dynamic channel reservation strategy based on dqn algorithm for multi-service leo satellite communication system,” *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 770–774, 2021.
 - [220] D. Deng, C. Wang, M. Pang, and W. Wang, “Dynamic resource allocation with deep reinforcement learning in multibeam satellite communication,” *IEEE Wireless Communications Letters*, pp. 1–1, 2022.
 - [221] X. Liao, X. Hu, Z. Liu, S. Ma, L. Xu, X. Li, W. Wang, and F. M. Ghannouchi, “Distributed intelligence: A verification for multi-agent drl-based multibeam satellite resource allocation,” *IEEE Communications Letters*, vol. 24, no. 12, pp. 2785–2789, 2020.
 - [222] C. Han, A. Liu, L. Huo, H. Wang, and X. Liang, “A prediction-based resource matching scheme for rentable leo satellite communication network,” *IEEE Communications Letters*, vol. 24, no. 2, pp. 414–417, 2020.
 - [223] X. Ding, L. Feng, Y. Zou, and G. Zhang, “Deep learning aided spectrum prediction for satellite communication systems,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 16314–16319, 2020.
 - [224] P. V. R. Ferreira, R. Paffenroth, A. M. Wyglinski, T. M. Hackett, S. G. Bilén, R. C. Reinhart, and D. J. Mortensen, “Multiobjective reinforcement learning for cognitive satellite communications using deep neural network ensembles,” *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 5, pp. 1030–1041, 2018.
 - [225] H. Al-Hraishawi, E. Lagunas, and S. Chatzinotas, “Traffic simulator for multibeam satellite communication systems,” in *10th Advanced Satellite Multimedia Syst. Conf. and the 16th Signal Process. Space Commun. Workshop (ASMS/SPSC)*, Sep. 2020, pp. 1–8.
 - [226] F. G. Ortiz-Gomez, M. A. Salas-Natera, R. Martínez, and S. Landeros-Ayala, “Optimization in vhts satellite system design with irregular beam coverage for non-uniform traffic distribution,” *Remote Sensing*, vol. 13, no. 13, 2021. [Online]. Available: <https://www.mdpi.com/2072-4292/13/13/2642>
 - [227] P. J. Honniah, N. Maturo, S. Chatzinotas, S. Kisseleff, and J. Krause, “Demand-based adaptive multi-beam pattern and footprint planning for high throughput geo satellite systems,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1526–1540, 2021.
 - [228] K. Rao, M. Cuchanski, and M. Tang, “Multiple beam antenna concepts for satellite communications,” in *Symposium on Antenna Technology and Applied Electromagnetics [ANTEM 1994]*. IEEE, 1994, pp. 289–292.
 - [229] S. K. Rao, “Advanced antenna technologies for satellite communications payloads,” *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 4, pp. 1205–1217, 2015.
 - [230] M. Schneider, C. Hartwanger, and H. Wolf, “Antennas for multiple spot beam satellites,” *CEAS Space Journal*, vol. 2, no. 1, pp. 59–66, 2011.
 - [231] K. Y. Jo, *Satellite communications network design and analysis*. Artech house, 2011.
 - [232] Y. Su, Y. Liu, Y. Zhou, J. Yuan, H. Cao, and J. Shi, “Broadband leo satellite communications: Architectures and key technologies,” *IEEE Wireless Communications*, vol. 26, no. 2, pp. 55–61, 2019.
 - [233] M. Takahashi, Y. Kawamoto, N. Kato, A. Miura, and M. Toyoshima, “Adaptive power resource allocation with multi-beam directivity control in high-throughput satellite communication systems,” *IEEE Wireless Communications Letters*, vol. 8, no. 4, pp. 1248–1251, 2019.
 - [234] S. Tani, K. Motoyoshi, H. Sano, A. Okamura, H. Nishiyama, and N. Kato, “An adaptive beam control technique for q band satellite to maximize diversity gain and mitigate interference to terrestrial networks,” *IEEE Transactions on Emerging Topics in Computing*, vol. 7, no. 1, pp. 115–122, 2016.
 - [235] S. Kisseleff, B. Shankar, D. Spano, and J.-D. Gayraud, “A new optimization tool for mega-constellation design and its application to trunking systems [international communications satellite systems conference],” 2019.
 - [236] M. Takahashi, Y. Kawamoto, N. Kato, A. Miura, and M. Toyoshima, “Adaptive multi-beam arrangement for improving throughput in an hts communication system,” in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.
 - [237] A. Pérez-Neira, J. M. Veciana, M. Á. Vázquez, and E. Lagunas, “Distributed power control with received power constraints for time-area-spectrum licenses,” *Signal Processing*, vol. 120, pp. 141–155, 2016.
 - [238] E. Lagunas, S. Maleki, S. Chatzinotas, M. Soltanalian, A. I. Pérez-Neira, and B. Ottersten, “Power and rate allocation in cognitive satellite uplink networks,” in *2016 IEEE International Conference on Communications (ICC)*. IEEE, 2016, pp. 1–6.
 - [239] D. Chen, J. Zhang, and R. Zhao, “Adaptive modulation and coding in satellite-integrated 5g communication system,” in *2021 IEEE 21st International Conference on Communication Technology (ICCT)*, 2021, pp. 1402–1407.
 - [240] E. İnceöz, R. Tutgun, and A. M. Y. Turgut, “Fpga based transmitter design using adaptive coding and modulation schemes for low earth orbit satellite communications,” in *2020 IEEE 5th International Symposium on Telecommunication Technologies (ISTT)*, 2020, pp. 39–44.
 - [241] J. Bas and A. A. Dowhuszko, “Time-packing as enabler of optical feeder link adaptation in high throughput satellite systems,” in *2020 IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 186–192.
 - [242] S. Andrenacci, D. Spano, D. Christopoulos, S. Chatzinotas, J. Krause, and B. Ottersten, “Optimized link adaptation for dvb-s2x precoded waveforms based on snr estimation,” in *2016 50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 502–506.

- [243] J. Zou, H. Xiong, D. Wang, and C. W. Chen, "Optimal power allocation for hybrid overlay/underlay spectrum sharing in multiband cognitive radio networks," *IEEE Trans. Veh. Commun.*, vol. 62, no. 4, pp. 1827–1837, 2013.
- [244] S. K. Sharma, S. Chatzinotas, and B. Ottersten, "In-line Interference Mitigation Techniques for Spectral Coexistence of GEO and NGE Satellites," *Int. J. Satell. Commun. Network*, vol. 34, no. 1, pp. 11–39, 2016.
- [245] K. Kim, I. A. Akbar, K. K. Bae, J.-S. Um, C. M. Spooner, and J. H. Reed, "Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio," in *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2007, pp. 212–215.
- [246] I. Jabandžić, F. Firyaguna, S. Giannoulis, A. Shahid, A. Mukhopadhyay, M. Ruffini, and I. Moerman, "The codysun approach: A novel distributed paradigm for dynamic spectrum sharing in satellite communications," *Sensors*, vol. 21, no. 23, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/23/8052>
- [247] F. Ling, *Synchronization in digital communication systems*. Cambridge University Press, 2017.
- [248] F. M. Gardner, *Phaselock techniques*. John Wiley & Sons, 2005.
- [249] U. Mengali, *Synchronization techniques for digital receivers*. Springer Science & Business Media, 1997.
- [250] J. Tong, R. Song, Y. Liu, C. Wang, Q. Zhou, and W. Wang, "Enhanced synchronization of 5g integrated satellite systems in multipath channels," in *2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, 2020, pp. 617–621.
- [251] Pierluigi Paganini. (2015, Aug.) Hacking the Iridium network could be very easy. Security Affairs. [Online]. Available: <https://securityaffairs.co/wordpress/39510/hacking/hacking-iridium-network.html>
- [252] Z. Wu, Y. Zhang, Y. Yang, C. Liang, and R. Liu, "Spoofing and anti-spoofing technologies of global navigation satellite system: A survey," *IEEE Access*, vol. 8, pp. 165 444–165 496, 2020.
- [253] P. Yue, J. An, J. Zhang, G. Pan, S. Wang, P. Xiao, and L. Hanzo, "On the security of LEO satellite communication systems: Vulnerabilities, countermeasures, and future trends," *arXiv preprint arXiv:2201.03063*, Jan. 2022.
- [254] W. Akoto. (2020, Feb.) Hackers could shut down satellites—or turn them into weapons. The Conversation. [Online]. Available: <https://theconversation.com/hackers-could-shut-down-satellites-or-turn-them-into-weapons-130932>
- [255] N. Rodríguez-Barroso, D. Jiménez-López, M. V. Luzón, F. Herrera, and E. Martínez-Cámara, "Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges," *Inf. Fusion*, vol. 90, pp. 148–173, Feb. 2023.
