
SYED JUNAID NAWAZ, (Senior Member, IEEE), SHREE K. SHARMA, (Senior Member, IEEE), SHURJEEL WYNE, (Senior Member, IEEE), MOHAMMAD N. PATWARY, (Senior Member, IEEE), MD ASADUZZAMAN, (Member, IEEE)

1Department of Electrical Engineering, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan. (e-mail: junaidnawaz@ieee.org and shurjeel.wyne@comsats.edu.pk)
2SnT - securityandtrust.lu, University of Luxembourg, Kirchberg, Luxembourg 1359, Luxembourg. (e-mail: shree.sharma@uni.lu)
3School of Computing and Digital Technology, Birmingham City University, Birmingham, UK. (e-mail: mohammad.patwary@bcu.ac.uk)
4School of Creative Arts and Engineering, Staffordshire University, UK. (e-mail: asad@ieee.org)

Corresponding author: Syed Junaid Nawaz (e-mail: junaidnawaz@ieee.org).

The first and third authors would like to acknowledge the financial support by EU ATOM-690750 research project approved under the call identifier H2020-MSCA-RISE-2015.

ABSTRACT The upcoming 5th Generation (5G) of wireless networks is expected to lay a foundation of intelligent networks with the provision of some isolated Artificial Intelligence (AI) operations. However, fully-intelligent network orchestration and management for providing innovative services will only be realized in Beyond 5G (B5G) networks. To this end, we envisage that the 6th Generation (6G) of wireless networks will be driven by on-demand self-reconfiguration to ensure a many-fold increase in the network performance and service types. The increasingly stringent performance requirements of emerging networks may finally trigger the deployment of some interesting new technologies such as large intelligent surfaces, electromagnetic-orbital angular momentum, visible light communications and cell-free communications – to name a few. Our vision for 6G is – a massively connected complex network capable of rapidly responding to the users’ service calls through real-time learning of the network state as described by the network-edge (e.g., base-station locations, cache contents, etc.), air interface (e.g., radio spectrum, propagation channel, etc.), and the user-side (e.g., battery-life, locations, etc.). The multi-state, multi-dimensional nature of the network state, requiring real-time knowledge, can be viewed as a quantum uncertainty problem. In this regard, the emerging paradigms of Machine Learning (ML), Quantum Computing (QC), and Quantum ML (QML) and their synergies with communication networks can be considered as core 6G enablers. Considering these potentials, starting with the 5G target services and enabling technologies, we provide a comprehensive review of the related state-of-the-art in the domains of ML (including deep learning), QC and QML, and identify their potential benefits, issues and use cases for their applications in the B5G networks. Subsequently, we propose a novel QC-assisted and QML-based framework for 6G communication networks while articulating its challenges and potential enabling technologies at the network-infrastructure, network-edge, air interface and user-end. Finally, some promising future research directions for the quantum- and QML-assisted B5G networks are identified and discussed.

INDEX TERMS 6G, B5G, Machine Learning, Quantum Communications, Quantum Machine Learning

I. INTRODUCTION

Research interests in data-driven adaptive and intelligent methods have strongly reemerged in recent years [1, 2]. This renewed interest has emerged in part due to the advancements in classical computing methods and partly due to the tremendous potential of parallelism offered by Quantum Computing (QC) and related quantum technologies. These advancements in computing methods have led to the consideration of deploying Machine Learning (ML) as a potential alternative to the conventional logic-based approaches. ML is not only believed to have a strong potential in the network operations spanning from autonomous man-
agement and service classification but also in addressing the re-configurability demands of the future systems. These data-driven learning and quantum-powered computing methods have a strong potential in realizing the ambitions of a service-driven fully intelligent 6th Generation (6G) communication network. In the emerging paradigm of increasing human and machine inter-connectivity, a significant proliferation in the number of network nodes and data traffic is expected [3, 4].

Towards provisioning this massive connectivity and efficiently processing the voluminous data available at the user and network sides of the Beyond 5G (B5G) networks, this paper proposes a novel framework based on QC-assisted ML and Quantum ML (QML) as enabling technologies. Specifically, in the next subsections, we discuss notable recent studies on B5G networks, and also survey recent works on ML and Quantum communications for B5G networks. The major contributions of this work are then presented.

A. WHY B5G NETWORKS?
The 5G wireless networks have recently started to be deployed in some parts of the world but the goal of a fully intelligent network furnishing everything as a service and rendering a completely immersive user experience remains elusive. With 5G reaching its limits in the next decade or so, the design goals for its successor have already begun to be explored in the literature.

The research community has now begun to discuss the vision for 6G communication networks under different labels such as B5G, 5G+, and 6G. In this context, a few articles discussing the vision and open challenges for 6G have recently appeared in the literature, for example, see [5–9]. The authors in [5] have discussed various performance requirements and potential technologies for 6G. In [6], a vision for 6G communications was presented and its perceived requirements were discussed based on an extrapolation of the evolution trends of previous mobile network generations, i.e., 1G through 5G. Recently, [7] provided an overview of the limitations of 5G networks in the context of meeting the growing network performance demands. The authors also discussed some revolutionary new technologies to meet these demands in 6G networks. It has been speculated in [7] that 5G will reach its performance limits within 10 years of its launch, and 6G will be required to deliver further increases of $100 \times$ and $50 \times$ in the individual and the downlink data rates, respectively. The authors in [8] have motivated the need for 6G through a gap-analysis between the original ambitions and maturing 5G networks. Moreover, the authors also presented a vision of future services and technologies based on a new communication infrastructure. The authors in [9] have pointed out some drawbacks of emerging 5G communication networks, and they have also discussed interesting 6G communication trends that can potentially address these shortcomings.

The 6G networks are widely projected to provide an increase of $100 \times$ in volumetric spectral and energy efficiency (in bps/Hz/m$^3$) relative to the 5G networks and they will have a very complex structure incurred from massive connectivity. The global mobile data traffic is forecasted to grow 55% annually between 2020 and 2030 [10]. This growing traffic will be generating 5,016 ExaByte (EB) data per month, by the year 2030. This tremendous amount of data may be harnessed, with strong processing and learning capabilities, to manage the network at different levels. To this end, ML and QC methods can play a significant enabling role.

B. MACHINE LEARNING FOR B5G NETWORKS
ML is a subbranch of Artificial Intelligence (AI) in which machines learn, perform, and improve their operations by exploiting the operational knowledge and experience gained in the form of data [1]. Based on the nature of available data and explicitness of the learning objectives, ML is usually classified into three major paradigms, i.e., supervised, unsupervised, and reinforcement learning. In this regard, authors in [11] have reviewed the history of these ML paradigms and their compelling applications in communication networks. The authors have also reviewed the ML prospects for the optimization of various performance metrics including data-rate, latency and reliability in the context of cognitive radios, heterogeneous networks, Internet of Things (IoT), and Machine-to-Machine (M2M) communications.

ML can potentially assist big data analytics to realize self-sustaining and proactive wireless networks [12]. Various potential applications of big data analytics and ML in enhancing the performance of communication networks have been pointed out in [4, 13]. For example, ML techniques can be significantly advantageous in addressing the access congestion problem in emerging ultra-dense IoT networks [14]. Furthermore, the scope of employing supervised and unsupervised ML methods across different layers of communication networks have been discussed in [15]. However, the success of data-driven learning solutions is directly linked to the availability of sufficiently large amounts of data and a robust processing capability. Also, the available state-of-art in ML is isolated in terms of ML techniques as well as their operations across different layers of the protocol stack of communication networks. To this end, one of the objectives of this paper is to provide a detailed classification of existing ML techniques along with their applications in B5G communications networks.

On the other hand, Deep Learning (DL) adopts an intensive system structure for representing and learning correlational structures in the available data by proceeding in a supervised, unsupervised, reinforcement, or hybrid fashion. For example, an Artificial Neural Network (ANN) with multiple (deep) transmissive hidden layers is referred to as a Deep Neural Network (DNN). The training and processing of data through conventional ML algorithms, executed on conventional Central Processing Units (CPUs) with a limited number of cores, has a limitation of large processing delays. The recent advances in parallel computing capabilities and distributed learning methods have enabled the deployment of data-driven DL approaches to complement the conventional model-based approaches. For example, the revisited learning
algorithms for exploiting the numerous amount of available processing cores in advanced Graphics Processing Units (GPUs) has demonstrated a remarkable performance gain. Furthermore, the advanced Tensor Processing Units (TPUs) have demonstrated a tremendous parallel processing potential with manifold speed-ups and power efficiency in executing ML algorithms. These advances have demonstrated a profound impact of DL-based solutions across various multi-dimensional signal processing applications, e.g., medical image processing, natural languages processing, and wireless communications, to name a few.

A comprehensive survey on the role of DL in mobile and wireless communication networks is presented in [16], where various DL platforms, architecture, and libraries suitable for applications in communication networks are indicated. Also, the motivation behind the use of DNNs in designing and operating the future wireless networks is extensively discussed in [17]. The accuracy of the estimated or prior-available statistics of radio propagation channels is of vital importance in enhancing the capacity of wireless communication links. The concept of auto-encoding an end-to-end communication system within a DNN for jointly optimizing the operations of both transmitter and receiver sides to best counter the channel impairments has recently emerged with a strong potential. For example, a DNN-based end-to-end learning system is proposed in [18], where a channel agnostic learning based system is proposed for learning the channel output through a conditional Generative Adversarial Net (GAN).

Bringing intelligence to the physical layer of the communication systems can empower smart estimation of parameters, mitigation of interference, and the management of resources [19]. As an example, DL capabilities can be utilized for channel estimation and symbol detection in Orthogonal Frequency-Division Multiplexing (OFDM) systems [20]. Furthermore, DL has also received significant research interests in dynamic allocation and management of radio resource for vehicular communications (i.e., vehicle-to-vehicle (V2V), vehicle-to-everything (V2X), etc), where high nodes mobility impose high dynamics in the channel characteristics. For example, a Deep Reinforcement Learning (DRL) based decentralized resource allocation mechanism can be utilized to support highly dynamic communication applications [21].

Massive Multiple-Input Multiple-Output (M-MIMO) systems and millimeter-Wave (mmWave) spectrum exploiting high spatial resolution and multi-gigahertz bandwidth, respectively, are believed to have an important role in addressing the capacity demands of future communication networks [12, 22]. One of the potential applications of employing DL methods in such mmWave M-MIMO systems is the estimation of radio channel quality [23], which is essential for the design of transmission techniques such as beamforming. Also, the potential of deploying DL for various other tasks across all the communication layers has also received notable attention [21, 24–26], e.g., intelligent localization, radio identification, routing, channel tracking, routing, and caching. The integration of DL capability with the smart city infrastructure can help in effective utilization of the available big data for accomplishing the dream of the cognitive smart world of the future [27].

However, DL methods lack an efficient mechanism for prior evaluation of the best choice of training algorithm, size and structure of DNN, and parameters setting that suits the model/problem under consideration. The hit-and-trial snoop- ing along a very large set of possibilities in structure, size, algorithm, and parameter-value makes the deployment of DL not only cumbersome but it can also lead towards the loss of balance between underfitting and overfitting. To this end, this paper provides a review of the existing related works, identifies the potential issues and discusses emerging DL methods including DNN, deep transfer learning, and deep unfolding.

C. QUANTUM COMMUNICATIONS FOR 5G NETWORKS

In the quest to meet the rapidly increasing demands of fast, reliable, secure, intelligent, and green communications; the demand for a high computational capability of the systems has also increased expeditiously. The inherent parallelism offered by the fundamental concepts of quantum mechanics and the prospects demonstrated through recent results of QC technology clearly indicate a definite potential to outperform the conventional computing systems [28]. This immense power of QC comes from the fundamental concepts of quantum superposition, quantum entanglement, or the no-cloning theorem [29]. The parallel processing of multi-dimensional large-sized data can be conveniently realized through QC in large tensor product spaces. The QC-assisted communications is another new research area, which is envisioned to hold promise for achieving extremely high data rates and link security in future 6G and beyond communications [30]. To this end, the reliable communication rate of quantum channels for amalgamated noiseless classical communication, quantum communication, and entanglement resources have begun to be investigated in the literature [31, 32].

The emerging solutions for enhancing the link capacity in future communication systems, e.g., power domain multiple access supported by Successive Interference Cancellation (SIC), have very high run-time computational power demands; thus there is a clear scope for exploiting QC. An example of multi-objective space exhaustive-search demanding problem in communications is to determine the optimal data-packet routes in multi-hop communication networks. A quantum-assisted solution to the above problem has been presented in [33], where an evolutionary quantum Pareto optimization algorithm has been proposed. Furthermore, the extension of classical turbo codes to quantum turbo codes with error correction for frequency selective channels has been proposed in [34]. Some examples of recent efforts on quantum-aided solutions for localization, multi-objective routing and load balancing, channel estimation and decoding, and multi-user transmission are discussed in [35], [36], [37], and [38], respectively. Moreover, [30] is a recent survey.
on the existing efforts on employing QC in solving various optimization problems encountered in different layers of wireless communication systems. Nevertheless, there is a need to conduct a comprehensive review of the recent studies on QC-assisted communications and pure quantum communications in order to draw a clear picture of the current state of understanding about these topics.

D. CONTRIBUTIONS OF THIS WORK

The advantages offered by QC and ML methods have collectively emerged into an exciting interdisciplinary framework of QML [39]. In this emerging framework, the ML techniques benefit from exploiting the quantum speedups, whereas the quantum devices’ uncertainties can be resolved with assistance from ML techniques [40]. By combining the established merits of quantum mechanics in producing counter-intuitive statistical patterns, and those of ML techniques in recognizing statistical patterns in data, the QML framework can generate and recognize statistical data patterns that are beyond the capabilities of classical computing or ML methods alone [41]. The research community has only recently begun to explore the applications of QML across various engineering disciplines, see e.g., [42, 43]. This nascent QML framework, having strong synergies with superimposed signals and enmeshed links, can find significant application in communication networks. To the authors’ best knowledge, this work is the first to explicitly investigate the deployment of the QML framework for future communication networks.

Previous works have separately characterized ML, QC, QML, and communication networks, see e.g., [1], [28, 29], [39–43], and [2, 3, 22], respectively. In the literature, the stand-alone application of ML and QC to future communication networks has been intensively studied, as summarized in Table 1, see e.g., [11–13, 15–17, 19, 21, 24, 26, 27] and [30–38], respectively. However, to-date there is no investigation of the joint deployment of QC and ML for future communications, either using QC-assisted ML or the QML framework. This work aims at bridging this gap by conducting a thorough review of QC, ML, and QML in the context of 6G and beyond communication networks. Specifically, the main contributions of this work are listed as follows:

- A review of the 5G target services and their enabling technologies is provided. Moreover, the major open challenges and enabling technologies envisioned for 5G communication networks are elaborated.
- The state-of-the-art of ML, including DL, is thoroughly reviewed in the context of increasingly stringent performance requirements of future communication networks. Moreover, various use cases and potential challenges in the application of DL methods to 5G networks are identified.
- The state-of-the-art of quantum communications, including QC-assisted communications, is comprehensively reviewed. Also, some open research problems in generalization, scalability, and algorithm-parallelization for QML-based communication networks are identified.
- A novel QC-assisted ML and QML-based framework for 6G communication networks is proposed. In the proposed framework context, various potential enabling technologies for 6G at network-infrastructure, network-edge, air interface, and user-side are also discussed thoroughly.
- To stimulate future research activities in the context of the proposed 6G framework, various research problems and some exciting future directions are also identified.

