WILEY | Hindawi

Research Article

A Novel Architecture for Future Classical-Quantum Communication Networks

Fabrizio Granelli ,[1] Riccardo Bassoli ,[2] Janis Nötzel,[3,4] Frank H. P. Fitzek,[2,5] Holger Boche,[3,4,6] and Nelson L. S. da Fonseca[7]

[1]*Department of Information Engineering
and Computer Science (DISI), University of Trento, Trento, Italy*
[2]*Deutsche Telekom Chair of Communication Networks, Institute of Communication Technology, Faculty of Electrical and
Computer Engineering, Technische Universität Dresden, Dresden, Germany*
[3]*Department of Electrical and Computer Engineering, Technische Universität München, München, Germany*
[4]*Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, 80799 Munich, Germany*
[5]*Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, Dresden, Germany*
[6]*Cyber Security in the Age of Large-Scale Adversaries (CASA)—Excellenzcluster, Ruhr Universität Bochum, Bochum, Germany*
[7]*Institute of Computing, State University of Campinas, Campinas 13083, Brazil*

Correspondence should be addressed to Fabrizio Granelli; fabrizio.granelli@unitn.it

The standardisation of 5G is reaching its end, and the networks have started being deployed. Thus, 6G architecture is under study and design, to define the characteristics and the guidelines for its standardisation. In parallel, communications based on quantum-mechanical principles, named quantum communications, are under design and standardisation, leading to the so-called quantum internet. Nevertheless, these research and standardisation efforts are proceeding in parallel, without any significant interaction. Thus, it is essential to discuss an architecture and the possible protocol stack for classical-quantum communication networks, allowing for an effective integration between quantum and classical networks. The main scope of this paper is to provide a joint architecture for quantum-classical communication networks, considering the very recent advancements in the architectural design of 6G and the quantum internet, also defining guidelines and characteristics, which can be helpful for the ongoing standardisation efforts. For this purpose, the article discusses some of the existing main standardisation processes in classical communications and proposed protocol stacks for quantum communications. This aims at highlighting the potential points of connection and the differences that may imply future incompatible developments. The standardisation efforts on the quantum internet cannot overlook the experience gained and the existing standardisation, allowing the creation of frameworks in the classical communication context.

1. Introduction

The rise of new fundamental theories in physics always opens the door for subsequent advancement in practical physics and theoretical engineering. A fundamental theory of the last century is quantum mechanics. In the last decade, quantum phenomena have been applied to various fields such as photonics, computing, and cryptography. Moreover, quantum mechanics has become the primary enabler for a disruptive evolution in communication and computing systems, addressing existing open challenges not possible before.

The first worldwide telephone services required direct links between all communications entities. Next, the

communication paradigm became circuit switching, which provides a dedicated circuit between a source and a destination. Next, the communication networks evolved to packet switching. This paradigm allows information to be sent into a finite set of bits (messages), stored and forwarded by each router (switch) throughout the network, allowing communication among multiple heterogeneous entities in a scalable way. Packet switching was adopted as the transfer mode for the Internet, being its fundamental architectural choice for the design of a scalable network of networks. Another essential adoption to promote scalability was the employment of the tcp-ip protocol suite.

Access networks have played a vital role in expanding the Internet, allowing users to access the Internet with different devices. Especially, cellular networks have become one of the significant types of access after the advent of smartphones. 5G and 6G future networks bring a new approach to end-user communications. In fact, future networks will be an ecosystem (or pan infrastructure) capable of interconnecting highly heterogeneous networks, supporting demanding requirements and several different verticals. This will be possible via network virtualization, which is the software-based implementation of network functions, running on general purpose hardware. This breakthrough opened the way for a new approach, called compute-and-forward [1]. This term reveals the new role that computing assumes in the management and operations of communication networks.

Currently, 5G is under deployment and its standardisation is going to reach its end with release 18, in 2024. That is why, the research and the design of 6G started in 2021 to prepare the ground for the respective standardisation effort, together with its subsequent deployment starting from 2030. In the design of the next generation of networks, critical trade-offs arise. For example, some services offered by 6G networks will target very low latency (1 ms), very high throughput (up to 1 Tb/s), and very high energy efficiency (10-100 times the one of 5G) [2, 3]. Next, the load of unprecedented security level should also be added to these requirements. These high requirements have motivated a search for technical and theoretical tools to build and model 6G networks. Since ultimately every communication network is designed through application of knowledge about physical systems to the design process, an interest has arisen into using the most advanced theory as a key ingredient to the design process. The study, design, and standardisation of the quantum internet started from these premises. In the EU, the qia and the quantum flagship started their work in 2018, also leading the IETF qirg. In their perspective, communications based on quantum-mechanical principles could lead to a new Internet based on quantum communications.

Indeed, the ongoing IETF Internet-Draft of qirg [4] states

[...] we are at a stage where we can start to physically connect our devices and send data, but all sending, receiving, buffer management, connection synchronisation, and so on, must be managed by the application itself at a level below conventional assembly language, where no common interfaces yet exist. Furthermore, whilst physical mechanisms

for transmitting quantum states exist, there are no robust protocols for managing such transmissions. [...]