- [256] J. Camp, "The world in 2010: many new entrants," *info.*, vol. 2, no. 2, pp. 167–185, Jan. 2000.
- [257] J. Ma and Z. Kan, "Role of mobile IPv6 for mobile networks and its remaining issues," in *Proc. SPIE 4586*, H. Wu and J. Vaario, Eds., International Society for Optics and Photonics. SPIE, Oct. 2001, pp. 119 – 133.
- [258] H. Cruickshank, "A security system for satellite networks," in *Proc. Fifth Satellite Systems for Mobile Commun. and Navigation*. IET, May 1996, pp. 187–190.
- [259] T.-H. Chen, W.-B. Lee, and H.-B. Chen, "A self-verification authentication mechanism for mobile satellite communication systems," *Comput. Electr. Eng.*, vol. 35, no. 1, pp. 41–48, 2009.
- [260] M.-S. Hwang, C.-C. Yang, and C.-Y. Shiu, "An authentication scheme for mobile satellite communication systems," *SIGOPS Oper. Syst. Rev.*, vol. 37, no. 4, p. 42–47, Oct. 2003.
- [261] X. Deng, J. Shao, L. Chang, and J. Liang, "A blockchain-based authentication protocol using cryptocurrency technology in LEO satellite networks," *Electronics*, vol. 10, no. 24, 2021.
- [262] E. Jedermann, M. Strohmeier, M. Schäfer, J. Schmitt, and V. Lenders, "Orbit-based authentication using tdoa signatures in satellite networks," in *Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2021, p. 175–180.
- [263] Y. Liu, L. Ni, and M. Peng, "A secure and efficient authentication protocol for satellite-terrestrial networks," *IEEE Internet Things J.*, 2022, early access.
- [264] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5G and beyond wireless networks," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 55–61, Oct. 2019.
- [265] W. M. Mahmoud, D. Elfiky, S. M. Robaa, M. S. Elnawawy, and S. M. Yousef, "Effect of atomic Oxygen on LEO cubesat," *Int. J. Aeronaut. Space Sci.*, vol. 22, no. 3, pp. 726–733, Jun. 2021.
- [266] C. Jiang, X. Wang, J. Wang, H.-H. Chen, and Y. Ren, "Security in space information networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 82–88, Aug. 2015.
- [267] S. Lohani and R. Joshi, "Satellite network security," in *2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3)*, Feb. 2020, pp. 1–5.
- [268] J. Ma, R. Shrestha, J. Adelberg, C.-Y. Yeh, Z. Hossain, E. Knightly, J. M. Jornet, and D. M. Mittleman, "Security and eavesdropping in terahertz wireless links," *Nature*, vol. 563, no. 7729, pp. 89–93, Nov. 2018.
- [269] B. Li, Z. Fei, C. Zhou, and Y. Zhang, "Physical-layer security in space information networks: A survey," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 33–52, Sep. 2020.
- [270] Y. Huang, J. Wang, C. Zhong, T. Q. Duong, and G. K. Karagiannidis, "Secure transmission in cooperative relaying networks with multiple antennas," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6843–6856, Jul. 2016.
- [271] N. Zhao, F. R. Yu, M. Li, Q. Yan, and V. C. M. Leung, "Physical layer security issues in interference- alignment-based wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 162–168, Aug. 2016.
- [272] M. G. Schraml, A. Knopp, and K.-U. Storek, "Multi-user MIMO satellite communications with secrecy constraints," in *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)*, Nov. 2019, pp. 17–22.
- [273] M. G. Schraml, R. T. Schwarz, and A. Knopp, "Multiuser MIMO concept for physical layer security in multibeam satellite systems," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1670–1680, Nov. 2021.
- [274] C. Gidney and M. Ekerå, "How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits," *Quantum*, vol. 5, p. 433, Apr. 2021.
- [275] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, Bangalore, India. New York: IEEE, Dec. 1984, pp. 175–179.
- [276] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Phys. Rev. Lett.*, vol. 67, pp. 661–663, Aug. 1991.
- [277] H.-K. Lo, M. Curty, and B. Qi, "Measurement-device-independent quantum key distribution," *Phys. Rev. Lett.*, vol. 108, p. 130503, Mar. 2012.
- [278] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, Switzerland, 2005.
- [279] M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo, "Finite-key analysis for measurement-device-independent quantum key distribution," *Nat. Commun.*, vol. 5, Apr. 2014.
- [280] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, "Tight finite-key analysis for quantum cryptography," *Nat. Commun.*, vol. 3, p. 634, Jan. 2012.
- [281] R. Renner and R. Wolf, "Quantum advantage in cryptography," *arXiv preprint arxiv:2206.04078*, 2022.
- [282] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li, F.-Z. Li, X.-W. Chen, L.-H. Sun, J.-J. Jia, J.-C. Wu, X.-J. Jiang, J.-F. Wang, Y.-M. Huang, Q. Wang, Y.-L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y.-A. Chen, N.-L. Liu, X.-B. Wang, Z.-C. Zhu, C.-Y. Lu, R. Shu, C.-Z. Peng, J.-Y. Wang, and J.-W. Pan, "Satellite-to-ground quantum key distribution," *Nature*, vol. 549, no. 7670, pp. 43–47, Sep. 2017.
- [283] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, Y. Li, Q. Shen, Y. Cao, F.-Z. Li, J.-F. Wang, Y.-M. Huang, L. Deng, T. Xi, L. Ma, T. Hu, L. Li, N.-L. Liu, F. Koidl, P. Wang, Y.-A. Chen, X.-B. Wang, M. Steindorfer, G. Kirchner, C.-Y. Lu, R. Shu, R. Ursin, T. Scheidl, C.-Z. Peng, J.-Y. Wang, A. Zeilinger, and J.-W. Pan, "Satellite-relayed intercontinental quantum network," *Phys. Rev. Lett.*, vol. 120, p. 030501, Jan. 2018.
- [284] S. Pirandola, "Satellite quantum communications: Fundamental bounds and practical security," *Physical Review Research*, vol. 3, no. 2, p. 023130, 2021.
- [285] S. Wang, W. Chen, Z.-Q. Yin, D.-Y. He, C. Hui, P.-L. Hao, G.-J. Fan-Yuan, C. Wang, L.-J. Zhang, J. Kuang, S.-F. Liu, Z. Zhou, Y.-G. Wang, G.-C. Guo, and Z.-F. Han, "Practical gigahertz quantum key distribution robust against channel disturbance," *Opt. Lett.*, vol. 43, no. 9, pp. 2030–2033, May 2018.
- [286] M. Á. Vázquez, A. Perez-Neira, D. Christopoulos, S. Chatzinotas, B. Ottersten, P.-D. Arapoglou, A. Ginesi, and G. Taricco, "Precoding in multibeam satellite communications: Present and future challenges," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 88–95, 2016.
- [287] L. M. Marrero, J. Duncan, J. Querol, N. Maturro, J. Krivochiza, S. Chatzinotas, and B. Ottersten, "Differential phase compensation

- in over-the-air precoding test-bed for a multi-beam satellite,” in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 1325–1330.
- [288] J. Krivochiza, J. C. M. Duncan, J. Querol, N. Maturo, L. M. Marrero, S. Andrenacci, J. Krause, and S. Chatzinotas, “End-to-end precoding validation over a live geo satellite forward link,” *IEEE Access*, pp. 1–1, 2021.
- [289] R. Hunger, *Floating point operations in matrix-vector calculus*. Munich University of Technology, Inst. for Circuit Theory and Signal Processing, 2005.
- [290] J. Kim, H. Lee, S.-E. Hong, and S.-H. Park, “Deep learning methods for universal miso beamforming,” *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1894–1898, 2020.
- [291] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, “Learning to optimize: Training deep neural networks for interference management,” *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.
- [292] “MLSAT — ESA TIA.” [Online]. Available: <https://artes.esa.int/projects/mlsat>
- [293] L. Pellaco, M. Bengtsson, and J. Jaldén, “Matrix-inverse-free deep unfolding of the weighted mmse beamforming algorithm,” *IEEE Open Journal of the Communications Society*, vol. 3, pp. 65–81, 2022.
- [294] H. Fenech, A. Tomatis, S. Amos, J. S. Merino, and V. Soumpholphakdy, “An operator’s perspective on propagation,” in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, 2014, pp. 3349–3352.
- [295] A. Kyrgiazos, B. Evans, P. Thompson, and N. Jeannin, “Gateway diversity scheme for a future broadband satellite system,” in *2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC)*, 2012, pp. 363–370.
- [296] A. Panagopoulos, P.-D. Arapoglou, J. Kanellopoulos, and P. Cottis, “Long-term rain attenuation probability and site diversity gain prediction formulas,” *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 7, pp. 2307–2313, 2005.