The rest of this paper is organized as follows. In Sec. II, the target services and technology innovations offered by 5G communication networks are identified. Moreover, the open research challenges and performance requirements of B5G communication networks are discussed. In Sec. III, the fundamentals of ML (including DL) and taxonomy of its applications in communication networks are thoroughly discussed. Sec. IV elaborates the fundamentals and state-of-the-art of quantum communications, QC-assisted communications, and QML-based communications. Sec. V proposes a novel framework for 6G and beyond communication networks. Detailed discussions on various exciting future research directions, potential enabling technologies, and open research problems in the context of the proposed framework are also conducted. Finally, the conclusions are drawn in Sec. VI.

Conventions: Quantum communications refers to the communication systems which are purely based on the quantum mechanics concepts. QC-assisted communications refers to the conventional communications exploiting quantum speed-ups. ML-assisted communications refers to communications exploiting ML methods (including DL). QC-assisted ML refers to conventional ML systems exploiting quantum speed-ups. QC-assisted ML based communications refers to the conventional communications exploiting both ML methods and QC speed-ups. QML-based communications refers to the communications exploiting the nascent framework of QML.

II. 5G AND BEYOND COMMUNICATION NETWORKS

The 5G mobile communication networks is envisioned to enable new services to everything at all-time through ultra-fast, low-latency, and ultra-reliable communication links. These networks are not only an evolution of existing networks, as shown in Fig. 1, but also they introduce revolutionary new communication technologies aimed at providing an immersive user experience. The preliminary 5G standardization efforts have matured through 3rd Generation Partnership Project (3GPP) Release 15 [44], which includes specifications for both the non-standalone and standalone operation of the 5G New Radio (NR). Further investigations and field trials are in progress [45] and the 5G standardization is expected to conclude with 3GPP Release 16 in the year 2020. Meanwhile, the commercial deployment of 5G NR in non-
standalone mode has already begun in major cities around the globe.

### A. 5G TARGET SERVICES

In the following, we elaborate on some of the major target services of 5G and discuss the technology innovations envisioned to materialize them. Some of these technologies are radically novel, whereas others may not have matured in time to be included in the 5G standards.

**a: Enhanced Mobile Broadband**

The 5G networks are aimed at providing a 1000-fold increase in the aggregate throughput and a 10-fold increase in the individual link throughput relative to the 4th Generation (4G) wireless networks [46]. The target throughput of up to 20 Gbit/s in the downlink and 10 Gbit/s in the uplink enable services such as ultra high definition video streaming, augmented reality, and TI. At the physical layer, the technology innovations to support these data rates include communications in the mmWave frequency band [47], wherein the large bandwidth can support high data rates; M-MIMO whereby the number of antenna elements at the Base Station (BS) is much larger in proportion to the serviced users [48] such that multiple data pipes can be established over the same time and bandwidth resource. Finally, the Ultra-Dense Network (UDN) strategy [49], which entails an aggressive deployment of multiple small-cells within a macro-cell, can also provide increased data rates to its associated users who are typically in close proximity of their small cell BS and enjoy favorable wireless propagation conditions.

**b: Ultra Reliable Low Latency Communications (URLLC)**

The provision of URLLC is a novel service paradigm offered in 5G networks. Both the reliability aspect, with packet error rates \( \leq 10^{-5} \), and end-to-end latencies \( \sim 1\text{ms} \) aim at supporting new use cases such as factory automation, autonomous driving, e-health, building automation, and smart cities, to name a few [50]. To enable these and other services, the 5G network infrastructure is based on the revolutionary novel concepts of Network Function Virtualization (NFV) [51] and end-to-end Network Slicing (NS) [52]. In the NFV approach to network design, many network services such as network address translation, domain name service, and caching are decoupled from propriety hardware and implemented in software that runs on off-the-shelf hardware. The NS concept allows multiple logical networks or slices to operate on a shared physical infrastructure. Each network slice has dedicated resources for computation and storage as well as data-traffic isolation from the other slices to create a true end-to-end virtual network. With NFV and NS, the physical network resources can be optimized to provide URLLC services for safety critical applications such as vehicular communications or remote-robotic-surgery. Another infrastructure evolution to support low latency communications is edge-computing architectures including Mobile Edge Computing (MEC) [53]. In the MEC paradigm, many of the data processing tasks are moved to the cellular BS or similar edge node, which also has the ability to cache content, thus minimizing the service time to its proximal network users.

**c: Massive Machine Type Communication (mMTC)**

With the advent of the IoT, a very large number of low-rate, low-power devices require an internet connection. These devices may be typically used for environment sensing and utility metering applications and only require intermittent communications with small data payloads [3]. The mMTC service aims at providing internet connectivity to such devices. While many of the 5G mMTC service components have been developed as part of the previous 3GPP releases, those services that need URLLC will require the 5G Core network deployment. The mMTC service is enabled by the flexible combination of NFV and NS, which may provide automated network functions to the mMTC devices without incurring heavy operational expenditures for the mobile service provider. Also, the Non-orthogonal Multiple Access (NOMA) scheme in 5G is seen as an enabler of mMTC connections by allowing grant-free uplink connections to the energy constrained mMTC devices and saving them the control signaling overhead [54]. One promising architecture to support mMTC as well as URLLC services is a collaborative edge-cloud framework which can utilize the benefits of both the cloud-computing and edge-computing towards handling a large amount of data and providing timely feedback to the end-users, respectively [55].

**d: Tactile Internet (TI)**

The IoT enables the interconnection of smart devices and the TI can be viewed as an evolution of the IoT to enable the real-world...
time control of the IoT [56]. The TI allows for a real-time interaction of humans and machines that supports haptic input with audiovisual feedback to control the machine operation in real-time. Some representative examples include remote-controlled robotics for hazardous or difficult-to-access scenarios in the manufacturing industry and tele-diagnosis and remote-robotic-surgery in the healthcare industry. The low end-to-end latency required by these services can be enabled with MEC, which is also supported by intelligent predictive content caching at the edge node [56].

**B. BEYOND 5G: OPEN CHALLENGES AND EMERGING TECHNOLOGIES**

While 5G networks have introduced many technology innovations, the network’s stringent performance requirements have also raised new design considerations. Below, we list some of these challenges and discuss how they may be addressed by some emerging technologies that can feature in the evolution towards 6G communication networks.

**a: Throughput**

Following the evolution trends of previous mobile network generations, the bit/s throughput targets in 6G networks are expected to increase by an order of magnitude relative to those of 5G. Additionally, the virtual reality applications once matured will require much higher data rates than those promised by 5G. For these reasons, individual user data rates of up to 100 Gbit/s are envisioned for 6G [7]. These high bit-rates can be supported in 6G by a large communication bandwidth, which is available in the higher range of the mmWave band between 100 GHz and 300 GHz. Also, large portions of free spectrum are available in the tera Hertz (THz) frequency band. Owing to the large propagation losses when communicating in these bands, the mmWave and THz communications in 6G will be typically employed for high bit-rate short-range communications. The visible light communications (VLC) using data-modulated white light-emitting diodes (LEDs) as transmitters and photo diodes as receivers, is another enabling technology that can support extremely high bit-rates in Line-of-Sight (LoS) connections [57]. These Gbit/s links are made possible by the fact that the optical spectrum’s bandwidth is significantly larger than that of the radio spectrum, and moreover, it is free to use. Furthermore, another promising technology to enhance the spectral efficiency of future wireless networks is full-duplex technology, which enables the concurrent sensing and transmission or concurrent transmission and reception over the same radio frequency channel [58].

**b: Network Capacity**

Traditionally, the cell-densification strategy has been the prime enabler of increasing network capacity. However, shrinking the cell-size (e.g., tiny-cells) also requires suitable management of the increased inter-cell interference to the cell-edge users. With the rise of smart cities, mMTC, and mobile users on land as well as airborne, cell-densification through static BSs alone cannot meet the exponential growth in capacity demands. This problem can be alleviated by hy-
brid cellular networks employing Unmanned Aerial Vehicles (UAVs) as mobile BSs [59]. These UAV BSs can not only offload the data-traffic of the static BSs but they can also relocate dynamically to provide a more favorable propagation channel to the edge users. Also, the increasing demands on the frequency resources can be addressed by sharing the mmWave band between satellite and terrestrial communication networks to give a more global mobile coverage [60]. This 3-Dimensional (3-D) nature of the 6G coverage specifications has led to volumetric descriptions of spectral efficiency requirements in bps/Hz/m³. The introduction of mobile BSs and dynamic spectrum sharing have opened up the possibility of employing ML tools to optimize these new network parameters such as route optimization for the UAV BSs [59] or efficient spectrum sharing [61].

c: Energy Efficiency
Increasing the mobile network’s energy efficiency helps to reduce both its operating expenditures and its carbon emissions. To this end the design efforts for 5G networks have considered energy-efficient approaches to network deployment and resource allocation, including new technologies such as M-MIMO and ultra-dense heterogeneous networks [62]. The energy efficiency is traditionally defined as the bitrate supported per Joule of energy consumed by the communication link. Therefore, if the 6G networks aim to provide more throughput and capacity than 5G networks at similar transmission power levels, then this requires a matching increase in the energy efficiency of 6G networks relative to that of the 5G networks. One promising approach is the use of programmable smart surfaces comprising reconfigurable planar meta-materials [63]. These surfaces can be used to coat walls or other structures and then programmed for the desired interaction with impinging electromagnetic waves to provide beam-steering for Signal-to-Noise Ratio (SNR) maximization or radiation absorption to reduce interference etc. The ML algorithms can be exploited to learn the wireless environment and formulate the appropriate configuration for desired objectives. Also, for conserving device battery life and for powering UAV BSs for uninterrupted operation, the paradigm of wireless power transfer, energy-harvesting and simultaneous wireless information and power transfer [64, 65] may also feature prominently in the 6G standardization efforts.

d: Backhaul and Access Network Congestion
The 6G backhaul traffic will require very low latency optical fiber-equivalent access networks to support the high data rates and quality of service requirements on 6G fronthaul communications. The backhaul congestion issue can be alleviated in part by deploying storage and computation resources at edge nodes in a MEC architecture [53], which can ensure low latency services to the node’s proximal users. Moreover, ML-based proactive content caching at the edge node can also be exploited to avoid backhaul congestion and further reduce the service latency [66, 67]. For the backhaul access network for indoor scenarios, wireless optical communications in visible light spectrum may be explored [68]. For outdoor scenarios, mmWave communications with low earth orbital satellites may provide backhaul service to the static as well as mobile UAV BSs [69].

Besides the congestion in the backhaul networks, congestion in the Random Access Networks (RAN) is another challenging issue to be addressed in the emerging ultra-dense wireless networks. For example, the RAN congestion in ultra-dense IoT networks may arise due to various reasons including the massive number of short-packet transmissions, huge signalling overhead per data packet, and very dynamic and sporadic nature of device transmissions [70]. In this direction, it is necessary to investigate suitable transmission scheduling, peak traffic minimization and access control techniques in the access networks of beyond 5G systems.

e: Data Security
An enormous amount of user-data is propagated and stored on mobile networks in the form of geo-tagged voice and text messages as well as mobile application activity logs. Securing this data from eavesdroppers and its un-authenticated use are of prime importance. In order to secure the 6G communication links, physical layer security schemes [71] may be deployed in tandem with conventional cryptography schemes. Also ML-based schemes for cyber-security [72] and quantum encryption [73] are promising approaches to be explored for securing communication links in future 6G networks.

The foregoing discussions suggest that the 6G communications will leverage robust learning capability at different network layers to perform diverse tasks such as network management, radio resource allocation, data security, and manipulation of smart surfaces to name a few.

III. MACHINE LEARNING FOR COMMUNICATIONS
ML is conventionally thought to have its application justified in the situations, where there is no exact mathematical model of the system available, a sufficiently large amount of training data is available, the system/model under study is stationary (slow varying) along time, and the numerical analysis is acceptable. The ML techniques have recently gained significant attention for the provision of data-driven solutions to various challenging problems in communication systems. The deployment of ML in communications is rapidly gaining popularity; in particular, to build self-sustaining and adaptive networks capable to meet the dynamic reconfigurability demands of the future devices and services. Furthermore, ML has a strong potential to replace the conventional mathematical model-based algorithmic solutions, given the availability of adequate data and computational power. In the following, we present the basics of ML and then discuss the scope of deploying ML at different layers, ends, and the types of communication networks.
A. FUNDAMENTALS AND TAXONOMY OF APPLICATIONS

The taxonomy of applications of different types of learning at different layers of communication systems, along with the available big data at different ends and layers of the network, are highlighted in Table 2. The ML techniques including supervised, unsupervised, and reinforcement learning have various applications in solving several problems across different protocol layers of communication systems, which are discussed in the sequel.

1) Supervised Learning

In supervised learning, the coefficients of intermediate stages are learned by exploiting the prior available set of inputs paired with their corresponding desired outputs. ML can potentially exploit the domain knowledge as well as the training data examples to learn the required behavior and perform the requisite operations. An ideal application of supervised ML can be pronounced as the scenario in which the true joint distribution of input and output parameters is available, which may be extracted from the available domain knowledge. However, there may be scenarios where the mathematical model or true distribution is not known; e.g., an accurate propagation channel model for Body Area Networks (BANs) is not available. In such learning problems, given the test data examples, a model from different classes of models (generative or discriminative) can be exploited to approximate the distribution for performing the learning process. Supervised learning is typically used for the classification and regression nature problems; while the typical examples of its implementation structure can be stated as ANNs, k-Nearest Neighbour (kNN), and Support Vector Machine (SVM).

A bank of ANNs is proposed for symbol decoding in MIMO-OFDM systems in [74]. The available information of transmitted (training/pilot symbols) and the corresponding received symbols at the physical layer of a communication system can be paired together to supervise the learning of ANNs for symbol decoding [74]. Supervised learning for channel compensation in vehicular communications may be challenging, where shortage in training data and/or time is imposed by the mobility of the nodes; this is because higher mobility causes higher Doppler spread which further causes reduction in the coherence-time and this eventually leads to fast variability in the channel statistics. For such fast time-varying channels, a hybrid learning method is proposed in [75], to assist in estimation and tracking of the channels.