Given the above premises, it is possible to see that not only the architectures of 6G and the quantum internet are still open research issues but also that the works are progressing independently, without any current clear integration. However, in order to solve the challenges stated by 6G and the quantum internet, a complete integration between the two communication networks is necessary. This implies it is now pivotal to discuss an architectural structure and the possible protocol stack for a future unique classical-quantum communication network. The design of this architecture is especially critical to allow for an effective and efficient integration between quantum and classical networks and to open the way for the study of possible protocols to manage the eventual "upper layers."

Thus, this paper aims at comparatively discussing in detail the status of the design and standardisation of classical and quantum communication networks. This aims at highlighting the potential points of connection and the differences that may imply future incompatible developments. Side by side, the article also describes the architectural and protocol stack's characteristics, focusing on key aspects like softwarization, layering, and synchronisation. Next, this work proposes an architectural design, with respective guidelines, in order to suggest an effective integration of 6G and quantum communication technologies, also beneficial for the future research on the quantum internet. We believe that since quantum communication networks will be built on top/next to the classical ones [2, 4], the discussions and standardisation efforts on the quantum internet cannot overlook the experience gained and the existing standardisation and realized frameworks in the classical communication context. In parallel, the current work on 6G must embrace the quantum communication technologies to go beyond its intrinsic limitations. The article starts with Section 2.1, which lists the main standardisation bodies in classical communications and introduces the research and design path towards 6G. Next, Section 2.2 summarises the characteristics of the classical Internet, especially focusing on the protocol stack and the layering problem. Section 2.3 explains the aspects of softwarization, focusing on its main architectural characteristics and standardisation effort. On the other hand, Section 3 shows the design and standardisation of the quantum internet, showing the layering aspects and the study of the protocol stack. Finally, Section 4 describes our new proposed architecture for future classical-quantum communications, taking advantage from the lessons learnt from classical and quantum communication technologies.

2. The Beginning of the Story: From the Internet and 1G to 5G

2.1. Main Standardisation of Classical Communication Networks. Major international standardisation bodies in

traditional telecommunications and networking include [1, Chapter 2].

2.1.1. International Organization for Standardisation. The International Organization for Standardisation was responsible for publishing the OSI model, which is a conceptual model that characterizes and standardizes the communication functions of a telecommunications or computing system without regard to its underlying internal structure and technology. The OSI model was defined in the document ISO/IEC 7498.

2.1.2. IEEE. The Institute of Electrical and Electronics Engineers published relevant standards such as IEEE 802.11 for WLAN, IEEE 802.3 (defining the physical layer and medium access characteristics of wired Ethernet), and IEEE 802.16 (for Wireless Wide Area Networks, so-called WiMAX).

2.1.3. 3rd Generation Partnership Project. The 3GPP covers cellular and mobile telecommunications technologies, including radio access, core network, and service capabilities. 3GPP is the main driver for the wireless 5G standardisation process with the current release 15/16/17.

2.1.4. ETSI. ETSI represents the recognized regional standard bodies dealing with telecommunications, broadcasting, and other electronic communications networks and services. ETSI partners with 3GPP to develop 4G and 5G mobile communication systems.

2.1.5. ITU-T. The mission of ITU-T is to ensure the efficient and timely production of standards covering all fields of telecommunications and ICT on a worldwide basis, as well as defining tariff and accounting principles for international telecommunication services.

2.1.6. Internet Engineering Task Force. The mission of the IETF is to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet. The IETF currently is the main driver for computing elements in communication networks through their standardisation activities on SDN and NFV.

2.1.7. Internet Research Task Force. The IRTF focuses on longer-term research issues related to the Internet while its parallel organization, the IETF, focuses on the shorter-term issues of engineering and standards development.

A primary aspect in all standarization efforts concern network synchronisation, which is a pillar to enable communication networks since it allows most of the protocols at every layer of the stack. Moreover, it is also fundamental for traffic engineering and assessment of most of the network performance metrics. Synchronisation can be performed in three possible domains [5]: time, phase, and frequency. Since the phase is also derived in the time, domain, time, and phase synchronisation are addressed concurrently.

A synchronisation scenario for 5G and beyond networks is depicted in Figure 1. Each network node usually has a clock (called slave clock measuring time $t_{s,i}$), which is synchronised with a so-called master clock time, t_m, equal to the reference time of the Global Navigation Satellite System (GNSS), obtained via a satellite link. The direct transmission of timing information from the gnss to each node of the network is impracticable because of technical limitations such as the coverage of indoor environments.

The ITU published its recommendation G.8271.1 in 2017, followed by an extension in 2020 [6]. These documents defined the maximum bounds on phase and time synchronisation error (see an extract in Table 1). Moreover, after the definition of the terminology to identify the devices involved in the synchronisation procedures, it stated the minimum accepted tolerance to phase and time synchronisation errors at the boundary of packet networks. Finally, this recommendation described the characteristics of the packet-based method for the distribution of time and phase synchronisation across a network.