- [297] M. Van de Kamp, “Short-term prediction of rain attenuation using two samples,” *Electronics Letters*, vol. 38, no. 23, p. 1, 2002.
- [298] L. De Montera, C. Mallet, L. Barthès, and P. Golé, “Short-term prediction of rain attenuation level and volatility in Earth-to-Satellite links at EHF band,” *Nonlinear Processes in Geophysics*, vol. 15, no. 4, pp. 631–643, 2008.
- [299] T. Maseng and P. Bakken, “A Stochastic Dynamic Model of Rain Attenuation,” *IEEE Transactions on Communications*, vol. 29, no. 5, pp. 660–669, 1981.
- [300] N. Jeannin, L. Castanet, I. Dahman, V. Pourret, and B. Pouponneau, “Smart gateways switching control algorithms based on tropospheric propagation forecasts,” *International Journal of Satellite Communications and Networking*, vol. 37, no. 1, pp. 43–55, 2019.
- [301] C. Huang, R. He, B. Ai, A. F. Molisch, B. K. Lau, K. Haneda, B. Liu, C.-X. Wang, M. Yang, C. Oestges, and Z. Zhong, “Artificial Intelligence Enabled Radio Propagation for Communications—Part I: Channel Characterization and Antenna-Channel Optimization,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 6, pp. 3939–3954, 2022.
- [302] M. A. Vázquez, P. Henarejos, and L. Blanco, “Deep Gateway Switching,” in *2022 27th Ka and Broadband Communications Conference (Ka) and the 39th International Communications Satellite Systems Conference (ICSSC)*, 2012.
- [303] J.-M. Shyu, *Cost Reduction of Satellite Systems: Lessons Learned from the ROCSAT-1 Program*. Dordrecht: Springer Netherlands, 2000, pp. 127–133. [Online]. Available: https://doi.org/10.1007/978-94-015-9395-3_17
- [304] V. González-Posadas, D. Segovia-Vargas, J.-L. Vázquez-Roy, E. Rajo-Iglesias, and C. Martín-Pascual, “Review and classification of active antennas,” *Applied Microwave and Wireless*, vol. 14, no. 3, pp. 76–89, 2002.
- [305] S. C. Cripps, *RF power amplifiers for wireless communications*. Artech house Norwood, MA, 2006, vol. 2.
- [306] E. McCune, “Chapter 3 - transmitter linearity and energy efficiency,” in *Linearization and Efficiency Enhancement Techniques for Silicon Power Amplifiers*, E. Kerhervé and D. Belot, Eds. Oxford: Academic Press, 2015, pp. 55–81. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780124186781000039>
- [307] T. Kobal, Y. Li, X. Wang, and A. Zhu, “Digital predistortion of rf power amplifiers with phase-gated recurrent neural networks,” *IEEE Transactions on Microwave Theory and Techniques*, 2022.
- [308] Y. Zhang, Y. Li, F. Liu, and A. Zhu, “Vector decomposition based time-delay neural network behavioral model for digital predistortion of rf power amplifiers,” *IEEE Access*, vol. 7, pp. 91 559–91 568, 2019.
- [309] X. Chen, Z. Lu, S. Zhang, S. Xu, and Y. Wang, “An inter-modulation oriented learning based digital pre-distortion technique via joint intermediate and radio frequency optimization,” *IEEE Transactions on Wireless Communications*, 2022.
- [310] P. Hailes, L. Xu, R. G. Maunder, B. M. Al-Hashimi, and L. Hanzo, “A survey of fpga-based ldpc decoders,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1098–1122, 2016.
- [311] M. Y. Zinchenko, A. M. Levadniy, and Y. A. Grebenko, “Ldpc decoder power consumption optimization,” in *2020 International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, 2020, pp. 1–5.
- [312] J. Kang, J. S. An, and B. Wang, “An efficient fec encoder core for vcm leo satellite-ground communications,” *IEEE Access*, vol. 8, pp. 125 692–125 701, 2020.
- [313] V. Pignoly, B. Le Gal, C. Jegou, B. Gadat, and L. Barthe, “Fair comparison of hardware and software ldpc decoder implementations for sdr space links,” in *2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2020, pp. 1–4.
- [314] T. Yoo and A. Goldsmith, “On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.
- [315] P. Viswanath, D. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.
- [316] ETSI, “Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications; part ii: S2-extensions (s2-x),” *Digital Video Broadcasting (DVB)*, vol. ETSI EN 302 307-2 V1.1.1, 2014-10. [Online]. Available: https://www.etsi.org/deliver/etsi_en/302300_302399/30230702/01.01.01_20/en_30230702v010101a.pdf
- [317] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, “Quality of Service and Max-Min Fair Transmit Beamforming to Multiple Cochannel Multicast Groups,” *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, 2008.
- [318] A. Guidotti and A. Vanelli-Coralli, “Geographical scheduling for multicast precoding in multi-beam satellite systems,” in *2018 9th Advanced Satellite Multimedia Systems Conference and the 15th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, 2018, pp. 1–8.
- [319] G. Taricco, “Linear precoding methods for multi-beam broadband satellite systems,” in *European Wireless 2014; 20th European Wireless Conference*, 2014, pp. 1–6.
- [320] Z. Sha, Z. Wang, S. Chen, and L. Hanzo, “Graph theory based beam scheduling for inter-cell interference avoidance in mmwave cellular networks,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3929–3942, 2020.
- [321] V. Joroughi, M. A. Vázquez, and A. I. Pérez-Neira, “Generalized multicast multibeam precoding for satellite communications,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 952–966, 2016.
- [322] M. A. Vázquez, M. R. B. Shankar, C. I. Kourogiorgas, P.-D. Arapoglou, V. Icolari, S. Chatzinotas, A. D. Panagopoulos, and A. I. Pérez-Neira, “Precoding, Scheduling, and Link Adaptation in Mobile Interactive Multibeam Satellite Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 5, pp. 971–980, 2018.
- [323] S. Dimitrov, S. Erl, B. Barth, S. Jaekel, A. Kyrgiazos, and B. G. Evans, “Radio resource management techniques for high throughput satellite communication systems,” in *2015 European Conference on Networks and Communications (EuCNC)*. IEEE, 2015, pp. 175–179.
- [324] S. Zhang, M. Jia, Y. Wei, and Q. Guo, “User scheduling for multicast transmission in high throughput satellite systems,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2020, no. 1, p. 133, Jun 2020. [Online]. Available: <https://doi.org/10.1186/s13638-020-01749-7>
- [325] D. Christopoulos, S. Chatzinotas, and B. Ottersten, “Multicast multi-group precoding and user scheduling for frame-based satellite communications,” *IEEE Trans. Wireless Commun.*, vol. 14, p. 9, September 2015.
- [326] P. J. Honniah, E. Lagunas, S. Chatzinotas, and J. Krause, “Interference-aware demand-based user scheduling in precoded high throughput satellite systems,” *IEEE Open Journal of Vehicular Technology*, vol. 3, pp. 120–137, 2022.
- [327] Y. D. Zhang and K. D. Pham, “Joint precoding and scheduling optimization in downlink multicell satellite communications,” in *2020*

- 54th Asilomar Conference on Signals, Systems, and Computers, 2020, pp. 480–484.
- [328] A. Bandi, B. S. Mysore R., S. Chatzinotas, and B. Ottersten, “Joint scheduling and precoding for frame-based multigroup multicasting in satellite communications,” in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.
- [329] A. Bandi, B. Shankar M. R., S. Chatzinotas, and B. Ottersten, “A joint solution for scheduling and precoding in multiuser MISO downlink channels,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 475–490, 2020.
- [330] Y. Chi, L. Liu, G. Song, C. Yuen, Y. L. Guan, and Y. Li, “Practical mimo-noma: Low complexity and capacity-approaching solution,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6251–6264, 2018.
- [331] L. Liu, C. Yuen, Y. L. Guan, and Y. Li, “Capacity-achieving iterative lmmse detection for mimo-noma systems,” in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–6.
- [332] L. Liu, Y. Chi, C. Yuen, Y. L. Guan, and Y. Li, “Capacity-achieving mimo-noma: Iterative lmmse detection,” *IEEE Transactions on Signal Processing*, vol. 67, no. 7, pp. 1758–1773, 2019.
- [333] J. Jiao, Y. Sun, S. Wu, Y. Wang, and Q. Zhang, “Network utility maximization resource allocation for noma in satellite-based internet of things,” *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3230–3242, 2020.
- [334] B. Makki, K. Chitti, A. Behravan, and M.-S. Alouini, “A survey of noma: Current status and open research challenges,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, 2020.