Another application of supervised learning at the physical layer for downlink communications can be optimal power allocation and interference cancellation. Applications of supervised learning are not only limited to the physical layer, but ML also has various popular applications in the network, application, transport, and other layers. Satellite links offer the advantage of global network coverage; however, the very high link latency limits its popularity. ML has a popular application in intelligent caching, which can help in reducing the latency in satellite links. Intelligent media/contents prediction has various other applications for enabling low latency communication in remote regions; e.g., intelligent caching and transfer caching at the nodes level in aeroplanes and ships for mesh-networks based airborne internet and oceanic broadband applications [76], respectively. Another potential scenario, where supervised ML can play an important role, is in determining the users association with BSs on the basis of contents/media demand. For land mobile radio communications, echo-state (supervised learning principle) neural network for proactive caching in Cloud Radio Access Networks (CRAN) to predict the contents’ demand and users’ mobility patterns (to predict user association) at the BS is proposed in [77]. The proposed learning method is shown to enhance the network sum effective capacity by about 30% compared to the baseline random caching approaches.

2) Semi-supervised and Unsupervised Learning

In semi-supervised learning, a small amount of annotated training data is available while most of the data is unlabelled; whereas, in unsupervised learning, no annotated training data is available. In unsupervised learning, the collection of available input data samples are exploited to train the system while no prior information of the desired system response is available. For example, at the physical layer, the received noisy data symbols can be used to train a system by clustering the sample points in the decision space for generating effective nonlinear decision boundaries for mapping of the symbols according to the constellation maps. Semi- and un-supervised learning is typically used for clustering and classification natured problems. The implementation structures of such learning methods can be named as: k-Means Clustering (kMC), Principal Component Analysis (PCA), and maximum likelihood learning, etc.

Unsupervised learning can potentially be applied for performing the wide range of tasks related to points clustering, features extraction, features classification, distribution estimation, and distribution specific samples generation. At the physical layer in highly dynamic scenarios of vehicular communications, where less coherence-time limits the available time and data for supervising the learning of channel equalizer; semi-supervised and unsupervised learning can make their way for assisting the channel equalization and tracking operations. The selection of encoding/precoding schemes for performance optimization is another potential application of unsupervised learning. Subsequently, at the higher layers, there are various potential applications of unsupervised and semi-supervised learning for grouping/pairing/clustering of nodes/points for optimal allocation of network/radio resources. Moreover, various potential applications for data analysis include: social networking trends analysis at the network side, phone-apps data analysis at user and networks side, ranking of web resources, data flow prediction, network state prediction, data dimensions reduction, spatial and temporal data analysis, data mining, malware detection and classification.
TABLE 2: Taxonomy of ML applications at different network layers and the required availability of big data.

<table>
<thead>
<tr>
<th>ML Scope</th>
<th>Learning Type</th>
<th>Available Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervised</td>
<td>Optimal modulation, interference cancellation, mobility prediction, spectrum sensing, radio resources optimization, localization, PL security, transmission optimization (lightpaths etc), nodes clustering, duplexing configuration, multiple access, beam switching, etc</td>
<td>Link preservation, channel tracking, on-demand beamforming, secure transmission, energy harvesting, transmit power selection, nodes selection, channel access management, modulation mode selection, coverage optimization, anti-jamming, radio identification, etc</td>
</tr>
<tr>
<td>Unsupervised and Semi-Supervised</td>
<td>Multi-objective routing, traffic control, network state prediction, source encoding/decoding, network-parameters prediction, intrusion detection, fault detection, anomaly detection, etc</td>
<td>Traffic load, services demands, adaptive rate control, network security, capacity and latency demand prediction, traffic prediction and classification, NS, etc</td>
</tr>
<tr>
<td>Reinforcement</td>
<td>Multi-objective routing, packet scheduling, access control, adaptive rate control, network security, capacity and latency demand prediction, traffic prediction and classification, NS, etc</td>
<td>Mobility tracks, traffic statistics, outage statistics, etc</td>
</tr>
</tbody>
</table>

3) Reinforcement Learning

Reinforcement learning is realized on the basis of a feedback performance indicator (termed as a reward) conceived from the environment after computing a specific output for a specific observation by adaptively converging to the ideal behavior through maximization of the reward (performance). This learning technique can be termed as a compromise between supervised and unsupervised learning, where the prior understanding of the ideal system performance provides indirect supervision while there is no available direct training data paired with the desired output. Typically, reinforcement learning is used for control and classification problems; whereas, some notable algorithm examples can be stated as Q-Learning (QL) and Markov decision process.

An agent can be associated with each serving station in a cellular network to assist in learning the optimal scheduling parameters to enhance the network Quality-of-Service (QoS) [78]. A promising application of reinforcement learning at the physical layer of communication networks is power control and optimization. In this regard, a model-free distributed reinforcement learning method for power allocation is proposed [79], in which Channel State Information (CSI) and QoS indicators are exploited to adapt the transmit power.

4) Genetic programming

Inspired from biological evolution, genetic programming evolutionarily evaluates the fitness objectives, given the constraints and limitations, to find an optimal solution to the subject problem. Genetic algorithms are among widely explored methods for resolving various optimization and estimation problems at different layers of communication systems. The Genetic Algorithm (GA) at physical layer of communication systems has been used for optimal antenna selection in MIMO systems, power control, and symbol detection in MIMO systems in [80, 81], [82], [83, 84], respectively. In [85], a detailed review of the scope and applications of evolutionary algorithms in wireless communications is presented. For some communication scenarios, there is no well-defined model of propagation channel available, a few examples to such scenarios are Underwater Acoustic Communication (UWAC) channels, mmWave channels, high mobility (dual-end) channels, molecular communication channels, etc. This makes the estimation and tracking of such channels a challenging task, where any error in channel estimate can significantly affect the symbol detection performance. The scope of GA for estimation of such channels has been investigated in the literature, see e.g., [84, 86] for GA based estimation of sparse channels (UWAC), etc. The scope of GA for intelligent cognitive radio has been encouraged in [87].
5) Learning Requirements and Capability

Model for an ML algorithm can be determined based on the amount and nature of the data in progression. The applications with a big amount of prior available training data, batch-learning algorithms can be applied. Batch-learning algorithms search through the space of all possible data knowledge structures while assuming unlimited available computing time. Such off-line approaches, in which the data is manually obtained, labelled, and then batch-processed, usually face the constraint of limited available data in practical applications. Therefore, the applications with real-time data processing requirements are not well-suited for such batch-learning algorithms. On-line training is a suitable solution for such streaming data applications. However, in online training, only a limited fixed time is available for processing each data sample. A typical application of off-line (batch) and online learning in communication systems can be intelligent caching and channel tracking, respectively. Model-based learning usually optimizes the performance indexes through available objectives functions with high computational efficiency. On the other hand, the pure data samples-based learning exploits all the available data samples to interpolate and/or extrapolate the samples, with high memory and time requirements. A typical application of model-based and samples-based learning can be symbol decoding and contents demand prediction, respectively. The communication prospects of learning requirements and capability of difference ML approaches have been investigated in [13].

B. ARTIFICIAL NEURAL NETWORKS FOR COMMUNICATIONS

ANN is a biologically inspired data processing structure which is designed to learn different operations from the observed data. ANNs are generally used to recognize any patterns within the input data by passing the data through different layers of simulated neural connections. An ANN is composed of connected input, hidden, and output layers of neurons, where each node (neuron) performs combining and/or limiting operations and each connection performs scaling operations. The layers of neurons may be fully connected, partially connected, pooled, feed-forward, recurrent, etc. With the growing applications of ANNs, the topologies of connections between the layers of neurons in a network are rapidly evolving, where a few notable structures can be named as Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Hopfield Network (HN), GAN, Echo State Network (ESN), Neural Turing Machine (NTM) etc. These structures define the flow of data in the network; e.g., in a Feed-Forward Network (FFN), each neuron is connected only to the neurons of the following layer, while an RNN allows the connections from the leading layers to be feedback to the previous layers. Training of ANNs is the process in which the weights of the connections between the neurons are learned. The training of ANNs is usually performed in a supervised learning fashion, where the prior available data labelled with the desired output is exploited to compute the error for adjustment in weights. The error can be quantified on the basis of different metrics, where a natural generic quantifier is Mean Square Error (MSE). The error can be iteratively propagated backwards from output towards the input layer, for quantizing the error at each layer and then updating the weights. A few notable training algorithms for ANNs can be named as gradient descent, conjugate gradient, Newton’s method, Quasi-Newton, and Levenberg-Marquardt, etc.

Deployment of ANNs in communication networks is not a new idea, instead, ANNs have been deployed to perform and assist in various operations of communication systems; e.g., ANN is proposed for symbol decoding for MIMO-OFDM systems in [74]. Recently, an ANN-based method for predicting channel features for large-scale multi-antenna BSs is proposed in [88]. Description of radio propagation channels characteristics for molecular communications is not well established in the literature [89]; for such communication scenarios, an ANN based receiver design is proposed in [90]. Also, ANNs assisted indoor localization method exploiting two variants of the fingerprinting approaches is proposed in [91]. Given the increasing complexity of the future communication networks, ANNs have a wide scope of its deployment in performing various diverse tasks, e.g., planning, optimization, estimation, tracking, controlling, and maintaining tasks, etc.

The size of an ANN (number of neurons and hidden layers) and the amount of available data determine its performance and requisite computational power, as illustrated in Fig. 2. The generalization of an ANN’s operations, for efficiently dealing with every new unseen input data sample, needs its training over a large amount of data (i.e., big data). However, the major limitation is the available computational power to deal with big data and deep structured ANNs. As anticipated in Fig. 2, big data-driven DL can enable the deployment of ANNs in learning complex (high dimensions, a high number of features, a high number of classes, etc) statistical structures and input-output operations of the future communication networks. ANNs are considered as the most notable enablers of DL mechanism [17], in which the understanding of the data’s distribution and important features are automatically learned to simulate an effective mapping function.

C. DEEP LEARNING FOR COMMUNICATIONS

DL is a subbranch of ML, in which the system intelligence is learned through the propagation of input data in massively connected multiple (deep) layers of the system, performing combining, limiting (thresholding), and/or other mathematical operations, in order to compute the output. DL methods can be supervised or unsupervised or a combination of both. A DL system automatically learns to model the high-level abstractions in the given data by extracting important features from it. The scope of emerging DL methods in wireless communication networks is thoroughly investigated in [17], where the deployment of DNNs in the future wireless networks is strongly motivated. Various diverse types of
applications of DL in communication networks have been indicated in the literature, which includes network planning, deployment, resources management, operations control, and maintenance etc. In this context, a DL based framework for optimization of downlink beamforming is proposed in [92]. Furthermore, a comprehensive survey on the developments of DL based mobile and wireless networks is presented in [16]. Moreover, a survey on the recent advancements of employing DL at the physical layer of wireless communication systems is conducted in [93], where the application of DL at the physical layer is categorized into blocked or without blocked structured. Also, introducing DL in emerging 5G communication networks is surveyed in [54], where DL at the physical layer of communication systems for introducing intelligent radio resource allocation mechanism is emphasized. In addition, DL for radio resource management in 5G networks is also suggested in [94], and a users location-aware DL method for run-time optimal users association for the maximization of sum-rate in M-MIMO based land mobile radio cellular networks is proposed in [95]. Another thorough literature survey indicating various emerging applications of DL in communication networks is presented in [96]; where the highlighted applications are: dynamic spectrum access, joint user association, data caching and offloading, security, connection preservation, traffic routing, resource sharing, and power control. There exist various realizations of DL, e.g., DNN, and Deep Boltzmann Machines etc.

1) Deep Neural Networks

Despite ANNs being promising, the limitation of required computational power for the training of an ANN creates challenges for their practical applications. Towards addressing this, the recent advances in Graphical processing units (GPUs) have provided the initial breakthrough by speeding up the training process through their ability to simultaneously perform multiple weight calculations operations. This has opened a new horizon of possibilities of deploying more complex structured ANNs. The DNN can be considered as a typical example of DL, which is an ANN containing a high number of hidden layers of neurons and a complex structure of connections between the neurons. The recent advancements in DNNs has enabled its deployment in even delay critical applications; which is achieved through offline training the DNNs and then performing the online operations (tracking, optimization, etc). An example of such delay critical applications at the physical layer of communication networks is optimal beamforming; where the latency induced by conventional iterative methods makes it outdated for the future networks. The major concern of required long training time in ANNs also stands in DNNs. Looking towards quantum algorithms for training such DNNs may be a futuristic solution, while there is a significant amount of research going on to devise intelligent and efficient learning methods for exploiting massive parallelism in the DNNs architecture.

DNN has also been investigated for auto-encoding an end-to-end communication system. A DNN based auto-encoder for jointly optimizing Bit-Error-Rate (BER) and Peak-to-Average Power Ratio (PAPR) in OFDM systems is proposed in [97]. Furthermore, DNN based architectures for symbol detection in MIMO systems have been proposed in [98], and a DNN-based localization method exploiting fingerprint-based and channel measurements for M-MIMO systems has been proposed in [99]. Moreover, a survey on opportunities and challenges in deploying DL at the physical layer of wireless communication systems has been discussed in [100], where the scope of DNNs for channel estimation and channel encoding/decoding tasks has been thoroughly investigated. The DNN for real-time radio resource management at the physical layer of communication networks has been proposed in [101]. Another popular application of DNNs at the physical layer of communication networks is power control and optimization [102].

2) Deep Transfer Learning

Deep transfer learning is another new direction of research to reduce the dependence of learning from the required large amount of data. The study of transferring the knowledge learned from the available data in a certain context to a new but similar scenario is referred to as transfer learning. This learning technique offers the advantage of facilitating the learning process by reducing the amount of required data, and relaxes the condition on the training data to be independently and identically distributed (i.i.d.) with the validation data. In this regard, a thorough survey of different transfer learning methods has been conducted in [103]. Deep transfer learning is another new approach, which combines DL methods with the transfer learning methods. A survey on deep transfer learning methods has been presented in [104], where the deep transfer learning methods have been classified into instance-, mapping-, network-, and adversarial-based transfer learning categories. In instance-, mapping-, network-, and adversarial-based transfer learning, the instances in source domain are exploited, mapping of instances from two domains into a new space with high similarity is exploited, pre-trained network is partially reused in source domain, and finding of transferable features is performed through adversarial methods in both the
domains, respectively.

3) Deep Unfolding

Unfolding concepts can be used to unfold a neural network for each iteration of the iterative algorithm into a layered structure and then unite it together to come up with an optimum ANN architecture that can be easily trained for the given problem, e.g., deep unfolding method in [105]. A detector design by unfolding iterative calculations into ANN layers for MIMO decode and forward relay channels has been proposed in [106]. However, in a generalized sense, determining the optimal size (number of neurons and layers) of an ANN for a problem under consideration (with known dimensions) remains an open research problem.

Today’s wireless communication networks are expected to experience a fundamental paradigm shift towards smart and intelligent radio environments [17]. The main question around the role of DL in such communication networks is: whether it will be an integral part of the future networks, but rather it is: when and how to trigger this integration. DL can be seen as an end-to-end solution for replacing the sequential blocks based processing methods for estimation and decoding of information at the receivers.