- [335] J. Zhao, X. Yue, S. Kang, and W. Tang, “Joint effects of imperfect csi and sic on noma based satellite-terrestrial systems,” *IEEE Access*, vol. 9, pp. 12 545–12 554, 2021.
- [336] S. Xie, B. Zhang, D. Guo, and W. Ma, “Outage performance of noma-based integrated satellite-terrestrial networks with imperfect csi,” *Electronics Letters*, vol. 55, no. 14, pp. 793–795, 2019. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/el.2018.7839>
- [337] Y. Mao, B. Clerckx, and V. O. Li, “Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming sdma and noma,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 133, May 2018. [Online]. Available: <https://doi.org/10.1186/s13638-018-1104-7>
- [338] —, “Energy efficiency of rate-splitting multiple access, and performance benefits over sdma and noma,” in *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, 2018, pp. 1–5.
- [339] B. Clerckx, Y. Mao, E. Jorswieck, J. Yuan, D. Love, E. Erkip, and D. Niyato, “A primer on rate-splitting multiple access: Tutorial, myths, and frequently asked questions,” 09 2022.
- [340] S. Gamal, M. Rihan, S. Hussin, A. Zaghloul, and A. A. Salem, “Multiple access in cognitive radio networks: From orthogonal and non-orthogonal to rate-splitting,” *IEEE Access*, vol. 9, pp. 95 569–95 584, 2021.
- [341] L. Yin and B. Clerckx, “Rate-splitting multiple access for multigroup multicast and multibeam satellite systems,” *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 976–990, 2020.
- [342] —, “Rate-splitting multiple access for multibeam satellite communications,” in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2020, pp. 1–6.
- [343] Z. W. Si, L. Yin, and B. Clerckx, “Rate-splitting multiple access for multigateway multibeam satellite systems with feeder link interference,” *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 2147–2162, 2022.
- [344] X. Li, Y. Fan, R. Yao, P. Wang, N. Qi, N. I. Miridakis, and T. A. Tsiftsis, “Rate-splitting multiple access-enabled security analysis in cognitive satellite terrestrial networks,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 11 756–11 771, 2022.
- [345] A. Vanelli-Coralli, A. Guidotti, T. Foggi, G. Colavolpe, and G. Montorsi, “5g and beyond 5g non-terrestrial networks: trends and research challenges,” in *2020 IEEE 3rd 5G World Forum (5GWF)*, 2020, pp. 163–169.
- [346] Y. Hokazono, H. Kohara, Y. Kishiyama, and T. Asai, “Extreme coverage extension in 6g: Cooperative non-terrestrial network architecture integrating terrestrial networks,” in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 138–143.
- [347] M. Werner, C. Delucchi, H.-J. Vogel, G. Maral, and J.-J. De Ridder, “Atm-based routing in leo/meo satellite networks with intersatellite links,” *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 1, pp. 69–82, 1997.
- [348] G. Stock, J. A. Fraire, and H. Hermanns, “Distributed on-demand routing for leo mega-constellations: A starlink case study,” in *2022 11th Advanced Satellite Multimedia Systems Conference and the 17th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, 2022, pp. 1–8.
- [349] N. Razmi, B. Matthiesen, A. Dekorsy, and P. Popovski, “On-board federated learning for dense leo constellations,” in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 4715–4720.
- [350] R. De Gaudenzi, O. Del Rio Herrero, G. Gallinaro, S. Cioni, and P. Arapoglou, “Random access schemes for satellite networks, from VSAT to M2M: a survey,” *International journal of satellite communications and networking*, vol. 36, no. 1, pp. 66–107, 2018.
- [351] O. Liberg, S. E. Löwenmark, S. Euler, B. Hofström, T. Khan, X. Lin, and J. Sedin, “Narrowband internet of things for non-terrestrial networks,” *IEEE Communications Standards Magazine*, vol. 4, no. 4, pp. 49–55, 2020.
- [352] C. Kissling and A. M. Dlr, “On the integration of random access and DAMA channels for the return link of satellite networks,” in *2013 IEEE International Conference on Communications (ICC)*, 2013, pp. 4282–4287.
- [353] R. De Gaudenzi and O. del Rio Herrero, “Advances in random access protocols for satellite networks,” in *2009 International Workshop on Satellite and Space Communications*, 2009, pp. 331–336.
- [354] M. M. Azari, S. Solanki, S. Chatzinotas, O. Kodheli, H. Sallouha, A. Colpaert, J. F. Mendoza Montoya, S. Pollin, A. Haqiqatnejad, A. Mostafaei, E. Lagunas, and B. Ottersten, “Evolution of non-terrestrial networks from 5G to 6G: A survey,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2633–2672, 2022.
- [355] B. Mao, F. Tang, Y. Kawamoto, and N. Kato, “AI models for green communications towards 6G,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 210–247, 2022.
- [356] D. de la Torre, F. G. Ortiz-Gomez, M. Salas-Natera, and R. Martínez, “Analysis of the traffic demand on very high throughput satellite for 5g,” in *Proceedings of the XXXV Simposio Nacional de la Unión Científica Internacional de Radio (URSI 2020)*, Malaga, Spain, 2020, pp. 2–4.
- [357] H. Al-Hraishawi, M. Minardi, H. Chougrani, O. Kodheli, J. F. M. Montoya, and S. Chatzinotas, “Multi-layer space information networks: Access design and softwarization,” *IEEE Access*, 2021.
- [358] H. Nishiyama, D. Kudoh, N. Kato, and N. Kadowaki, “Load balancing and qos provisioning based on congestion prediction for geo/leo hybrid satellite networks,” *Proceedings of the IEEE*, vol. 99, no. 11, pp. 1998–2007, 2011.
- [359] J. Wu, Y. Peng, M. Song, M. Cui, and L. Zhang, “Link congestion prediction using machine learning for software-defined-network data plane,” in *2019 International Conference on Computer, Information and Telecommunication Systems (CITS)*. IEEE, 2019, pp. 1–5.
- [360] Y. Li, R. Yu, C. Shahabi, and Y. Liu, “Diffusion convolutional recurrent neural network: Data-driven traffic forecasting,” in *International Conference on Learning Representations (ICLR '18)*, 2018.
- [361] P. Henarejos, M. A. Vazquez, and L. Blanco, “Traffic congestion prediction in satellite broadband communications,” in *Proceedings of the 39th International Communications Satellite Systems Conference (ICSSC)*, 2022.
- [362] H. Al-Hraishawi, S. Chatzinotas, and B. Ottersten, “Broadband non-geostationary satellite communication systems: Research challenges and key opportunities,” in *IEEE Int. Conf Commun. (ICC)*, Jun. 2021, pp. 1–6.
- [363] 3GPP, “System architecture for the 5G System (5GS),” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 23.501, 12 2021, version 17.3.0. [Online]. Available: <http://www.3gpp.org/DynaReport/23501.htm>
- [364] —, “Management and orchestration; Concepts, use cases and requirements,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 28.530, 12 2021, version 17.2.0. [Online]. Available: <http://www.3gpp.org/DynaReport/28530.htm>
- [365] B. Tiomela Jou, O. Vidal, P. Foulon, D. Pham-Minh, I. Keesmaat, M. Boutin, S. Watts, C. Politis, K. Liolis, R. Sperber, P. S. Khodashenas, and L. Goratti, “Integrated SaT5g general network architecture,” issue: D3.1. [Online]. Available: https://www.sat5g-project.eu/wp-content/uploads/2019/04/SaT5G-D3.1-Integrated-SaT5G-general-network-architecture_ADS_v1.00_Inter...pdf
- [366] VITAL consortium. H2020 VITAL project. [Online]. Available: <http://www.ict-vital.eu/>
- [367] B. T. Jou, O. Vidal, J. Cahill, F. Arnal, L.-M. Honssin, M. Boutin, and D. K. Chau, “Architecture options for satellite integration into 5g

- networks,” in *2018 European Conference on Networks and Communications (EuCNC)*, pp. 398–9, ISSN: 2575-4912.
- [368] T. Ahmed, A. Alleg, R. Ferrus, and R. Riggio, “On-demand network slicing using SDN/NFV-enabled satellite ground segment systems,” in *2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft)*, pp. 242–246.
- [369] Y. Drif, E. Chaput, E. Lavinal, P. Berthou, B. Tiomela Jou, O. Gremillet, and F. Arnal, “An extensible network slicing framework for satellite integration into 5g,” *International Journal of Satellite Communications and Networking*, vol. 39, no. 4, pp. 339–357, 2021.
- [370] Y. Drif, E. Lavinal, E. Chaput, P. Berthou, B. T. Jou, O. Grémillet, and F. Arnal, “Slice aware non terrestrial networks,” in *2021 IEEE 46th Conference on Local Computer Networks (LCN)*, 2021, pp. 24–31.
- [371] O. Kodheli, A. Guidotti, and A. Vanelli-Coralli, “Integration of satellites in 5g through leo constellations,” in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.
- [372] O. Kodheli, A. Astro, J. Querol, M. Gholamian, S. Kumar, N. Maturo, and S. Chatzinotas, “Random access procedure over non-terrestrial networks: From theory to practice,” vol. 9, pp. 109 130–109 143, conference Name: IEEE Access.
- [373] Mobile-edge computing - introductory technical white paper. [Online]. Available: https://portal.etsi.org/portals/0/tbpages/mec/docs/mobile-edge_computing_-_introductory_technical_white_paper_v1%2018-09-14.pdf
- [374] Z. Zhang, W. Zhang, and F.-H. Tseng, “Satellite mobile edge computing: Improving qos of high-speed satellite-terrestrial networks using edge computing techniques,” *IEEE Network*, vol. 33, no. 1, pp. 70–76, 2019.
- [375] Q. Li, S. Wang, X. Ma, Q. Sun, H. Wang, S. Cao, and F. Yang, “Service coverage for satellite edge computing,” *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 695–705, 2022.
- [376] P. Mach and Z. Becvar, “Mobile edge computing: A survey on architecture and computation offloading,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.
- [377] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, “A survey on mobile edge networks: Convergence of computing, caching and communications,” *IEEE Access*, vol. 5, pp. 6757–6779, 2017.
- [378] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, “Web caching and zipf-like distributions: evidence and implications,” in *IEEE INFOCOM ’99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, vol. 1, 1999, pp. 126–134 vol.1.
- [379] P. Blasco and D. Gündüz, “Learning-based optimization of cache content in a small cell base station,” in *2014 IEEE International Conference on Communications (ICC)*, 2014, pp. 1897–1903.
- [380] —, “Multi-armed bandit optimization of cache content in wireless information networks,” in *2014 IEEE International Symposium on Information Theory*, 2014, pp. 51–55.
- [381] E. Baştuğ, M. Bennis, and M. Debbah, “A transfer learning approach for cache-enabled wireless networks,” in *2015 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2015, pp. 161–166.
- [382] A. Sengupta, S. Amuru, R. Tandon, R. M. Buehrer, and T. C. Clancy, “Learning distributed caching strategies in small cell networks,” in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, 2014, pp. 917–921.
- [383] L. Marini, J. Li, and Y. Li, “Distributed caching based on decentralized learning automata,” in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 3807–3812.
- [384] W. Wang, R. Lan, J. Gu, A. Huang, H. Shan, and Z. Zhang, “Edge caching at base stations with device-to-device offloading,” *IEEE Access*, vol. 5, pp. 6399–6410, 2017.
- [385] B. N. Bharath, K. G. Nagananda, D. Gündüz, and H. V. Poor, “Caching with time-varying popularity profiles: A learning-theoretic perspective,” *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 3837–3847, 2018.
- [386] O. K. Shoukry and M. B. Fayek, “Evolutionary content pre-fetching in mobile networks,” in *2013 12th International Conference on Machine Learning and Applications*, vol. 1, 2013, pp. 386–391.
- [387] W. Chen and H. V. Poor, “Content pushing with request delay information,” *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1146–1161, 2017.
- [388] H. Linder, H. Clausen, and B. Collini-Nocker, “Satellite internet services using dvb/mpeg-2 and multicast web caching,” *IEEE Communications Magazine*, vol. 38, no. 6, pp. 156–161, 2000.
- [389] C. Brinton, E. Aryafar, S. Corda, S. Russo, R. Reinoso, and M. Chiang, “An intelligent satellite multicast and caching overlay for cdns to improve performance in video applications,” in *Proc. AIAA Int. Commun. Satellite Syst. Conf.*, 2013, pp. 386–391.
- [390] E. Bastug, M. Bennis, and M. Debbah, “Living on the edge: The role of proactive caching in 5g wireless networks,” *IEEE Communications Magazine*, vol. 52, no. 8, pp. 82–89, 2014.
- [391] E. Baştuğ, M. Bennis, E. Zeydan, M. A. Kader, I. A. Karatepe, A. S. Er, and M. Debbah, “Big data meets telcos: A proactive caching perspective,” *Journal of Communications and Networks*, vol. 17, no. 6, pp. 549–557, 2015.
- [392] B. Paul, A. R. Chiriyath, and D. W. Bliss, “Survey of rf communications and sensing convergence research,” *IEEE Access*, vol. 5, pp. 252–270, 2017.
- [393] A. Herschfeld, A. R. Chiriyath, S. Srinivas, and D. W. Bliss, “An introduction to spectral convergence: Challenges and paths to solutions,” in *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*, 2021, pp. 1–6.
- [394] Q. Yin, Q. Hu, H. Liu, F. Zhang, Y. Wang, Z. Lin, W. An, and Y. Guo, “Detecting and tracking small and dense moving objects in satellite videos: A benchmark,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–18, 2022.
- [395] C. Xiao, S. Wu, Y. Wang, M. Li, W. An, and Z. Chen, “Rsmot: Remote sensing multi-object tracking network with local motion prior for objects in satellite videos,” in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, pp. 1904–1907.
- [396] Y. Li, C. Bian, and H. Chen, “Object tracking in satellite videos: Correlation particle filter tracking method with motion estimation by kalman filter,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.
- [397] T. Sarma, C. Srinivas, and A. Venugopal, “Design and development of elevation independent tracking system for the remote sensing satellite data acquisition,” in *2007 International Conference on Signal Processing, Communications and Networking*, 2007, pp. 204–207.
- [398] O. Dubovik, G. Schuster, F. Xu, Y. Hu, H. Bösch, J. Landgraf, and Z. Li, “Grand challenges in satellite remote sensing,” *Frontiers in Remote Sensing*, vol. 2, p. 619818, 02 2021.
- [399] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, “Joint radar and communication design: Applications, state-of-the-art, and the road ahead,” *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [400] S. Hu, X. Yuan, W. Ni, and X. Wang, “Trajectory planning of cellular-connected UAV for communication-assisted radar sensing,” *IEEE Tran. Commun.*, 2022.
- [401] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, “Integrated sensing and communications: Towards dual-functional wireless networks for 6g and beyond,” *IEEE journal on selected areas in communications*, 2022.
- [402] L. You, X. Qiang, C. G. Tsinos, F. Liu, W. Wang, X. Gao, and B. Ottersten, “Beam squint-aware integrated sensing and communications for hybrid massive mimo leo satellite systems,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 2994–3009, 2022.
- [403] S. Aliaga, A. J. Alqaraghuli, and J. M. Jornet, “Joint terahertz communication and atmospheric sensing in low earth orbit satellite networks: Physical layer design,” in *2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2022, pp. 457–463.
- [404] H. Choi and S. Pack, “Cooperative downloading for leo satellite networks: A drl-based approach,” *Sensors*, vol. 22, no. 18, p. 6853, 2022.
- [405] L. Dalin, W. Haijiao, Y. Zhen, G. Yanfeng, and S. Shi, “An online distributed satellite cooperative observation scheduling algorithm based on multiagent deep reinforcement learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 11, pp. 1901–1905, 2020.
- [406] R. Deng, B. Di, H. Zhang, L. Kuang, and L. Song, “Ultra-dense leo satellite constellations: How many leo satellites do we need?” *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 4843–4857, 2021.
- [407] R. Radhakrishnan, W. W. Edmonson, F. Afghah, R. M. Rodriguez-Osorio, F. Pinto, and S. C. Burleigh, “Survey of inter-satellite communication for small satellite systems: Physical layer to network layer view,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2442–2473, 2016.
- [408] B. Palmintier, C. Kitts, P. Stang, and M. Swartwout, “A distributed computing architecture for small satellite and multi-spacecraft missions,” 2002.

- [409] SmartSat, "Machine learning onboard satellites," SmartSat Technical Report no. 010, SmartSat, Adelaide, Australia, Tech. Rep., 2021.
- [410] X. He, K. Zhao, and X. Chu, "Automl: A survey of the state-of-the-art," *Knowledge-Based Systems*, vol. 212, p. 106622, 2021.
- [411] G. Ghiasi, T.-Y. Lin, and Q. V. Le, "Nas-fpn: Learning scalable feature pyramid architecture for object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 7036–7045.
- [412] S. Xuan, S. Li, M. Han, X. Wan, and G.-S. Xia, "Object tracking in satellite videos by improved correlation filters with motion estimations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 2, pp. 1074–1086, 2020.