4) Deep Learning for Cognitive Communications

Cognitive radio techniques enables a radio system to sense, learn, and adapt based on the context of the surrounding environment [87]. The sensing, learning, and adaption may allude to the sensing of radio spectrum, user demands and spatial environment. In this regard, several survey papers including [107, 108] exist, which discuss various aspects of intelligent cognitive radio wireless networks. In [109], supervised and unsupervised ML-based cooperative spectrum sensing algorithms for cognitive radio networks have been proposed. A deep reinforcement learning based power control method for spectrum sharing in cognitive radios has been proposed in [110]. Such dynamic sharing of spectrum aims at enhancing the spectrum utilization efficiency through the provision of access of under-utilized spectrum to the secondary network users [111]. Furthermore, the scope of ML for anomaly and fault diagnosis, intrusion detection and prevention, and network configuration and optimization has been reviewed in [112]. Also, authors in [113] presented the features and advantages of self-organizing networks along with a thorough literature survey, where various methods to improve the efficiency of such networks have been indicated. Moreover, the characterization of learning problems in cognitive radios and the importance of ML in achieving full cognitive networks has been discussed in [114]. In addition, the role of ML for cognitive network management has been investigated in [115], where the realization of ML for automating the management of Fault, Configuration, Accounting, Performance, and Security (FCAPS) has been thoroughly studied.

IV. QUANTUM TECHNOLOGY AND QML-ASSISTED COMMUNICATIONS

QC and ML can create close synergies with each other towards providing their joint benefits in communication systems. The enormous amount of parallelism offered by QC has motivated the start of new disciplines like “Quantum Information Science” and “Quantum Computer Science” [116, 117]. This concept of parallelism comes from quantum Physics concept of qubit, entanglement, and superposition. A qubit can simultaneously hold both the binary states ‘0’ and ‘1’; subsequently, any n interacting qubits can simultaneously represent 2^n unique binary patterns, which is unlike a single binary pattern at-once in the classical computers. These quantum mechanics concepts are well recognized for generating counter-intuitive statistical data patterns which classical computers are unable to produce effectively [41]. The capability of classical ML methods for recognizing statistical data characteristics in the given data, and also for producing data with the same statistical characteristics has also been well established (Classical ML is discussed in Sec. III). The tasks of ML involve manipulation and classification of a large amount of data in the form of large-dimensional vectors, where the required time polynomial is proportional to the data dimensions. The QC has a recognized potential in conveniently manipulating such large-dimensional data vectors in large tensor product spaces. Also, it is envisioned that the combination of QC and ML features together in the framework of QML can produce and recognize the statistical data patterns which classical computers and classical ML are unable to perform effectively. At the initial stage, QML is being defined to exploit QC to accelerate the intelligent data analysis methods. However, in the long run, it is foreseen to lead towards a completely redefined model of ML for quantum computers. This section revolves around the three fundamental questions: “Why quantum communications?”, “What is QML?” and “How QML can contribute to 6G and Beyond communication networks?”.

A. QUANTUM AND QC-ASSISTED COMMUNICATIONS

In this section, we first provide an introduction and the basic principles of quantum communications and then discuss the applications of quantum techniques in various sectors of communication systems. Subsequently, we highlight the potential enablers for quantum communications along with the relevant discussion from the existing literature.

1) Fundamentals of Quantum Communications

Quantum mechanics is expected to play a significant role in various sectors of our everyday life, ranging from high-endurance materials and pharmaceuticals to communications and computing [118]. Any communication or computing device built from the elementary particles is subject to follow the axioms of quantum mechanisms which are usually analogous to the postulates of the Euclidean geometry. In the communications and computing domains, the existing protocols can be enhanced with more efficient algorithms.
by exploiting the physical phenomena available in the quantum world with the utilization of quantum principles and tools. Furthermore, quantum techniques can be significantly useful in investigating computationally efficient solutions to classical signal processing problems. In summary, quantum principles provide significant benefits to the communication networks including enhanced channel capacity, the ability to transmit an unknown quantum state, i.e., quantum teleportation, and to deliver secure information, i.e., quantum cryptography by utilizing a number of advanced communication protocols which will not be possible with the classical techniques [119].

Quantum communication is an emerging branch of telecommunications engineering, which has been motivated from the principles of quantum mechanics and is based on the exchange of quantum states [120]. This novel field of research area aims to utilize the quantum theories/principles to enhance the capacity of future communications systems as well as to incorporate new functionalities. The quantum-assisted communications can enhance various aspects of the existing classical communication networks including channel estimation, optimal Multi-User Detection (MUD), the design of optimal precoding matrix and the optimal routing by employing the quantum algorithms [30]. One important advantage of utilizing quantum domain in communications is high degrees of freedom. By replacing the conventional physical communications channel with the nano-scale objects, i.e., photons, electrons, governed by the quantum principle, in terms of logical values of 0 and 1, it is possible to utilize the linear combinations of these logical values. As an example, for a polarized photon of $P = aP_v + bP_h$, with $P_v$ and $P_h$ denoting the vertical and horizontal polarization, respectively, the values of $a$ and $b$ can be adjusted towards optimizing the communication protocols [118].

Regarding the quantum information sources, a single photon source can be considered as an ideal source for generating quantum information and this can generate pulses having the mean number of pulses equal to one and zero variance [120]. However, the realization of such an ideal photon source in practical quantum communications is very challenging since it occupies very large space and requires trained technicians. In this regard, there are other sources of light which can be used to approximate the ideal photon source, for example, fainted lasers and four-wave mixing process for the fiber-optic quantum communication systems. Out of these, fainted lasers is mostly used in quantum cryptographic key distribution systems while the four-wave mixing process is useful for optical processing devices such as parametric amplifiers and wavelength converters.

For enabling the quantum communications, the information signal can be encoded in different ways such as by modulating the photons’ polarization, which is usually detected by utilizing single-photon detectors, and the photons’ phase, usually measured by the homodyne detection [120]. For the first approach, the polarization is not preserved while transmitting quantum signals via optical fibers and it is necessary to have a non-intrusive polarization control to preserve the information being transmitted in the quantum domain. While the second approach based on the photons’ phase does not require the polarization control but an optical carrier (either propagated along with the quantum signal or generated locally at the receiving side) is needed in order to retrieve the phase information.

Qubit or quantum bit is the quantum version of the classical binary bit and is a fundamental unit of quantum information in QC and communications. It represents a two-level quantum-mechanical system, for example, up and down spins of an electron and vertical and horizontal polarizations of a photon. The state of a Qubit can be represented utilizing any selected orthogonal basis and the most commonly used basis is the computational basis corresponding to the states of $|0\rangle$ and $|1\rangle$ [30]. In this computational basis $\{ |0\rangle, |1\rangle \}$, the quantum state $|q\rangle$ of a Qubit system can be expressed in the following way

$$|q\rangle = a|0\rangle + b|1\rangle, \quad (1)$$

where $a, b \in C$ denote the amplitudes of the quantum state in the considered computational basis and $|a|^2 + |b|^2 = 1$. When $a = 0, b = 1$ and $|q\rangle$ corresponds to the classical bit of 1 while when $a = 0, b = 1$ and $|q\rangle$ corresponds to the classical bit of 0. On the other hand, if $a = b = \frac{1}{\sqrt{2}}$, another state of $|q\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ is obtained which exhibits a symmetry with regard to other states.

For representing the quantum states geometrically, 2-D representation and 3-D representation (Bloch sphere) are utilized for the real-valued and complex-valued amplitudes of the quantum, respectively. Some of the quantum algorithms using only the real-valued amplitudes include the Grover’s Quantum Search Algorithm (QSA), Dürr-Hoyer QSA and Boyer-Brassard-Hoyer-Tapp QSA while some other algorithms including quantum counting algorithm and Shor’s algorithm utilize the complex-valued amplitudes of the quantum states [30]. In this regard, authors in [30] provided the fundamentals of QC by using linear algebra, and then provided a review of existing quantum algorithms along with the applications of quantum principles in wireless communications systems.

Quantum communications aims to utilize the quantum nature of information, thus providing novel challenges and opportunities for designing the 6G and beyond communication protocols. In comparison to the classical binary based communications systems, quantum communications has the great potential to provide absolute randomness and security, to carry much more information and to significantly enhance the transmission quality. Furthermore, quantum-based techniques are able to execute the tasks much faster and beyond the capability of the classical systems [121]. However, quantum communications face mainly two challenges towards designing new communication protocols. The first challenge is regarding the construction of network entities with quantum Internet which requires quantum switches/routers and repeaters, which becomes difficult due to no-cloning theorem.
Another challenge is regarding the capacity measures of quantum communication channels. Although the capacity of classical channels has been well understood within the framework of classical information theory, the capacity of quantum channels is not completely understood and various measures are available in the literature. This is due to the reason that quantum channels can have different possibilities in terms of delivering information including quantum information, entanglement-assisted classical information [123] and private classical information [124].

Furthermore, quantum channel/error correction coding is of significant importance for the practical design of quantum-assisted communication protocols to approach closer to the theoretical achievable capacity. Since the information via the quantum channels is carried in quantum states, the encoding and decoding processes are fundamentally different from the classical encoding and decoding schemes [121]. Another fundamental part of the quantum theory is the measurement which depicts the amount of information which can be gathered about a quantum system. Although the classical meaning of measurement is well understood, its quantum notion has been an important discussion topic and has many variants in the existing literature [125]. One way of interpreting the quantum notion of measurement is that it causes to suddenly collapse or jump into one of the many possible states with some probability. In this regard, authors in [125] have shown that shared randomness is available if necessary, quantum measurements can be asymptotically represented by the amount of classical communication equivalent to the quantum notion of the mutual information of the measurement.

Some of the promising quantum communication protocols to expand the possibility of classical data transmission in quantum-based systems include quantum key distribution (QKD) [126, 127], quantum teleportation [128] and dense coding [129]. Also, like in classical communication networks, quantum networks can utilize frequency and wavelength division multiple access techniques to the problem of channel access in the presence of several users. In addition to these techniques, other multiple access techniques by utilizing the orbital angular momentum of single photons and by using coherent states can also be utilized in quantum communications networks. Furthermore, spread spectrum based multiple access techniques, which can send the photons of multiple users via the medium (optical or free-space) by sharing the frequency band, time window and the route, seem promising in the context of quantum communications [130].

2) Applications of Quantum Communications

Quantum principles can be applied in various sectors of communications ranging from underwater communications and terrestrial wireless networks to the satellite networks. One of the widely-discussed application areas of quantum communications is optical fiber communications in which the conventional approach is based on the classical electromagnetic fields and may suffer from the undesired fluctuations. Also, noise having the quantum-mechanical origin may limit the performance of photodetectors. To address these issues, optical communication systems can be designed under the quantum-mechanical framework [131].

Another promising application area is to enhance the security using quantum communication protocols in aquatic scenarios due to the increasing number of vehicles sailing on the ocean surface. In this regard, authors in [132] carried out the feasibility analysis of Quantum Key Distribution (QKD) protocols in the aquatic scenarios and showed the significance of employing QKD protocols in the underwater environment.

In addition, Satellite Communications (SatCom) is another important area where quantum techniques can be employed for various purposes. For example, authors in [126] discussed and analyzed the applicability of QKD protocols in quantum-assisted SatCom systems in order to perform secure communication between ground stations and the satellite. Also, another promising application area is the quantum Internet, which enables the transmission of Qubits from one quantum computer to another [119]. In addition, another application area of quantum techniques could be TeraHertz (THz) communication system which is recently being investigated in the research community. To this end, authors in [133] discussed the properties of THz frequency bands and the essential conditions for the application of quantum communications in this frequency band.

Another important application of quantum communications is quantum teleportation, which utilizes the quantum entanglement principle to transfer a particular quantum state to another place with the quantum devices by using the classical bits rather than the quantum bits [128]. The main challenge in employing quantum teleportation in wireless systems is that EPR (named after Einstein, Podolsky, and Rosen) pairs, i.e., entangled pairs of qubits cannot be set up and shared instantaneously in wireless quantum devices since EPR pairs cannot be distributed to the quantum devices via the air. This leads to the need of designing a new quantum mechanism which is capable of performing teleportation from one site to another without requiring to have a mutual exchange of EPR pairs between the sites. To address this, a novel approach of quantum routing mechanisms by executing quantum circuits in parallel at the intermediate nodes has been recently proposed in [134] and it has been shown that the proposed quantum routing approach is independent of the number of routing hops and is closer to the optimum in terms of time taken to teleport a quantum state.

In terms of practical implementation, a quantum annealing chipset is already commercially available from the company D-Wave [135]. Also, due to the recent developments in the quantum stabilizer codes towards mitigating the decoherence effects in quantum circuits, gate-based architecture which comprises of computational blocks with the quantum gates has attracted significant attention. Furthermore, D-Wave 2000Q3 having a total of 2000 qubits and IBM Q Experience4 with a total of 20 qubits are already available.
and IBM has a recent plan of finalizing a gate-based quantum computer with 50 qubits by 2020 [30].

3) Potential Enablers for Quantum Communications

This section discusses the potential enablers and related critical issues of quantum communications.

a: Quantum Entanglement

One of the issues in quantum communication is the effective transmission of information over a noisy quantum channel and there are several attempts in the literature to characterize the achievable rate of transmitting classical and quantum information over a noisy quantum channel. For example, the achievable rate for the transmission of classical information over a noisy quantum channel is given by the Holevo-Schumacher-Westmorel (HSW) coding theorem [136], which generalizes the Shannon’s theorem in quantum settings. Also, regarding the transmission of quantum data over a quantum channel, the achievable rate is given by the Lloyd-Shor-Devetak (LSD) coding theorem [124, 137]. Subsequently, the article [138] investigated the case where both the classical and quantum information can be simultaneously transmitted over a quantum channel by employing a time-sharing strategy.

Authors in [139] investigated the tradeoffs for channel coding both quantum and classical information over a noiseless entanglement-assisted quantum channel and proved that the proposed entanglement-assisted classical and quantum capacity theorem provides the achievable rates in the considered scenario. In addition to this quantum entangled based communication [139, 140], there are recent attempts in developing quantum entanglement-assisted quantum turbo codes [141] and the squashed entanglement of a quantum channel, which is an additive function of a tensor product of any two quantum channels [142].

b: Quantum-dot Cellular Automata (QCA)

One of the main issues with the Complementary Metal-Oxide Semiconductor (CMOS) technology is the physical limitation in terms of the feature sizing [143]. To address this issue, Quantum-dot Cellular Automata (QCA) seems to be a promising enabler, which is a nano-scale computing mechanism and serves as a basis for binary computation which has fundamental differences from the current transistor technology [144, 145]. In other words, QCA utilizes cells of quantum dots to store and transfer information, with each cell comprising of four quantum dots structured at the corners of a square. In this direction, several theoretical and modelling work related to QCA are already available in the literature [143, 146–148].