- [413] R. F. Ahmad, A. S. Malik, A. Qayyum, and N. Kamel, "Disaster monitoring in urban and remote areas using satellite stereo images: A depth estimation approach," in *2015 IEEE 11th International Colloquium on Signal Processing & Its Applications (CSPA)*, 2015, pp. 150–155.
- [414] 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, 2017. [Online]. Available: <https://www.mdpi.com/2072-4292/9/7/678>
- [415] I. del Portillo Barrios, B. Cameron, and E. Crawley, "Ground segment architectures for large leo constellations with feeder links in ehf-bands," 03 2018, pp. 1–14.
- [416] E. Juan, M. Lauridsen, J. Wigard, and P. E. Mogensen, "5g new radio mobility performance in leo-based non-terrestrial networks," in *2020 IEEE Globecom Workshops (GC Wkshps)*, 2020, pp. 1–6.
- [417] N. K. Lyras, C. N. Efrem, C. I. Kourogiorgas, A. D. Panagopoulos, and P.-D. Arapoglou, "Optimizing the ground network of optical meo satellite communication systems," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3968–3976, 2020.
- [418] N. Torkzaban, A. Gholami, J. S. Baras, and C. Papagianni, "Joint satellite gateway placement and routing for integrated satellite-terrestrial networks," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [419] A. Alvaro, L. Pascual, A. Abad, P. Pinto, A. Alvarez-Herrero, T. Belenguier, C. Miravet, P. Campo, L. F. Rodriguez, M. Reyes, J. Socas, and J. Bermejo, "Caramuel: The future of space quantum key distribution in geo," in *2022 IEEE International Conference on Space Optical Systems and Applications (ICSOS)*, 2022, pp. 57–65.
- [420] C. N. Efrem and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Communications Letters*, vol. 9, no. 7, pp. 1101–1104, 2020.
- [421] H. Dahrouj, R. Alghamdi, H. Alwazani, S. Bahanshal, A. A. Ahmad, A. Faisal, R. Shalabi, R. Alhadrami, A. Subasi, M. T. Al-Nory, O. Kittaneh, and J. S. Shamma, "An overview of machine learning-based techniques for solving optimization problems in communications and signal processing," *IEEE Access*, vol. 9, pp. 74 908–74 938, 2021.
- [422] K. Çelikbilek, Z. Saleem, R. Morales Ferre, J. Praks, and E. S. Lohan, "Survey on optimization methods for leo-satellite-based networks with applications in future autonomous transportation," *Sensors*, vol. 22, no. 4, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/4/1421>
- [423] A. Cornejo, S. Landeros-Ayala, J. M. Matias, and R. Martinez, "Applying learning methods to optimize the ground segment for hts systems," in *2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS)*, 2020, pp. 1–4.
- [424] F. Ortiz, E. Lagunas, W. Martins, T. Dinh, N. Skatchkovsky, O. Simeone, B. Rajendran, T. Navarro, and S. Chatzinotas, "Towards the application of neuromorphic computing to satellite communications," in *39th International Communications Satellite Systems Conference (ICSSC)*. IEEE, 2022, pp. 1–7.
- [425] F. Akopyan, J. Sawada, A. Cassidy, R. Alvarez-Icaza, J. Arthur, P. Merolla, N. Imam, Y. Nakamura, P. Datta, G.-J. Nam, B. Taba, M. Beakes, B. Brezko, J. B. Kuang, R. Manohar, P. Risk, B. Jackson, and D. S. Modha, "Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 10, pp. 1537–1557, 2015.
- [426] M. Davies, N. Srinivasa, T. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y. Weng, A. Wild, Y. Yang, and H. Wang, "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan 2018.
- [427] "Introducing Qualcomm Zeroth Processors: Brain-Inspired Computing — Qualcomm." [Online]. Available: <https://www.qualcomm.com/news/onq/2013/10/10/introducing-qualcomm-zeroth-processors-brain-inspired-computing>
- [428] S. Furber and P. Bogdan, Eds., *SpiNNaker - A spiking neural network architecture*, ser. NowOpen. Hanover, MD: now, Mar. 2020.
- [429] J. Schemmel, A. Grübl, S. Hartmann, A. Kononov, C. Mayr, K. Meier, S. Millner, J. Partzsch, S. Schiefer, S. Scholze, R. Schüffny, and M.-O. Schwartz, "Live demonstration: A scaled-down version of the brain-scales wafer-scale neuromorphic system," in *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2012, pp. 702–702.
- [430] B. V. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. R. Chandrasekaran, J.-M. Bussat, R. Alvarez-Icaza, J. V. Arthur, P. A. Merolla, and K. Boahen, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699–716, 2014.
- [431] A. Neckar, S. Fok, B. V. Benjamin, T. C. Stewart, N. N. Oza, A. R. Voelker, C. Eliasmith, R. Manohar, and K. Boahen, "Braindrop: A mixed-signal neuromorphic architecture with a dynamical systems-based programming model," *Proceedings of the IEEE*, vol. 107, no. 1, pp. 144–164, 2019.
- [432] S. Moradi, N. Qiao, F. Stefanini, and G. Indiveri, "A scalable multicore architecture with heterogeneous memory structures for dynamic neuromorphic asynchronous processors (dynaps)," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 106–122, 2018.
- [433] C. Frenkel, M. Lefebvre, J.-D. Legat, and D. Bol, "A 0.086-mm² 12.7-pj/sop 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28-nm cmos," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 1, pp. 145–158, 2019.
- [434] AMD Inc., "First space-grade Versal AI Core devices to ship early 2023," p. 1, nov 2022. [Online]. Available: <https://www.amd.com/en/press-releases/2022-11-15-amd-announces-protect/discretionary{\char\hyphenchar\font}{\}\}\}\completion-class-b-qualification-for-first-space-grade>
- [435] I. Rodriguez, L. Kosmidis, O. Notebaert, F. J. Cazorla, and D. Steenari, "An On-board Algorithm Implementation on an Embedded GPU: A Space Case Study," in *Design, Automation and Test in Europe Conference and Exhibition, DATE 2020*. Grenoble, France: Institute of Electrical and Electronics Engineers Inc., mar 2020, pp. 1718–1719.
- [436] G. Pang, "The AI Chip Race," *IEEE Intelligent Systems*, vol. 37, no. 2, pp. 111–112, 2022, accessed: 2022-9.
- [437] LUXonis, "BW1099EMB - Myriad X SoM," Datasheet, nov 2020.
- [438] NVIDIA Corp., "NVIDIA Jetson Roadmap," p. 1, 2022. [Online]. Available: <https://developer.nvidia.com/embedded/develop/roadmap>
- [439] —, "NVIDIA Jetson Nano System-on-Module," 2014. [Online]. Available: https://developer.download.nvidia.com/assets/embedded/secure/jetson/Nano/docs/JetsonNano{_}DataSheet{_}DS09366001v1.1.pdf
- [440] —, "NVIDIA Jetson TX2 Series System-on-Module Data Sheet," NVIDIA Corp., Datasheet, 2021.
- [441] —, "NVIDIA Jetson TX2 NX System-on-Module," NVIDIA Corp., Tech. Rep., mar 2022.
- [442] —, "NVIDIA Jetson Xavier NX Series System-on-Module," NVIDIA Corp., Datasheet, 2021. [Online]. Available: <https://developer.nvidia.com/embedded/downloads>
- [443] —, "NVIDIA Jetson AGX Xavier Series System-on-Module," Datasheet, 2021.
- [444] —, "NVIDIA Jetson Orin NX Series," NVIDIA Corp., Datasheet, 2022. [Online]. Available: http://www.papersearch.net/view/detail.asp?detail{_}key=10000715
- [445] —, "NVIDIA Jetson AGX Orin Series," Tech. Rep., mar 2022. [Online]. Available: <http://ci.nii.ac.jp/naid/40005137127/>
- [446] —, "NVIDIA Jetson Orin Nano Series," NVIDIA Corp., Datasheet, 2022. [Online]. Available: http://www.papersearch.net/view/detail.asp?detail{_}key=10000715
- [447] —, "Introducing Jetson Orin Nano," 2022. [Online]. Available: <https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-orin/>
- [448] Qualcomm, "Qualcomm Cloud AI 100 Purpose-built for high performance, low-power AI processing in the cloud." Qualcomm Technologies Inc., Tech. Rep., 2020.
- [449] —, "Qualcomm Cloud AI 100 Edge Development Kit," Qualcomm Technologies Inc., Product Brief, 2020. [Online]. Available: <https://www.qualcomm.com/products/cloud-artificial-intelligence/cloud-ai>
- [450] —, "Qualcomm Cloud AI 100 MLPerf v2.0 Inference Benchmarks," Qualcomm Technologies Inc., Tech. Rep., 2022.
- [451] AMD Inc., "AMD Instinct MI200 Series Accelerator," AMD Inc., AMD Inc., Tech. Rep., 2021.

- [452] —, “AMD Instinct MI210 Accelerator,” AMD Inc., AMD Inc., Tech. Rep., 2022.
- [453] AMD XILINX Inc., “Versal AI Edge Series,” AMD XILINX Inc., Product Brief, 2021. [Online]. Available: <https://www.xilinx.com/products/silicon-devices/acap/versal-ai-edge.html#{#}applications>
- [454] XILINX Inc., “ACAP at the Edge with the Versal AI Edge Series,” XILINX Inc., White Paper, 2021. [Online]. Available: <https://docs.xilinx.com/v/u/en-US/wp518-ai-edge-intro>
- [455] AMD XILINX Inc., “Versal ACAP AI Edge Series,” AMD XILINX Inc., Product Selection Guide, 2022. [Online]. Available: <https://docs.xilinx.com/v/u/en-US/versal-ai-edge-product-selection-guide>
- [456] XILINX Inc., “AI Core Series Versal™ AI Core Series Product Selection Guide,” XILINX Inc., Product Selection Guide, 2019.
- [457] AMD XILINX Inc., “Versal Architecture and Product Data Sheet: Overview,” AMD XILINX Inc., San Jose, CA, Data Sheet, nov 2022.
- [458] —, “AI Inference with Versal AI Core Series,” AMD XILINX Inc., San Jose, CA, Tech. Rep., 2022.
- [459] M. Andersch, V. Sarge, and P. Micikevicius, “Tensor Core DL Performance Guide,” NVIDIA Corp., Tech. Rep., 2019.
- [460] XILINX Inc., “Beamforming Implementation on AI Engine,” XILINX Inc., San Jose, CA, Tech. Rep., jan 2021. [Online]. Available: <https://docs.xilinx.com/t/en-US/xapp1352-beamforming-ai-engine/Summary>
- [461] —, “Xilinx AI Engines and their applications,” XILINX Inc., White Paper, 2020. [Online]. Available: <https://docs.xilinx.com/v/u/en-US/wp506-ai-engine>
- [462] D. Steenari, K. Forster, D. O’Callaghan, M. Tali, C. Hay, M. Cebecauer, M. Ireland, S. McBerren, and R. Camarero, “Survey of High-Performance Processors and FPGAs for On-Board Processing and Machine Learning Applications,” in *European Workshop on On-Board Data Processing (OBDP)*, ESA, Ed. Online: ESA, jun 2021, p. 28. [Online]. Available: <https://atpi.eventsair.com/QuickEventWebsitePortal/obdp-2021/website/Agenda/AgendaItemDetail?id=1dea8c4d-04a0-4f40-8d9c-2bee77a904d7>
- [463] R. Ginosar, D. Goldfeld, P. Aviely, R. Golan, A. Meir, F. Lange, D. Alon, T. Liran, and A. Shabtai, “Ramon Space RC64-based AI/ML Inference Engine,” in *European Workshop on On-Board Data Processing (OBDP)*, jun 2021, pp. 1–33. [Online]. Available: <https://zenodo.org/record/5521124>
- [464] L. Kosmidis, I. Rodríguez, Á. Jover, S. Alcaide, J. Lachaize, J. Abella, O. Notebaert, F. J. Cazorla, and D. Steenari, “GPU4S (GPUs for Space): Are we there yet?” in *European Workshop on On-Board Data Processing (OBDP)*, ESA, Ed. Online: ESA, jun 2021, pp. 1–8. [Online]. Available: <https://atpi.eventsair.com/QuickEventWebsitePortal/obdp-2021/website/Agenda/AgendaItemDetail?id=c719851f-003e-47ad-81ca-0255dcd1f651>
- [465] I. Rodríguez, L. Kosmidis, J. Lachaize, O. Notebaert, and D. Steenari, “Design and Implementation of an Open GPU Benchmarking Suite for Space Payload Processing,” *Universitat Politècnica de Catalunya*, pp. 1–6, 2019.
- [466] D. Steenari, L. Kosmidis, I. Rodríguez-Ferrandez, A. Jover-Alvarez, and K. Forster, “OBPMark (On-Board Processing Benchmarks) - Open Source Computational Performance Benchmarks for Space Applications,” in *European Workshop on On-Board Data Processing (OBDP)*, ESA, Ed. On-line: ESA, jun 2021, pp. 14–17. [Online]. Available: <https://atpi.eventsair.com/QuickEventWebsitePortal/obdp-2021/website>
- [467] L. Kosmidis, J. Lachaize, J. Abella, O. Notebaert, F. J. Cazorla, and D. Steenari, “GPU4S: Embedded GPUs in Space,” in *22nd Euromicro Conference on Digital System Design, DSD 2019*. Kallithea, Greece: Institute of Electrical and Electronics Engineers Inc., aug 2019, pp. 399–405.
- [468] L. Kosmidis, I. Rodríguez, Á. Jover, S. Alcaide, J. Lachaize, J. Abella, O. Notebaert, F. J. Cazorla, and D. Steenari, “GPU4S: Embedded GPUs in space - Latest project updates,” *Microprocessors and Microsystems*, vol. 77, p. 103143, sep 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0141933120303100>
- [469] H. Marques, K. Foerster, M. Bargholz, M. Tali, L. Mansilla, and D. Steenari, “Development methods and deployment of machine learning model inference for two Space Weather on-board analysis applications on several embedded systems,” in *ESA Workshop on Avionics, Data, Control and Software Systems (ADCSS)*. ESA, nov 2021, pp. 1–14.
- [470] AMD XILINX Inc., “5G Beamforming with Versal AI Core Series,” AMD XILINX Inc., San Jose, CA, Tech. Rep., 2021.
- [471] —, “XQR Versal for Space 2.0 Applications,” AMD XILINX Inc., San Jose, CA, Product Brief, 2022.
- [472] J. Cornebise, I. Oršolić, and F. Kalaitzis, “Open high-resolution satellite imagery: The worldstrat dataset—with application to super-resolution,” *arXiv preprint arXiv:2207.06418*, 2022.
- [473] A. Van Etten, D. Lindenbaum, and T. M. Bacastow, “Spacenet: A remote sensing dataset and challenge series. arxiv 2018,” *arXiv preprint arXiv:1807.01232*, 1807.
- [474] M. Schmitt, S. A. Ahmadi, and R. Hänsch, “There is no data like more data-current status of machine learning datasets in remote sensing,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 1206–1209.
- [475] S. Ayesha, M. K. Hanif, and R. Talib, “Overview and comparative study of dimensionality reduction techniques for high dimensional data,” *Information Fusion*, vol. 59, pp. 44–58, 2020.
- [476] J. Zhang, J. Yu, and D. Tao, “Local deep-feature alignment for unsupervised dimension reduction,” *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2420–2432, 2018.
- [477] G. T. Reddy, M. P. K. Reddy, K. Lakshmana, R. Kaluri, D. S. Rajput, G. Srivastava, and T. Baker, “Analysis of dimensionality reduction techniques on Big Data,” *IEEE Access*, vol. 8, pp. 54 776–54 788, 2020.
- [478] R. M. Gahar, O. Arfaoui, M. S. Hidri, and N. B. Hadj-Alouane, “A distributed approach for high-dimensionality heterogeneous data reduction,” *IEEE Access*, vol. 7, pp. 151 006–151 022, 2019.
- [479] O. A. Abdullah, H. Al-Hraishawi, and S. Chatzinotas, “Deep learning-based device-free localization in wireless sensor networks,” in *IEEE Wireless Commun. Net. Conf. (WCNC)*, Mar. 2023, pp. 1–6.
- [480] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, “Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations,” in *Int. Conf. on machine learning*, 2009, pp. 609–616.
- [481] R. Abdulhammed, H. Musafer, A. Alessa, M. Faezipour, and A. Abuzneid, “Features dimensionality reduction approaches for machine learning based network intrusion detection,” *Electronics*, vol. 8, no. 3, 2019.
- [482] H. Shirmard, E. Farahbakhsh, A. Beiranvand Pour, A. M. Muslim, R. D. Müller, and R. Chandra, “Integration of selective dimensionality reduction techniques for mineral exploration using ASTER satellite data,” *Remote Sensing*, vol. 12, no. 8, 2020.
- [483] G. Furano, G. Meoni, A. Dunne, D. Moloney, V. Ferlet-Cavrois, A. Tavoularis, J. Byrne, L. Buckley, M. Psarakis, K.-O. Voss, and L. Fanucci, “Towards the use of artificial intelligence on the edge in space systems: Challenges and opportunities,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 35, no. 12, pp. 44–56, 2020.