The term “Quantum-dot” in QCA represents the portion of matter, i.e., semiconductor, whose excitons are concentrated in three spatial dimensions and its electrical properties are in between those of the discrete molecules and of bulk semiconductors. On the other hand, the term “Cellular Automata (CA)” represents the dynamical systems having discrete space and time, and also can be considered dynamical systems with the infinite dimension [145]. The CAs define the mathematical models for the systems in which several simple components interact with each other to generate the complicated behavior patterns.

c: Quantum Hardware Capacity

One of the crucial issues for the application of quantum technology in communications related applications is the presence of harmful quantum perturbations, whose harmful effects can be mitigated by employing Quantum Error Correction Codes (QECCs) [141]. The performance of QECCs can be enhanced by employing the entanglement assistance in the context of a symmetric depolarizing channel [141]. In this regard, authors in [32] have provided a detailed analysis on the capacity of an entanglement-assisted quantum channel while considering the realistic quantum devices, and also provided an EXtrinsic Information Transfer (EXIT) chart-based design methodology for the QECCs to enhance their performance in asymmetric quantum channels. With the help of simulation results, it has been demonstrated that the proposed EXIT chart based techniques are useful tools to analyze and design quantum coding schemes.

In the above context, authors in [149] provided a comprehensive survey on the recent development of quantum-like models which can better represent various factors involved in the human decision-making process, namely, ambiguity, uncertainty, emotions and risks. Furthermore, the article [150] developed a QDT based approach for quantitative predictions in the arbitrary scenarios including the ones where the utility theory fails. In contrast to the previous quantum-like models which mainly utilize several fitting parameters for the construction of some models to describe particular effects of a use-case, the QDT model proposed in [150] is considered as a generic theory applicable to any type of decision making, and its mathematical structure is common to both the decision theory and quantum measurements. The proposed QDT model is based on the generalization of the von Neumann theory [151] of quantum measurements to the non-conclusive measurements and the composite events comprised of noncommutative operators.

d: Quantum Key Distribution

The crucial problem in the traditional Vernam one-time pad cryptosystem is to deliver a secret key to two legitimate parties. This issue can be addressed by the QKD, also called the quantum cryptography, which provides a secret key to two legitimate parties in a Vernam one-time pad cryptosystem [127]. In this QKD approach, quantum mechanism provides the unconditional guarantee of the security of the key. There are several QKD protocols available in the literature, namely, BB84, E91, B92 and BBM92, of which BB84 is the most popular and widely used QKD scheme.

The QKD can be used to enhance security in various networks including optical networks, terrestrial wireless networks and satellite networks. Recently, authors in [126]
investigated the application of QKD in the satellite communication system to perform secure quantum communication between ground stations and the satellite. The performance of QKD in satellite networks gets degraded in the presence of high attenuation caused due to noise and atmospheric effects. To address this issue, suitable quantum error correction methods can be employed. Furthermore, the article [152] analyzed the feasibility of trust-free long-haul QKD in future quantum communication networks by combining the measurement device-independent QKD and a quantum repeater, which is considered as one of the key ingredients of trust-free networks.

e: Quantum Decision Theory (QDT)
The classical decision-making process is mostly based on the expected utility theory and its performance significantly degrades in the scenarios having the risk and uncertainty [150]. In most of the classical decision-making process, the possibility of making correct predictions can be strongly affected by the nature of the surrounding environment such as the unknown stochastic or varying environment. Furthermore, in the scenarios having incomplete or partially reliable information or incomplete preference relations, any prediction is likely to be just partial and qualitative. To address this, Quantum Decision Theory (QDT) seems to be a promising approach and has been already investigated in some existing literature [149,150]. Also, the process of representing all steps of a decision process mathematically in order to allow quantitative prediction is significantly important not only for the decision theory but also for developing artificial quantum intelligence, which can work only for the operations defined in mathematical terms [153].

f: Quantum Game Theory (QGT)
With the recent advances in quantum information and quantum computation, there has been a trend of formulating classical game theory using quantum probability amplitudes towards analyzing the impact of quantum superposition, entanglement and interference on the agents’ optimal strategies [154]. The Quantum Game Theory (QGT) in general replaces the classical probabilities of game theory by quantum amplitudes by creating the possibility of new effects coming from entanglement or superposition. The main difference between the classical game and the quantum game is that classical games perform calculations in the probability space whereas quantum games operate in the Hilbert space.

Quantum game theoretic techniques can be utilized in investigating suitable solutions in quantum communication [155] and quantum information processing [156]. In this regard, the article [154] provided an introduction to the quantum theory along with some related works and discussed some well-known quantum games including quantum penny flip, Eisert’s quantum prisoners’ dilemma and quantum Parrondo’s games. Furthermore, the recent article [157] analyzed the existing works on quantum games from three perspectives, namely, co-authorship, co-occurrence and co-citation, and also reviewed main quantum game models and applications.

g: Quantum-proof Randomness Extractors
For several applications in computation, information theory and cryptography, randomness is a fundamental aspect and the objective of randomness extraction is to transform the sources of correlated and biased bits into nearly uniform bits [158]. The extractors which can work in the presence of quantum side information are quantum-proof, and also the extractors are with one bit output are regarded as the quantum-proof [159].

Quantum-proof randomness extractors can be considered as an important building block for implementing classical and quantum cryptography in security applications [160]. Mainly, the randomness extractors setting of this block provides a nice framework to study the capability and limitations of a quantum memory over the classical one. The study on the behavior of randomness extractors in the scenarios with quantum adversaries can be based on the theory of operator spaces, which is also known as quantized functional analysis. The extractors in general approximately map a weakly random system into uniform random bits by utilizing the perfectly random bits, called the seed. There exists one interesting generalization of extractors, called condensers, which is considered as an intermediate step towards building the extractors [161].

4) Notable Applications of QC-Assisted Communications
This section surveys a few notable recent application examples of QC-assisted communications.

a: Quantum-assisted Multi-User Detection (QMUD)
The practical implementation of classical optimal classical detectors such as Maximum Likelihood (ML) MUD is often limited by their very high implementation complexity. To address this, one of the promising approaches could be Quantum-assisted MUD (QMUD) [162]. With the recent advances in quantum cryptography and quantum error correction, there have been substantial research efforts towards investigating the feasibility of QMUDs. In this regard, the article [162] presented a comprehensive review and tutorial on quantum search algorithms and their applications. Furthermore, an ML QMUD was proposed by considering the legitimate combinations of the users’ transmitted symbols at the receiver and it was shown that the performance of the proposed ML QMUD matches to that of the classical QMUD.

b: Quantum-Aided Multi-User Transmission
In addition, the QSA can be utilized in reducing the complexity of vector perturbation precoding and enhancing the performance of multi-user transmission in wireless networks. In this regard, authors in [38] proposed quantum-assisted Particle Swarm Optimization (PSO) algorithms in both the discrete and continuous modes with the objective of performing vector perturbation precoding and reducing
transmission power at the BS if a rank-deficient multi-user system while minimising the average BER at the mobile users. Via numerical results, it was shown that quantum-assisted precoding provides better BER performance as compared to the conventional PSO algorithm, while keeping the same computational complexity. Also, the superiority of Quantum-assisted precoder over the classical precoder has been illustrated in the scenarios having limited feedback of CSI from the users to the BS. In this regard, low-complexity soft-output quantum-assisted MUD has been investigated in various settings by considering different multi access schemes including Space Division Multiple Access (SDMA), Orthogonal Frequency Division Multiple Access (OFDMA), Code Division Multiple Access (CDMA) and Interleave Division Multiple Access (IDMA) [163–165].

In a rank deficient multiple-access system in which the number of users is higher than the number of receive antenna elements at the BS, low-complexity heuristic MUD does not provide the desired performance. Furthermore, the complexity of optimal Maximum A posteriori Probability (MAP) MUD increases exponentially with the number of users and the number of bits per transmit symbol. To address this, authors in [165] employed quantum search assisted MUD to reduce the search space and with this soft-input soft-output MUD approach, only a fixed subset of the best multi-level symbols having a near optimal cost function needs to be evaluated to achieve the near-optimal bit error rate performance. Subsequently, the EXIT chart was utilized to design the proposed QMUD assuming the Gaussian distribution of the MUD’s output and the performance was evaluated for multi-carrier interleave-division multiple-access systems. Furthermore, another article [166] exploited the advantages of QMUD in the uplink of a multi-user system by considering the transmission of a vide stream from a reference user to the BS. The employed QMUD detects the signals transmitted by all the users instead of considering other users’ signals as interference.

c: Quantum-assisted Indoor Localization for mmWave and VLC

There is a recent trend of employing mmWave and Visible Light Communications (VLC) technologies in indoor localization applications. One of the main issues with these technologies in practical applications is to achieve the desired localization accuracy. Also, it may not be possible to utilize the triangulation approach due to the limitations in the infrastructure and scenarios [35]. Although fingerprinting based localization method could be employed in both the Radio Frequency (RF)-based and VLC-based applications, the complexity of searching the fingerprinting database can be expensive for the scenarios requiring high accuracy. One of the promising approaches to address this complexity reduction issue is to employ a QSA, which aims to find the minimum entry in the unsorted database with \( N \) elements by utilizing only the \( O(\sqrt{N}) \) Cost Function Evaluations (CFE).

In this regard, authors in [35] showed the possibilities of utilizing QSA for reducing the computational complexity of mm-Wave based and VLC-based localization algorithms while achieving the same performance as that of a full search.

d: Quantum-assisted Joint Routing and Load Balancing

One of the crucial challenges in wireless networks involving mobile networking devices such as smartphones and tablets is to optimize the routing of message flow in order to maximize the utilization of bandwidth and power. One of the issues in achieving this is nodes’ social selfishness, which makes nodes to choose certain paths for optimizing specific utility but without considering the impact on the degradation of the overall network’s performance [167]. This may lead to the creation of the bottlenecks in the network flow, and to address this issue, the design of socially-aware load balancing may be significantly useful, and it is important to consider nodes’ user-centric social behavior in addition to the conventional conflicting objectives such as power consumption and path delay. In this context, a multi-objective optimization approach can be utilized based on the socially-aware Pareto optimal routing, however, finding the set of Pareto-optimal solutions has huge complexity since the problem is usually NP-hard. The recently emerging quantum technologies including quantum computation [168] and quantum information processing can significantly reduce the complexity of finding Pareto-optimal solutions by utilizing the concept of Quantum Parallelism (QP). As compared to the Hardware parallelism (HP) for complexity reduction (which provides complexity reduction in the order of \( O(K) \), \( K \) being the number of independent parallel processes), the QP can achieve the complexity reduction in the order of \( O(\sqrt{N}) \), where \( N \) being the database length. In this regard, authors in [36] employed a multi-objective decomposition quantum optimization algorithm for the joint optimization of routing and load balancing in socially-aware networks.

e: Quantum-assisted Channel Estimation and Detection

The performance enhancement of MIMO-OFDM system with the joint channel estimation and MUD has been depicted in several existing works [169, 170]. In this joint channel estimation and MUD process, QC can play a significant role due to its inherent parallelism for reducing the complexity, and for enhancing the estimation as well as detection performance [37]. In this regard, authors in [37] proposed a quantum-aid repeated weighted boosting algorithm for the channel estimation and employed in the uplink of MIMO-OFDM systems along with a MAP MUD and a near-optimal QMUD. The performance of the proposed quantum-based scheme was shown to be superior to that of the conventional repeated weighted boosting algorithm, and also the impact of channel impulse response prediction filters, Doppler frequency and power delay profile of the channels were analyzed.
B. FUNDAMENTALS OF QUANTUM MACHINE LEARNING

This section revisits the ML methods discussed in Sec. III-A in the context of quantum-assisted algorithms for ML and the QML framework.

1) Overview of Quantum Learning Methods

Quantum principles based on emerging computing technologies will bring entirely new modes of information processing. An overview of supervised, unsupervised, and reinforcement learning methods for QML is discussed in the sequel.

a: Supervised and Unsupervised QML

As elaborated in Sec. III-A, supervised learning infers the required functionality from the given labeled training data; while unsupervised learning attempts to find hidden patterns and structures in the given unlabeled data. In different ML tasks, the scale of speed-ups achieved by QML over the classical ML algorithms transpires in different fashions. For various learning tasks, QML algorithms can provide exponential speed-ups over classical ML algorithms [42], involving large dimensional data, in both supervised and unsupervised learning approaches. Alongside, QML can also enhance security and privacy in communication networks. In this regard, supervised and unsupervised learning for clustering and classification tasks have been thoroughly explored in [42] and quantum improvements in supervised and unsupervised learning have been reported in [43]. Also, in [171], fundamental learning concepts and the applications of QML have been comprehensively discussed, where some discussed notable QML tasks are quantum pattern recognition, quantum classification, quantum process tomography and regression, boosting QC, and adiabatic QC. Furthermore, training, model selection, and error estimation aspects of QC powered supervised ML have been thoroughly discussed in [172].

Unsupervised QML algorithms have been discussed in [173], where the process of (partially or totally) converting a classical algorithm to its quantum counterpart has been explained. A distributed setting based k-medians clustering method for cost-efficient communication protocols has also been described [173], which estimates the sum of distances instead of simple sequential additions. Moreover, quantum algorithms for neighborhood graph, outlier detection, and smart initialization of cluster center have been proposed.

b: Quantum Reinforcement Learning

Reinforcement learning is an interactive and generalized form of learning. As discussed in Sec. III-A3, an agent learns the required optimal behavior through reinforced rewards and penalties from the environment. Quantum-speedup for reinforced learning is an emerging framework with a strong potential in the agent-environment paradigm. The interactive setting of two-parties (agent and environment) can conveniently be extended for a quantum information treatment. The superposition and parallelism concepts of quantum mechanics can be used to represent and identify the eigenstates in quantum-powered reinforcement learning, by observing a random quantum state simulated through the collapse postulate of quantum measurement. The reward from the environment can be used to update the probability of Eigen actions in a parallel fashion. The probability of the Eigen action is determined by the probability amplitude, which is parallelly updated according to rewards. In [174], a quantum value updating algorithm for quantum-powered reinforcement learning has been proposed.

Some fundamental characterization (based on convergence, balancing, and optimality) to study performance tradeoffs between exploration and exploitation of quantum parallelism for speeding up reinforcement learning is conducted. Advances in quantum-powered reinforcement learning have been discussed in [175]; where a solution to the bottleneck of required oracularized variants of task environments, has also been proposed. In the context of communication systems, quantum inspired reinforcement learning method for optimal spectrum assignment has been discussed in [176].