- [484] W. Li and M. Liewig, “A survey of ai accelerators for edge environment,” in *Trends and Innovations in Information Systems and Technologies: Volume 2 8*. Springer, 2020, pp. 35–44.
- [485] K. Berggren, Q. Xia, K. K. Likharev, D. B. Strukov, H. Jiang, T. Mikolajick, D. Querlioz, M. Salinga, J. R. Erickson, S. Pi et al., “Roadmap on emerging hardware and technology for machine learning,” *Nanotechnology*, vol. 32, no. 1, p. 012002, 2020.
- [486] V. Monzon Baeza, F. Ortiz, S. Herrero Garcia, and E. Lagunas, “Enhanced communications on satellite-based iot systems to support maritime transportation services,” *Sensors*, vol. 22, no. 17, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/17/6450>
- [487] S. Cheng, Y. Gao, X. Li, Y. Du, Y. Du, and S. Hu, *Blockchain Application in Space Information Network Security: Third International Conference, SINC 2018, Changchun, China, August 9–10, 2018, Revised Selected Papers*, 01 2019, pp. 3–9.
- [488] L. Pellaco, N. Singh, and J. Jaldén, “Spectrum prediction and interference detection for satellite communications,” 10 2019.
- [489] H. Xu, P. V. Klaine, O. Onireti, B. Cao, M. Imran, and L. Zhang, “Blockchain-enabled resource management and sharing for 6g communications,” *Digital Communications and Networks*, vol. 6, no. 3, pp. 261–269, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352864820300249>
- [490] S. Laaper, “Using blockchain to drive supply chain transparency future trends in supply chain.” [Online]. Available: <https://www2.deloitte.com/us/en/pages/operations/articles/blockchain-supply-chain-innovation.html>
- [491] R. Xu, Y. Chen, E. Blasch, and G. Chen, “Exploration of blockchain-enabled decentralized capability-based access control strategy for space situation awareness,” *Opt. Eng.*, vol. 58, no. 4, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352864820300249>
- [492] H. Ibrahim, M. A. Shouman, N. A. El-Fishawy, and A. Ahmed, “Literature review of blockchain technology in space industry: Challenges and applications,” in *2021 International Conference on Electronic Engineering (ICEEM)*, 2021, pp. 1–8.

- [493] L. Clark, Y.-C. Tung, M. Clark, and L. Zapanta, "A blockchain-based reputation system for small satellite relay networks," in *2020 IEEE Aerospace Conference*, 2020, pp. 1–8.
- [494] F. Tang, C. Wen, L. Luo, M. Zhao, and N. Kato, "Blockchain-based trusted traffic offloading in space-air-ground integrated networks (sagin): A federated reinforcement learning approach," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 12, pp. 3501–3516, 2022.
- [495] S. Prakash, M. Stewart, C. Banbury, M. Mazumder, P. Warden, B. Plancher, and V. J. Reddi, "Is tinyml sustainable? assessing the environmental impacts of machine learning on microcontrollers," *arXiv preprint arXiv:2301.11899*, 2023.
- [496] J. ur Rehman, H. Al-Hraishawi, and S. Chatzinotas, "Quantum approximate optimization algorithm for knapsack resource allocation problems in communication systems," in *IEEE Int. Conf Commun. (ICC)*, Rome, Italy, Jun. 2023, pp. 1–6.
- [497] H. Al-Hraishawi, J. ur Rehman, M. Razavi, and S. Chatzinotas, "Characterizing and utilizing the interplay between quantum technologies and non-terrestrial networks," 2022.
- [498] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, ai-based, or both?" *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7331–7376, 2019.
- [499] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proceedings of the IEEE*, 2023.
- [500] M. Kochupillai, M. Kahl, M. Schmitt, H. Taubenböck, and X. X. Zhu, "Earth observation and artificial intelligence: Understanding emerging ethical issues and opportunities," *IEEE Geoscience and Remote Sensing Magazine*, 2022.
- [501] F. G. Ortíz-Gómez, R. Martínez-Rodríguez-Osorio, and S. Landeros-Ayala, "Method of optimizing the costs of a satellite network in ka and q/v bands in the feeder link," in *35th AIAA International Communications Satellite Systems Conference*, 2017, p. 5442.

TABLE V
AI CHIPSET/ACCELERATORS OFF-THE-SHELF: CORE UNITS

Device	Provider	Core units	
		CPU	On-chip accelerator
Myriad Family	Intel	2x Leon 4 RISC	Image/Video PA SHAVE
Jetson Nano	NVIDIA	4x ARM Cortex-A57MP	128-CUDA Maxwell
Jetson TX2 Family	NVIDIA	2x ARM64b Denver	256-CUDA Pascal
		4x ARM Cortex-A57MP	
Jetson Xavier NX Family	NVIDIA	6x ARM64b Carmel	384-CUDA Volta
			2x NVDLA
			2x PVA
			48 Tensor
Jetson AGX Xavier Family	NVIDIA	8x Carmel ARM64b	512-CUDA Volta
			2x NVDLA
			2x PVA
			64 Tensor
Jetson Orin NX Family	NVIDIA	6x/8x ARM64b Cortex-A78AE	1024-CUDA Ampere
			1x/2x NVDLAv2
			1x PVAv2
			32 Tensor
Jetson AGX Orin Family	NVIDIA	8x/12x ARM64b Cortex-A78AE	1782/2048-CUDA Ampere
			2x NVDLAv2
			1x PVAv2
			56/64 Tensor
Jetson Orin Nano Family	NVIDIA	6x ARM64b Cortex-A78AE	512/1024-CUDA Ampere
			16/32 Tensor
Cloud AI 100 Family	Qualcomm	Snapdragon 865 MP	Cloud AI 100
		Kryo 585 CPU	
Instinct MI200 Family	AMD	CDNA2	6656/14080-Stream Proc.
			104/220 Core Units
Versal AI Edge Family	AMD XILINX	2x ARM64 Cortex-A72	8-304 AI Engines/ML
		2x ARM Cortex-R5F2	90-1312 DSP Eng
			43k-1139k System Logic Cells
Versal AI Core Family	AMD XILINX	2x ARM64 Cortex-A72	128-400 AI Engines
		2x ARM Cortex-R5F2	928-1968 DSP Eng
			504k-1968k System Logic Cells

TABLE VI
AI CHIPSET/ACCELERATORS OFF-THE-SHELF: COMPUTER CAPACITY PER
OPERATION

Device	Computer capacity in operations			Power (W)
	18 (Ops)	FP16 (FLOPs)	FP32 (FLOPs)	
Myriad Family	NR	NR	NR	1~2
Jetson Nano	NR	512G	NR	5-10
Jetson TX2 Family	NR	x	NR	7.5-20
Jetson Xavier NX Family	21T	x	NR	10-20
Jetson AGX Xavier Family	30-32T	10T	NR	10-40
Jetson Orin NX Family	70-100T Sparse	x	x	10-25
	35-50T Dense			
	20T Sparse			
Jetson AGX Orin Family	108-170T Sparse	58-85T	3.3-5.3T	15-60
	92-105T Sparse	67-106T		
Jetson Orin Nano Family	2-4T Sparse	x	x	5-15
	1-1.33T Dense			
Cloud AI 100 Family	70-400T	35-200T	x	15-75
Instinct MI200 Family	181-383T	181-383T	45.3-95.7T (Matrix)	300-560
			22.6 - 45.9T	
Versal AI Edge Family	5-202T	NR	0.4-16.6T	6-75
	0.6-9.1T		0.1-2.1T	
	1-17T			
Versal AI Core Family (VC1902)	133T	NR	8T	~87W
	13.6T		3.2T	
	29T			

TABLE VII
AI CHIPSET/ACCELERATORS OFF-THE-SHELF: DIMENSION AND
AVAILABILITY

Device	Form Factor	Status	Comments
Myriad Family	8~14×9.5~14 mm ²	Released	
Jetson Nano	80×100 mm ²	Released	
Jetson TX2 Family	45~50×69.6~87 mm ²	Released	
Jetson Xavier NX Family	45×69.6 mm ²	Released	
Jetson AGX Xavier Family	87 × 100 mm ²	Released	
Jetson Orin NX Family	45 × 69.6 mm ²	In development	Planned for Jan 2023
Jetson AGX Orin Family	87 × 100 mm ²	In development	Planned for Dec 2022
Jetson Orin Nano Family	45 × 69.6 mm ²	In development	Planned for Jan 2023
Cloud AI 100 Family	46~68.9×110~169.5 mm ²	Released	PCIe Interface
Instinct MI200 Family	111×267 mm ²	Released	PCIe Interface
Versal AI Edge Family	23~45×23~45 mm ²	Released	No development board available
Versal AI Core Family	35~45×35~45 mm ²	Released	VCK190 190×241.3 mm ²