2) Generative and Discriminative QML Models

In generative models, the actual distribution of each class is learned, while the conditional probability distributions are predicted through different transformational theorems. In the discriminative model, the focus is to learn the decision boundaries between the classes by modeling the conditional probability distributions. A tensor network inspired QC approach to both discriminative and generative learning models has been proposed in [177], where a solution to the near-future quantum devices with a limited number of physical qubits and high error-rate are targeted. A significant amount of today’s learning ideology is based on generative models. In [178], a quantum generative model based QML algorithm has been proposed. It has been demonstrated that the representation of probability distributions in the proposed quantum generative model compared to the classical generative model has exponential speedup in learning and inference. A quantum restricted Boltzmann machine network algorithm for unsupervised generative models has been proposed in [179]; where generative models are shown to outperform discriminative models in terms of classification performance. Moreover, the construction aspects of the algorithm for quantum circuits and computers are discussed. Writing a good quantum algorithm may be a challenging task at the initial stage with a limited hold on the basic knowledge, ML can offer the trick of the trade by learning the quantum algorithms.

3) Quantum SVMS and ANNS

The origin of the motivation of proceeding towards Quantum Neural Networks (QNNs) is in the essential quantum manipulations happening in a living brain, and in the exploitation of advancements in both QC and ANNs. Outspanning the fundamental concepts of quantum information processing and
Among various difficulties in realizing the QNN, the maintenance of coherence during quantum parallel distributed processing and implementation of interconnections (massive) between neurons in the form of entanglement of qubits, are the most notable. There also exist literature on the realization of physical systems for QNNs, see e.g., [183], which include nuclear magnetic resonance and quantum dots.

Based on the known unprecedented potential of QC in solving problems beyond the conceivable reach of classical computing methods, QNNs can be seen as outperforming the classical ANNs at-least at a similar rate. A QC based perceptron learning model has been proposed in [184], which overcomes one of the major obstacles in advancing the growth of QNNs, i.e., ANNs being a nonlinear function. Perceptron is a fundamental building block of ANNs (and SVMs), with known tight performance bounds on computational and statistical complexity of perceptron training. This rigorously enables the clear demonstration of any improvements achieved through any advancements. In this context, in [184], the QC perceptron has been shown to achieve non-trivial improvements in the computational and statistical complexity of the learning model.

The implementation of Quantum SVMs (QSVMs) has been presented in [185], where an exponential speed-up of QSVMs over classical SVMs has been reported. The core concept in offering this improvement of the QSVM framework is the use of an efficient matrix inversion operation required for computing training data inner-product (kernel) matrix, which exploits a non-sparse matrix exponentiation technique. Furthermore, a Quantum Sparse SVM (QSSVM) for minimizing $\ell_1$-norm of feature weights vectors has been proposed in [186]. Moreover, sparse structured vectors are encountered in various applications of wireless communication systems, e.g., the mmWave propagation channels are usually sparse in angular domain [22] and underwater acoustic communication (UWAC) channels are usually sparse in delay domain [84]. In this regard, a sparse representation approach for wireless communication systems has been discussed in [187]. In this context, the sparse representation of features in wireless communication systems can be a potential application of QSSVMs.

The advantages of QNNs and QSVMs over classical ANNs and SVMs in terms of processing speed, faster learning, smaller scale, scalability, and reliability motivates the exploration of these methods in resolving many diverse problems in wireless communication networks (e.g., resources optimization, nodes coordination, estimation of parameters, etc.). Towards this direction, the advancements in QC and ML methods in the last decade has now opened new horizons of realizing Deep QML methods, e.g., Deep QNNs (DQNNs) [180].

4) Quantum Deep Learning

Quantum-assisted DL is receiving significant attention towards enhancing various performance metrics of communication networks. The classical DL faces various challenges; where a substantial challenge is to figure out the training method for complex topologies of ANNs (which are of similar complexity to that of the natural structure of the human brain). Automatically conceiving the optimal size and topology of an ANN for the problem under consideration is another research challenge of classical DL. The derivation of DL models for simulating complex neural topologies and data flow mechanisms is not naturally supported by the classical computing architectures. In this regard, QC based algorithms for DL are envisioned to have a profound impact on the evolution of ML methods. Quantum DL algorithms can not only outperform conventional learning algorithms in terms of processing time but also in terms of enrichment in modeling quality.

An example of very ambitious deep, wide, and complex ANN, holding a balance between underfitting and overfitting, operating at evolved parallel processing framework, is illustrated in Fig. 3. Supervised ML is generally prone to overfitting, which is defined as a situation when the model produces a good- and bad-fitting for training and unseen data, respectively. Optimum setting of ANNs (i.e., structure, size, memorizing capacity, etc) to hold a balance between overfitting and underfitting is a critical requirement.

The emerging concepts of QNN and DNN can be postulated together to formulate DQNNs. There is a remarkable scope for conducting research work on this modern concept of deep QML for devising methods for clustering, classification, recognition, optimization, estimation, and other AI operations by exploiting its magnificent capability of quickly modeling several layers of abstraction in the given raw data. Recognizing this scope, the research community has already started advancing the classical training algorithm for deep QML. In this regard, authors in [188] proposed a quantum algorithm for training a DQNN. Another quantum algorithm for training a deep restricted Boltzmann machine has been presented in [189]. More importantly, the crossover between QC, DNN, and information processing is an exciting interdisciplinary area stimulating progress in all the three disciplines. This observation quickly suggests that DQNNs have a tremendous scope of meeting the requirements of bringing full intelligence to individual nodes, a swarm of nodes, and a network of nodes in 6G and beyond communications.

5) Parallelization, Scalability, and Generalization

In the emerging era of a data-centric world, the massive amount of available data will require the innovative robust provisions of storing, processing, and analyzing the data. A crucial concern for the QML framework will be the improvisation of storing and processing capabilities for enabling advanced data analytics through effective handling of the available unstructured data. The provision of data security in such a framework will be another crucial issue [190]. A simple solution to these problems can be the provision of data storing and processing capabilities in parallel, distributed, and batch fashions. To this end, a few promising platforms...
FIGURE 3: An example of deep, wide, and complex artificial neural network structure. Evolution from sequential to parallel data processing and optimum balance between overfitting and underfitting is illustrated.

can be named as Hadoop, Spark, Flink, Beam etc [191]. QML algorithms can play a crucial role because of their abilities to process information through quantum superposition, which can significantly speed-up the storage and computations with the assurance of high data security. In the context of analyzing the huge amount of data through QML, it is crucial to characterize the data structuring methods for effective representation of data in quantum superposition framework. Despite having the privileges of available arrangements for interfacing the classical memory units, it is highly desirable to devise advanced QML mechanism to process/store the data in parallel, distributed, and batch fashions as quantum-data.

To implement a fully scaleable computational device, the existing technology needs the ability to maintain quantum coherence among the qubits in a scalable fashion with very high certainty. Superconducting qubits and ion traps are the popular forms of the current architecture for quantum computation, which are arrays of interacting qubits that are continuously controlled via external pulses to implement the desired operations [192]. However, this approach will face scalability issues even if they are highly capable of maintaining quantum coherence for a longer duration. More sophisticated classical control units are required to develop the unmodulated quantum devices, which are capable of implementing the desired algorithm. More specifically, fault-tolerant quantum computers are required to solve harder problems. Better hardware devices and methods will be developed in the near future to implement quantum error correction using relatively small-scale experiments with quantum error-correcting codes [193]. An error-corrected qubit with more enhanced control is believed to be sooner or later available. With the higher number of physical qubits, the fault-tolerant QC system will be able to efficiently solve the classical hard problems. However, this may take some time to develop such fault-tolerant quantum devices. These fault-tolerant quantum devices will go far in the context of computations through the execution of large circuits having more accurate quantum gates. In summary, significant advancements are required for hypothesizing the new insights and innovations to have fully parallelized, scalable, and generalized quantum algorithms and devices.

C. CHALLENGES IN ENABLING QUANTUM AND QML-ASSISTED COMMUNICATIONS

This section briefly discusses the open research problems in the development of quantum communications, quantum computers, and QML. A few challenges, which require attentions of the research community in enabling the timely provision of QC facilities at the edge and cloud of 6G and beyond communication networks, are highlighted. The development of highly consistent and controllable qubits and quantum logical operations is a fundamental need in the realization of reliable large-scale quantum computers, where quantum error correction methods can be used in bringing improvements in system reliability. Provision of a high degree of precision and sensitivity in quantum devices (sensors, measurement, etc) that enables the full exploitation of quantum entanglement concepts is of vital importance.

An important milestone in realizing quantum communications and quantum internet is the development of long-distance quantum communication channels. Long-distance quantum communication can be suitably realised through the physical platform of photons, where an open research problem is the loss of photons in quantum channels. Use of
repeaters, in principle, can overcome this drawback, through subdividing the large distances into small sections which are suitable for entanglement to be teleported. Moreover, to implement such quantum repeaters, the decoherence effects imposed by the quantum channels need to be dealt with. Development of transducers to photonic states can help in traveling long distances with minimal decoherence. These transducers also have another research potential application in interconnecting different leading physical platforms, viz: superconducting circuits, ultra cold atoms, spins in semiconductors, and trapped-ions. A single-photon quantum device has already been realized, however it currently operates at low temperatures. The advancements may be happening very quickly; however, to make the quantum devices operational at normal (practical) temperatures, a lot of dedicated effort is needed.

Looking back at the evolution of different generations of communication networks, it can be observed that the development cycle of each generation typically takes a decade. Preceded by this time frame, for the development of 6G, the resolution of challenges and limitations in the provision of large-scale reliable quantum devices need dedicated efforts from the research community. The capability of various physical quantum platforms (e.g., superconducting and trapped-ions) in realizing multiple qubits together for performing quantum logic operations with high reliability is well established now. In 2016, it was envisioned in [194] that the short- and long-term goals for next 5 and 10 years are to realize quantum computers with up to 100 and 1000 qubits, respectively. Recently, Google has announced a 72 qubit superconducting quantum computer [195].

The quantum computers simulators available today can only simulate a small number of circuits, i.e., very limited offered capacity [196]. This is because the simulation of a quantum computer on a classical computer is a computationally hard problem. Such simulators require an exponential amount of operations to model the exponential behavior of quantum systems on classical computers. Parallelization can partially facilitate the resolution by allowing the simulation of more qubits in less time. To this end, the concept of grid computing may further assist in realizing the ambitions by conceding the coordinated resource sharing and access to dynamic multi-institutional virtual organizations [197]. In order to make the QML for wireless communications a reality, an expedition in research work on QML base communications can be achieved through the provision of such classical grid computing facilities assisted commercially available simulators of quantum devices to the research community. This will facilitate the development of novel QML algorithms in parallel to the development of quantum computers.

There are numerous other challenges and open research problems in the fields of quantum communications, quantum computers, and QML; which requires a separate dedicated article to thoroughly survey and review all of them.

V. PROPOSED FRAMEWORK FOR 6G NETWORKS AND FUTURE RESEARCH DIRECTIONS

The 5G networks have now entered into the commercialization phase, which makes it rational to launch a strong effort to draw future vision of the next generation of wireless networks. The increasing size, complexity, services, and performance demands of the communication networks necessitate a deliberation for envisioning new technologies for enabling and harmonizing the future heterogeneous networks. An overwhelming interest in AI methods is seen in recent years, which has motivated the provision of essential intelligence to 5G networks. However, this provision is only limited to perform different isolated tasks of optimization, control, and management nature. The recent success of quantum-assisted and data-driven learning methods in communication networks (discussed in previous sections) has a clear motivation to consider these as enablers of future heterogeneous networks. This section proposes a novel framework for 6G networks, where quantum-assisted ML and QML are proposed as the core enablers along with some promising communication technology innovations. An illustration of the proposed framework is presented in Fig. 4, which indicates various emerging technologies, complex and heterogeneous network structure, multi-space massive connectivity, and a wide range of available big data across different layers, sides, and applications are indicated. The discussion on key thrust areas of future research in the context of the proposed framework is categorized into “Network-Infrastructure and -Edge” and “Air Interface and User-End” sections as detailed in the following.

A. NETWORK-INFRASTRUCTURE AND -EDGE

An extension of the conventional land-mobile radio cellular communication networks to the multi-space highly-mobile radio-to-optic services-oriented cell-free communication networks is suggested. To meet the increasingly stringent performance demands, extending network connectivity to everyone and everywhere is envisioned. Such integration includes a wide range of communication applications across multi-dimensional physical space, e.g., underwater (sensors, submarines, etc), ocean (sensors, ships, etc), land (indoor and outdoor users, Massive-IoT (M-IoT) devices, inter- and intra-vehicle, etc), air (UAVs, drones, aeroplanes, High Altitude Platforms (HAPs), etc), space (satellites, space shuttles, space mission robots, etc), human body (in-body sensors, brain interface, etc). An example of the left horizons for the provision of high-performance all-time network connectivity is airborne internet access. In a traveling friendly smart World of the future, the passengers traveling across remote (oceanic) regions in ships and aeroplanes will also demand the same provisions of network services which are available to the land/home users. For enabling harmony across such massively connected complex 6G networks operating in the co-existence of its predecessor, a tremendous learning and processing capability will be required. Various important research directions for enabling intelligent operations at
network-infrastructure and network-edge in 6G networks are discussed in the following subsections.

1) Intelligent Proactive Caching and Mobile Edge Computing

Intelligent proactive caching refers to the concept of buffering the data at the nodes (IoT devices, BSs, etc) intelligently on the basis of their popularity/demand-rate. This concept helps in reducing the delay and power consumption in data routing and delivery, and it offers a significant performance improvement for all type of users; e.g., the smartphones designed for previous generations also equally benefit from this. Providing intelligence to the nodes to smartly grade the popularity of contents has received an overwhelming response, see e.g., ML advised enhanced caching in [198]. Also, the DL for processing, classifying, and manipulating contents to compute their importance for enabling proactive caching at nodes/edge has also been actively studied in the literature, see e.g., DL based caching method in [199, 200]. However, enabling this concept of proactive caching requires the processing of a very large amount of data to...
evaluate/estimate the media/content popularity. In the context of big data processing, QC to accelerate the content/media processing has a potential research application in proactive caching [30]. Independently and jointly investigating the scope of QC and ML for proactive caching in the emerging big-data era are the potential future research directions.

The demands of mobile users tend to exhibit a predictable pattern in media/data interests and data patterns. Intelligently caching the data at serving stations (e.g., BSs) in the proximity of mobile users can enable offloading heavy traffic from the network backhaul and reducing the network latency for popular contents, through an instantaneous service from the network edge. This promising and emerging paradigm of MEC has also received a joint interest with multiple-access methods, referred to as Multi-Access Edge Computing, leveraging real-time access to the radio network. This enables new possibilities to jointly optimize the radio resources and data network performance features. Enabling these interesting concepts to demand the provision of intelligence and strong computational capability at the network edge, which can be foreseen in the shape of QC, ML, and QML in the future.

2) Multi-Objective Optimization and Routing Optimization

Various diverse type of data analysis tasks involves optimization of objectives given tweaking parameters and their constraints. The optimization problems can be classified into various types based on the nature of the objective and/or penalty functions, amount of objectives, and equality/inequality constraints. QML is observed to exponentially speed-up the optimization problems involving quadratic objective functions subject to equality constraints and involving penalty functions subject to inequality constraints [41]. In this regard, a quantum approximate optimization algorithm based on alternating qubit rotations for penalty function problems has been proposed in [201]. Furthermore, optimization through QML is not only a subfield of QC and ML but it is increasingly emerging to redefine QC in the context of software design, hardware development, and their applications.

The multi-objective approach for efficient routing of data packets in communication networks (e.g., M-IoT), wireless sensor networks (WSN), mesh-networks in airborne internet access etc) is another emerging paradigm. This refers to the optimization of multiple objectives/performance-metrics in a routing problem (e.g, optimization of average delay and expected transmission count etc). Both ML and QC has been independently considered for this computationally-tedious and intelligence-needy task of searching in multiple spaces with multiple tight constraints (battery resources etc) to draw a global optimal packet-route in order to optimize the overall network performance, see e.g. in [202] and [33], respectively. To this end, a QC-assisted routing optimization algorithm, named as nondominated quantum optimization algorithm, for self-organizing networks has been proposed in [168]. Moreover, the framework of DL has a recognized potential in multi-objective optimization, see e.g., [203]. This research topic is directly related to various other interesting research topics; e.g., intelligent proactive caching.

In the application scenarios of mobile mesh networks like airborne and oceanic broadband etc, the remote (oceanic regions) flying aircrafts (or sailing ships) cannot always be served from the optical fiber supported ground stations, while the satellite links are expensive and have high latency, the solution lies in the mixture of proactive caching, transfer caching, multi-objective routing, and deep (transfer) learning. Quantum-assisted ML and QML can be seen as the enablers for solving massive-objectives optimization tasks of the massively-complex communication networks of the future, e.g., massive-objective routing in M-IoT for enabling smart world.

3) Massive-IoT and Big Data Analytics: Realizing Smart Green World

The concept of future smart, intelligent, and green cities, aims at offering manifold new people-centered services for enhancing the quality of people’s life. Realization of this concept is only possible through the use of the latest technologies. In this regard, IoT and AI (e.g., ML) are being considered as the core smart-cities enabling technologies. The IoT concept involves the extension of network connectivity to a plethora of devices provided with sensing, detecting, actuating, data mining, and analyzing capabilities. Such devices may include sensor nodes, cameras, sensors installed vehicles (private/public), road traffic monitoring systems (visual and sensor-based), UAVs, fire/earthquake/other sensors (disasters alerting and monitoring), user smartphones, etc.

Leveraging intelligence to the IoT devices originates the new frontier of “IoT meets AI” [27]. This new frontier has also already cast a significant impact in defining and characterizing the concept of future smart cities (covering all aspects from planning to overhauling the city services). Among many, extending the provisions for enabling smart cities, is an important object of emerging 5G networks. In many different shapes, various provisions of people-centered services have already emerged, e.g, health care, home utility management, city transportation network management, city alert, and rescue services management (fire, flood situations etc). This has been possible through various separate platforms, e.g., increased number of cameras in a city has enabled safe-city concept, sensing capability in the smartphones has enabled services like healthcare, etc. However, a composite concept of a green smart city with everything-a-service, at a low-cost, to improve the quality of life of everyone-in-the-city may be fully achieved in 6G and beyond communications.

In the concept of fully integrated smart cities in 6G and beyond communications, the voluminous amount of data instantaneously produced from the massive amount of IoT devices (almost everything connected to the network) can only be effectively utilized to provide instantaneous (runtime) services if a very high computational capability is leveraged to the system. Moreover, this high computational load at IoT devices will make them more power hungry. QC, energy harvesting, energy efficient routing (ML optimized),
and wireless power transfer concepts come to the rescue by offering a massive capacity of accelerating the processing speed and power requirements of future IoT devices. The types of learning methods suitable for different tasks along with the available data at different layers of communication systems is presented in Table 2. The quantum speed-ups for different tasks (e.g., classification, learning, etc) can help in realizing various diverse types of applications of big data analytics, see e.g., [204]. This motivates the exploitation of QC and ML for effective big data analytics for enabling M-IoT based green smart cities of the future.

The future vision of this concept of smart cities will eventually shape the smart world/planet [205–207]; where the combined role of M-IoT, QC, ML, and big data analytics is of vital importance. There are various exciting applications of IoT, which includes, but not limited to, IoT for the industry, IoT for agriculture, IoT for smart offices and homes, IoT for healthcare, IoT for elderly care, IoT for farming, IoT for education, IoT for customer experience, etc. To realize these concepts, the open research challenges of M-IoT big data analytics needs attention of data scientists and communication engineers, the problems include: privacy and security of data, management of exhaustive data read/write operations, integration of heterogeneous types of devices, energy efficient routing, proactive caching, accelerating processing capability, and network support for massive number of devices, etc. QC and QML team with excitement to stand a central role in enabling smart world through assistance in M-IoT and runtime big streaming data analytics in 6G and beyond communications.

4) Security and Privacy
Providing privacy and security is a big challenge in the emerging world of everything connected to the network (e.g., the privacy of big data in M-IoT). In this context, the development of novel and complete security/privacy solutions is a demand of future communication networks. There are various attention-seeking open research problems from providing ultimate privacy in data mining and data processing to providing highly secure communication links. For example, in the paradigm of integrating ML in almost everything has recognized exposure of new types of privacy and security vulnerabilities, while the current understanding of these aspects is very limited [208]. Another example is, the enabling of secure multihop data routing in heterogeneous communication networks (say e.g., in M-IoT) itself opens various multidimensional research topics related to security/privacy, e.g., authentication of diverse types of IoT devices, the runtime encryption of streaming big data in ad-hoc networks, etc.

Furthermore, several unique security and privacy enabling solutions have emerged in the recent years, e.g., quantum cryptography and physical layer security etc. The fusion of ML (despite ML being itself vulnerable) for enabling physical layer security has also been used in the literature, see e.g., an ML-based method for antenna design to enable physical layer security in ambient backscatter communications in [209]. The ambient backscatter communications is a sustainable and independent communication solution for enabling M-IoT, through the exploitation of the already existing radio signals in the environment. The role of quantum cryptography, through its features of generating secret keys to multiple legitimate parties in a Vernam one-time pad cryptosystems, is highly recognized as a strong potential for the future of security/privacy [127]. The fusion of these QC features with intelligent systems can materialize into a holistic approach for enabling ultimate security and privacy in big data and massive connectivity era of 6G and beyond communications.

5) Harmonization and Interoperability of Networks
The 6G wireless networks are envisioned to be driven by on-demand self reconfigurability and interoperability with complete harmonization in the co-existence of all of its predecessors. In enabling this ambition, the evaluation of real-time state information of everything in the hybrid 1~6G network including network-infrastructure, network-edge, air interface, and user-side. For the realization of such consciousness and responsiveness in massively connected heterogeneous networks of future, very robust processing and learning capabilities will be required. In this context, QC-assisted ML and QML being capable of manipulating multi-state, multi-dimensional, and large-sized data, can be seen as the potential enablers.

B. AIR INTERFACE AND USER-END
The emerging paradigms of software-defined metamaterials based configurable leaky wave antennas and Large Intelligent Surfaces (LISs) combined with large-scale multiple antenna systems operating at a very broad range of frequency spectrum (microWave, multiTerraHz, visible light, etc) in distributed and undistributed fashions will completely redefine the physical framework of the air interface. Subsequently, such a physical framework combined with promising multiple-access and modulation paradigms like NOMA-MEC and Orbital Angular Momentum-Shift Keying (OAM-SK) can extend enormous provisions for simultaneously boosting all KPIs (e.g, capacity, energy efficiency, etc). In enabling these ambitions, quantum learning methods can play a vital role through instantaneous learning and manipulations of the available very large number of tweaking parameters (from: M-MIMO, antenna states, MEC, NOMA, OAM-SK, UAVs, etc) for searching global optimum solutions in a harmonized fashion across hybrid 1~6G heterogeneous network infrastructure. In this context, this section highlights various future research directions for realizing the air interface and user-side of 6G communication networks.

1) Configurable Multi-Antenna Systems
The large-scale multiple antenna systems have a strong potential in enhancing capacity and energy efficiency. The nature of wireless fading channels caused by different operating environments imposes different types of challenges on
the accuracy of communication. Considering the operating environment and network setting, the selection of antennas subset to serve a certain user with an optimal tradeoff between different performance quantifiers (e.g., data rate and power consumption etc) is an interesting research area. In this direction, the deployment of ML methods for this task has recently gained a vast response, see e.g., [210] and [211] for ML-based antenna subset and beam selection methods, respectively.

Furthermore, the use of reconfigurable antennas in multi-antenna systems is believed to provide a significant additional performance gain. Reconfigurable antennas are defined as capable of dynamically adapting their beam patterns by optimally selecting an antenna-state based on the available/estimated knowledge of CSI for each antenna-state; an example of such antennas can be named as: directional metamaterial reconfigurable leaky wave antenna. Evolution and revolutions in metamaterials and antennas are happening; e.g., a new form of antenna is the fluid antenna, that can be shaped to any required form. Future metamaterials and antennas can be seen as controllable through software. Using learning algorithms for smart selection among antenna-states software-defined antenna shape to optimize the overall system performance is another potential research direction, e.g., [212]. These promises provide a strong hope for enabling deployment of multi-antenna systems also at the user-side.

The processing and learning capability required for manipulating multi-dimensional, large-sized, and highly-dynamic (V2V) streaming data in such large-scale software-defined antenna shape to optimize the overall system performance is another potential research direction, e.g., [212]. These promises provide a strong hope for enabling deployment of multi-antenna systems also at the user-side.

2) Optical, mmWave, and TeraHz Communications

The plenty of available unused radio spectrum in the mmWave and TeraHz (THz) bands can be potentially utilized to meet the ever-increasing capacity demands. However, to enable its utilization, a dedicated effort is required for studying, modeling, characterizing, licensing, and regularizing these bands. The radio propagation characteristics associated with these bands are vastly different from those in the conventional microwave band. For example, the dominant specular reflections (in contrast to the dominant scattering in the microwave) and high isotropic pathloss in mmWave bands make it limited to only short-distance and LoS communications (tiny cells with low elevated BSs). This makes the highly-directional transmissions essential for enabling communications (e.g., mmWave). The establishment of initial access link in mmWave communications relies on searching in a high-resolution angular domain, which makes it a key challenge in enabling mmWave communications. The scanning of the angular domain for determining the initial transmission directions can be carried randomly or sequentially with the targets to optimize the access delay and overall system performance. For achieving a very high beam-forming gain, the searching over a very large beam space can significantly reduce the initial access performance. In this regard, an ML-based initial access method using RNNs in standalone mmWave communications has been proposed in [213].

Furthermore, quantum-assisted ML can be considered for the rescue for searching, tracking, learning, and predicting the candidate 3-D directions with quantum speed-ups. Moreover, in vehicular communication context, the carrier frequency being a linear scaling factor causes very high Doppler shift/spread even in very low mobility conditions (e.g., even vehicle engine vibrations can cause high Doppler shift [22]). This further reduces the coherence time, resulting in very fast time-variability in the channel characteristics. ML and quantum-assisted ML in modeling, characterizing, estimating, and tracking these dynamic channels can make a natural application. Moreover, in the context of required sensitivity for 3-D spatial directionality for communication over these bands, an instantaneous and accurate nodes localization and tracking for beamforming can be achieved through quantum-assisted learning methods.

Moreover, VLC is considered a new strong opportunity for enabling 5G communications, as there is no support for VLC is emerging 5G communication networks. This new technology offers various rich advantages, including but not limited to, large available bandwidth, power efficiency (lights are not just lights), no-interference with radio frequency bands, spatial reuse mechanism is easily degradable, human health friendly, and suitable in scenarios where radio frequencies are not suitable (e.g., underwater communications, etc). An important research direction in enabling VLC applications is achieving a tight localization accuracy. Various ML methods for indoor localization problems exist in the literature, see e.g., [214]. Moreover, ML for indoor localization in the context of VLC has also been studied in the literature, see e.g., [215]. Among many localization methods, the fingerprinting method can achieve considerable accuracy at the cost of high fingerprinting searching complexity. QC-assisted indoor localization method exploiting the offered quantum speed-ups for searching fingerprinting has also been investigated in the literature, see e.g., QSA in [35]. In summary, amalgamating these QC- and ML-assisted indoor localization methods in QML can open new research direction for precise and fast localization.

In addition to indoor applications, various outdoor applications of VLC are also suggested in the literature, e.g., underwater and V2V etc. In optical communications, the orthogonality offered by different states of OAM has strongly emerged into the research topics of OAM-multiplexing and OAM-SK. OAM systems are not limited to optical com-
munications, the electromagnetic (EM) OAM systems have also demonstrated a tremendous potential in achieving extraordinary spectral efficiency [216]. In this context, ML for operating and managing such multiplexing and modulation methods for optimization of optical/EM communication systems performance is a potential research direction. An ML-based method for adaptive \( m \)-ary demodulation of light beams carrying OAMs over free-space turbulence channels has been proposed in [217]. Also, an analysis of the use of ML for recognizing the intensity patterns in OAM-SK signals for underwater optical communications has been conducted in [218]. In the V2V context, the head- and back-lights of vehicles have been suggested as Tx/Rx units for establishing communication links, in [7]. In conclusion, precise and instantaneous localization and tracking of fast moving vehicular nodes is a crucial challenge in the outdoor applications of VLC, where QML can be explored as an enabler.

3) Tiny-Cells and Cell-Free Communications

The wireless communication networks are conventionally divided into cells for efficient spatial re-utilization of radio resources (e.g., macro-, micro-, pico-, femto-, small-, tiny-cells, etc). The enormous increase in the number of network devices and limited radio resources have led the evolution of cellular networks to tiny-sized cells (bringing users very close to the BSs) for more rigorous use of the resources. In the recent years, the concept of cell-free M-MIMO networks has emerged, which is defined by a massive amount of spatially distributed BSs (typically single antenna) serving a relatively small number of single antenna user devices through Time Division Duplex (TDD) operations, by exploiting the estimated CSI at the BSs. The user-centric transmission not only overcomes the inter-cell interference encountered in the conventional cellular networks but also provides macro-scale diversity [219].

Another cell-free communication concept revolves around introducing mobility in the BSs, e.g., mobile HAP serving the users on the ground. The footprint of such HAP forms a cell on the ground, which evolves along time due to the mobility and trajectory of HAPs. These HAPs can extend the network coverage to the remote users which are beyond the reach of land BSs, or to the users involved in situations like remote scientific exploration campaigns, coping with disastrous situations, etc. Various other interesting recent concepts are directly linked with this futuristic concept of flying BSs, which can be named as proactive caching, optimal resource allocation, trajectory prediction, multi-objective routing, user association, network topology reconfiguration, etc. In this context, Quantum-assisted ML and QML can play a central role for the best exploitation of available resources and large-dimensional data for enabling cell-free intelligent communications in 6G and beyond communication networks, through assistance in all the tasks spanning from proactive caching at flying BS to the estimation of a massive amount of channels in cell-free M-MIMO.

4) Auto-Encoder

End-To-End learning aims at representing the entire communication system from a transmitter to the receiver with a single learning block. This fascinating concept allows the learning of transmitter and receiver behavior for jointly optimizing all the operations based on an end-to-end error in recovery accuracy. The main operations of a typical communication receiver (like demodulation, channel estimation, channel equalization, symbol decoding, etc.) are classically performed in sequence to decode the information from the received corrupted symbols. Whereas, in an ML-based end-to-end system, an equivalent of all the operations can be combined within a single block, say e.g., a DNN block, where all the operations are realized in the form its layers. The scope of DL as an end-to-end solution to channel estimation and symbol detection tasks in OFDM systems has been investigated in [20]. Another conditional GAN-based end-to-end communication system has been proposed in [18]. Furthermore, a DL based end-to-end system design, referred to as auto-encoder, to jointly optimize both transmitter and receiver components in a point-to-point communication scenario has been proposed in [19]. The proposed end-to-end system model uses cascaded DNNs implementing data transmission, propagation channel, and receiving operations; where the layers representing the known propagation channel are fixed (not trainable). The information symbols (baseband) are feed as input to the DNN based end-to-end system and the symbol estimate is processed at the output. Also, in [220], the end-to-end system concept has been extended for performing equalization and synchronization tasks for frequency-selective channels in OFDM systems. In addition, authors in [221] proposed an unsupervised learning based approach to autoencoder concept for minimizing reconstruction loss through artificial impairment layers to model the channel. In addition, the autoencoder based end-to-end system concept has been extended to MIMO systems in [222]; where both open- and closed-loop systems assumed with and without CSI feedback have been studied, respectively. Moreover, end-To-End learning based over-the-air transmission method exploiting transfer learning concepts has been proposed in [223], where various challenges in training of such systems under realistic channel conditions are indicated (e.g., missing channel gradient etc). These studies involve the assumption of prior available channel statistics (i.e., available differentiable channel model).

The training of DNNs is usually performed through back-propagating the gradient of the loss function, however, the unavailability of prior knowledge of channel statistics, in end-to-end DNN systems, prevents the back-propagation of gradients. In this regard, a channel model less novel learning method combining supervised and reinforced learning for end-to-end systems has been recently proposed in [224]. In this method, the end-to-end accuracy loss for each decoded symbol at the receiver is fed back as a reward from the environment for loss optimization without requiring the gradients from channels.
This popularity of auto-encoding of a single end-to-end communication link in a DNN, is a strong motivation to extend this idea of auto-encoding an entire heterogeneous wireless access network in a QC-assisted very deep ANN to enable an instantaneous unique encoding response for service provisions through hybrid 1~6G network infrastructures. This QML based notion of auto-encoding the entire (very-complex and -dynamic) wireless access networks of the future can provide a strong potential in finding entirely unique solutions for best utilization of network resources and delivery of services.

5) Learning at User-Side
Considering the lack of computational and energy resources available at the user nodes, various tasks which are naturally of the user’s side are today preferred to be performed at the serving station or cloud-side of the network. For example, in TDD M-MIMO systems, the data for downlink is pre-coded at the BS in order to relieve the user nodes from the burden of CSI estimation and data decoding. The number of channels to be estimated in a Frequency Division Duplex (FDD) M-MIMO system at user-side is directly proportional to the number of antennas at the BS. For such an FDD M-MIMO system, a dictionary learning-based channel estimation method has been proposed in [225]. The TDD scheme is usually preferred over FDD in such large-scale multi-antenna systems for the reasons to avoid the tedious task of estimation of the massive amount of downlink channels at the user-end, however, this causes the pilot contamination problem imposing limitations over the capacity, which can get severe in high mobility scenarios.

The software-defined (fluid) antennas are expected to provide rich diversity at the user nodes, of similar level which M-MIMO systems can provide at the BS side. The antenna tweaking parameter for configuring antennas at the user-side opens the possibility of intelligently manipulating them through ML for overall performance optimization. Moreover, with the advent in computing methods (e.g., QC) and evolution in computationally efficient ML methods, it can be foreseen that the constraints on computational capability and related problems (e.g., battery life etc) at the user nodes may completely vanish in the future. These aspects will open new horizons of possibilities for better exploitation of network resources. Quantum-assisted DL combined with deep transfer learning methods can potentially come up with intelligent and dynamic solutions for optimal utilization of the network resources with consideration of the provision of intelligence at not only the network-side but also at the user-side of the communication links.

6) Multiple-Access
The need for privileging the access of wireless medium to a massive amount of users in an ultra-efficient fashion has led to the evolution of multiple-access mechanisms. Conventionally, the orthogonalization has been achieved through a clear distinct allocation of resources to the users through slicing of available resources in time, frequency, code, or space domains. With an increase in users, the mechanism for spatial re-utilization of the resources emerged, which further evolved into the idea of cognitive radio. As of today, a further massive increase in the number of network users/devices is causing the transition from conventional orthogonal multiple access methods to (random) NOMA methods. The NOMA scheme utilizes an additional domain of power, while it mostly revolves around the hybrid of different conventional orthogonal multiple schemes, i.e., division in joint multi-dimensional space of power, angle, and code etc. In NOMA, advanced signal processing methods are exploited to suppress the interference in order to accurately decode the data symbols, e.g., SIC. The hardware support for such SIC at user nodes is also released as NOMA-chipset [226].

The NOMA principle is also believed to be a convenient method in realizing massive connectivity in the context of IoT in emerging 5G networks [227–229]. However, there are a number of recognized fundamental performance limiting factors of NOMA, which include: high computation complexity for SIC in massive number of users context, inefficient transmission-time consumption, and estimation/feeding of CSI for a large number of users (specially for very fast time-varying channels, e.g., mmWave V2V channels). Combining the operations of channel estimation, channel equalization, and symbols decoding in a single block can help in resolving these limitations [30]. Moreover, the advanced ML concepts of DL, online-learning, transfer learning, and auto-encoder can together help to overcome these limitations. Online adaptive ML approach for detection in NOMA in the context of 5G networks has been proposed in [230]. Also, ML for optimal user clustering and power allocation in mmWave NOMA has been proposed in [231]. Furthermore, a DL-aided NOMA scheme has been proposed in [232], where DL has been used for learning channel conditions in an end-to-end learning fashion. In addition, NOMA-based MEC network has also received considerable attention in the recent years, see e.g., [233, 234]. Both NOMA and MEC being intelligence and computing power demanding technologies, QML can be seen as a strong potential enabler.

In the context of achieving the speed-ups of multi-user transmission/detection optimization, various quantum-aided schemes have been proposed in the recent years. As discussed in Sec. IV-A4b, quantum-assisted multi-user detection/transmission has been studied in various scenarios including direct-sequence spreading SDMA-OFDM, CDMA and SDMA, Multi-Carrier IDMA, and NOMA systems in [163], [164], [165], and [38], respectively. In the perspective of both ML and QC being rationally considered for improving operations of various multi-access schemes (including NOMA), it can be speculated that QML can take a vital role in reinventing the multi-access methods in 6G and beyond communications.
7) Intelligent Cognitive Radio and Self-Sustaining Wireless Networks

Leveraging intelligence to fully automate future communication networks for enabling operations like self-management, self-optimization, self-healing, and self-protection, is a clear need for future networks. The software-defined cognitive radios are designed to achieve reliable communication with minimal use of natural resources through intelligent operations (e.g., intelligent spatial reuse etc) learned from the environment (e.g., radio scene analysis). All the tasks for operating cognitive radios, e.g., sensing spectrum gaps, network (spatial) interference analysis, CSI estimation, power control, and dynamic resources management, etc, have natural connections with the deals offered by ML methods. Therefore, ML methods have thoroughly been reviewed for cognitive radios, see e.g., [235]. The role of AI in enabling cognitive radio networks has been strongly endorsed in [236], Alos, the scope of DL and evolutionary game theory for dynamic spectrum access in a cognitive radio network has been discussed in [237]. With the growing size and complexity of communication networks, the future of cognitive radios seems to be in the sensing and optimization through advanced ML methods assisted with quantum speed-ups.

The concepts of everything-connected-to-the-network and everything-as-a-service are rapidly evolving. The emerging 5G communication networks are expected to enable a variety of new services which may have various diverse requirements. When confronted with increasing services demands and network complexity, leveraging intelligence to the network can play a vital role. The 5G era is expected to lay the foundation of intelligent communication networks [54] by introducing some basic AI based operations, such as intelligent resource management, intelligent management of services provision, intelligent control, etc. A complete AI solution for intelligent cognitive and self-sustaining networks may emerge in B5G communications. The amount of available configurable parameters at a communication node is rapidly increasing, which is projected to be 2000+ parameters in a typical 5G node [54]. To this end, 5G cellular networks are expected to establish a framework of employing primary intelligence to the network by realizing self-organizing features. In the emerging architecture of 5G networks, the network design is centralized with lacking capability of complete robustness in design, dynamicity in the services types and flexibility in end-to-end NS, for recognition, management, and provision of new types of services.

The aforementioned indications assert that the truly intelligent self-sustainability in communication networks will be materialized in 6G and beyond communication with QC-assisted ML and QML as the potential enablers.

VI. CONCLUSIONS

In this paper, we have provided a comprehensive review of the emerging technologies including ML, QC and QML, and put forward our vision for QC- and QML-assisted framework towards enabling beyond 5G wireless networks. First, the target services offered by emerging 5G communication networks and the open research challenges for B5G communication networks have been detailed. Subsequently, the state-of-the-art of quantum, QC-assisted, ML-assisted, QC-assisted ML, and QML-assisted communications have been thoroughly reviewed. Furthermore, a QC-assisted ML and QML based framework for 6G communication networks has been proposed. In the context of the proposed framework, detailed discussions on various promising new technologies, open research problems, and future research directions have been provided.

More importantly, various potential enabling technologies for network-infrastructure, network-edge, air interface, and user-side of the proposed 6G framework have been identified and discussed. At the network-infrastructure and -edge levels: the role of the proposed framework for intelligent proactive caching, intelligent MEC, multi-objective routing optimization, resource allocation, massive-IoT management, big data analytics, interoperability harmonization, secure links assurance, and data privacy assurance aspects have been thoroughly discussed and recommended. Moreover, at air interface and user-end levels: various enablers for the proposed framework including mmWave communications, terahz communications, optical communications, VLC, small- and tiny-cells based communications, cell-free communications (UAV BSs and distributed M-MIMO), end-to-end autoencoding, learning at user-side, multiple access for massive connectivity, cognitive and self-sustainable radio networks, large scale multi-antenna systems, LIs, and fluid-antennas have been discussed in detail along with the associated challenges and potential future research directions.

REFERENCES


VOLUME xx, 201x


Quantum ML for 6G Communication Networks: Vision for the Future


SHREE K. SHARMA (S’12-M’15-SM’18) is currently a Research scientist at the SnT, University of Luxembourg. Prior to this, he worked as a Postdoctoral Fellow at the University of Western Ontario, Canada, and also worked as a Research Associate at the SnT being involved in different European, national and ESA projects after receiving his PhD degree in Wireless Communications from the University of Luxembourg in 2014. His current research interests include 5G and beyond wireless, Internet of Things, machine learning, edge computing and optimization of distributed communications, computing and caching resources.

He has published more than 85 technical papers in scholarly journals and international conferences, and has over 1600 google scholar citations. He is a Senior Member of IEEE and is the recipient of several prestigious awards including “2018 EURASIP Best Journal Paper Award”, “Best Paper Award” in CROWNCOM 2015 conference and “FNR Award for Outstanding PhD Thesis 2015” from the FNR, Luxembourg. He has been serving as a Reviewer for several international journals and conferences; as a TPC member for a number of international conferences including IEEE ICC, IEEE GLOBECOM, IEEE PIMRC, IEEE VTC and IEEE ISWCS; and an Associate Editor for IEEE Access journal. He organized a special session in IEEE PIMRC 2017 conference, worked as a Track Co-chair for IEEE VTC-fall 2018 conference, and has recently published an edited book on “Satellite Communications in the 5G Era” with the IET as a lead editor.

MOHAMMAD N. PATWARY (SM’11) received the B.Eng. degree (Hons), in electrical and electronic engineering from the Chittagong University of Engineering and Technology, Bangladesh, in 1998, and the Ph.D. degree in telecommunication engineering from The University of New South Wales, Sydney, Australia, in 2005. He was with General Electric Company of Bangladesh from 1998 to 2000 and with Southern Poru Communications, Sydney, from 2001 to 2002, as Research and Development Engineer. He was a Lecturer with The University of New South Wales from 2005 to 2006, and then a Senior Lecturer with Staffordshire University, U.K., from 2006 to 2010. He was then a Full Professor of Wireless Systems and Digital Productivity and the Chair of the Centre of Excellence on Digital Productivity with Connected Services, Staffordshire University, until 2016. He is currently a Full Professor of Telecommunication Networks and Digital Productivity and Head of the Intelligent Systems and Networks (ISN) research group at School of Computing and Digital Technology, Birmingham City University, UK. He is Principal Data Architect for a large scale 5G testbed in the UK to accelerate digital productivity & develop urban connected community. His current research interests include - sensing and processing for intelligent systems, wireless communication systems design and optimization, signal processing and energy-efficient systems, future generation of cellular network architecture and business modelling for Data-economy.

SHURJEEL WYNE (S’02, M’08, SM’13) received his Ph.D. from Lund University, Sweden in March 2009. Between April 2009 and April 2010, he was a Post-Doctoral Research Fellow funded by the High-Speed Wireless Center at Lund University. Since June 2010 he has been with the Department of Electrical Engineering at COMSATS University Islamabad (CUI), Islamabad, Pakistan, where he serves as an associate professor. His research interests include wireless channel characterization, multi-antenna systems, cooperative communications, physical layer security, and vehicular communications. He is a co-recipient of the best paper award of the Antennas and Propagation Track at IEEE VTC2013-Spring.

MD ASADUZZAMAN Md Asaduzzaman (M’16) received the B.Sc. and M.Sc. degrees in applied statistics from the University of Dhaka, Dhaka, Bangladesh, in 1999 and 2001, respectively; the M.Sc. degree in bioinformatics from Chalmers University of Technology, Gothenburg, Sweden, in 2007; and the Ph.D. degree in operational research from the University of Westminster, London, U.K., in 2010. He is a Senior Lecturer of statistics and operational research with Staffordshire University, Stoke-on-Trent, U.K., where he has been a faculty member since 2014. He was a Lecturer and an Assistant Professor of applied statistics with the Institute of Statistical Research and Training, University of Dhaka. His primary research interests include queueing, other stochastic models and mathematical programming for performance measure, capacity and resource planning and management in healthcare, and telecommunication and other communication networks. He is also interested in statistical computing, large-scale data mining, and analysis in earth, environmental sciences, and healthcare.

Dr. Asaduzzaman has received several awards, including the first runner-up prize of “The Doctoral Award” in 2011 from the Operational Research Society, U.K. and “Dean’s Honor Award” from the University of Dhaka, Bangladesh in 2013.