Side by side, the Institute of Electrical and Electronics Engineers (IEEE) released the updated version of the standards IEEE 1588 [7], in 2019. In particular, this standard sets a packet-based synchronisation protocol called ptp. The main network architecture of ptp is depicted in Figure 2. The IEEE standard sets a network consisting of a k number of devices, where one represents the master clock with clock time t_m, which spread timing information to $k-1$ slave devices, owning clock time $t_{s,i}$, with $i = \{1, \cdots, k-1\}$. The reference time is sent to the prtc at the master clock from a gnss. Next, the time signal is passed to the Telecom Grandmaster (T-GM), which encapsulates timing information in the packets to be sent through the network.

The communication network between a master and its slaves consists of three kinds of devices: routers without support for packet-based synchronisation and routers with T-BC and T-TC. The T-BC has multiple ports and it can become a master or a slave as well. Nevertheless, it cannot be an end user (e.g., a sensor and an actuator). Another role is the termination and regeneration of timestamp messages. The T-BC devices can measure the residence time rt_i, which is the time a packet resides in the device from input to output ports. This calculation is separately performed for downstream (rt_i^{ms} from master to slave) and upstream (rt_i^{sm} from slave to master) communication. Side by side, the t-tc is fundamental to achieve a synchronisation accuracy in the order of microseconds or below. In fact, it can measure not only the residence time in a router/switch but can also measure processing and queuing delays. Next, T-TC devices can also be peer-to-peer, which implies the capability to also measure the link-propagation latency between similarly equipped ports at the opposite sides of the respective link.

The 3GPP has been active in the definition of the requirements, technologies, and protocols for aerial communications since 2017. The standardisation effort started analyzing the required additional characteristics to be added to LTE to provide optimal connectivity to UAV [8]. Next, in 2019, 3GPP started the work on integrating 5G upcoming networks with UAVs as base stations. This was published in the Release 17 [9]. Currently, the trend is to expand this preliminary approach to a full so-called three-dimensional
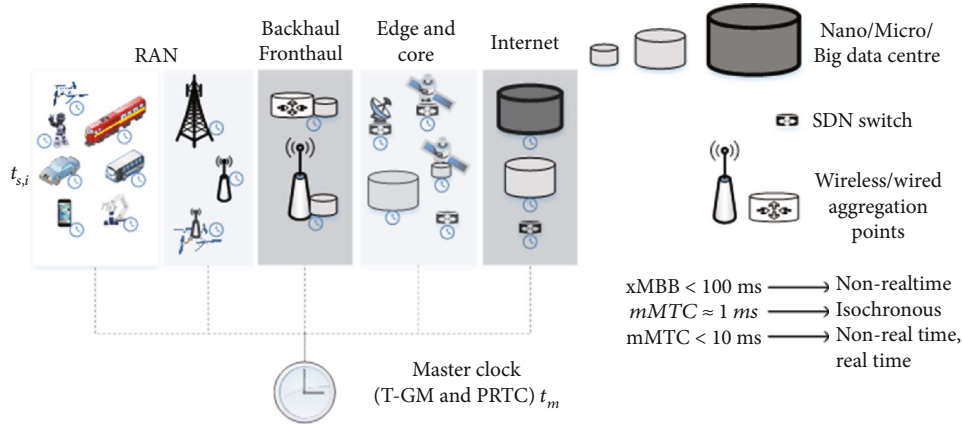
FIGURE 1: 5G and beyond communication network scenario with specification of master (time t_m) and slave clocks (time $t_{s,i}$).

TABLE 1: Synchronisation classes defined by the recommendation ITU G.8271.1, with the respective time/phase requirements [6].

Class	Error requirement	Possible application
Class 1	500 ms	Billing
Class 2	$100 - 500\,\mu s$	IP delay monitoring
Class 3	$5\,\mu s$	LTE TDD for large cells
Class 4	$1.5\mu s$	LTE TDD for small cells, NR TDD
Class 5	$1\mu s$	WiMAX TDD
Class 6	65 ns	LTE NR MIMO

networking in the context of future 6G architecture, where UAVs, haps, and satellites will seamlessly converge to a unique communication network architecture. This was clearly stated in the recently-published deliverable D5.1 of the EU flagship Hexa-X project [10], which focuses on the guidelines, enablers, and technologies for the future 6G architecture.

Figure 3 depicts the three 5G architectures using satellite and the 6G architectural vision of three-dimensional networking. The transparent 5G architecture allows ues to directly connect to satellites. Next, the traffic can also be routed to the terrestrial gNodeBs via the ntn gateways. On the other hand, the regenerative 5G architecture considers the placement of the gNodeB and its related computing tasks within the satellite-aerial platforms as well. Then, such a gNodeB can also communicate with the terrestrial core network and the Internet via the ntn gateways. The hybrid 5G architecture is the most flexible one since it also includes the softwarization and subsequent functional split of the gNodeB. In fact, the gNodeB is split into a du and cu, which can be placed somewhere between the aerial-satellite and the terrestrial platforms. This is the preliminary version of the 6G three-dimensional architecture that has been envisioning by now. The current architectural vision that will be standardised by 2030 will create a sort of "continuum," in which the softwarized network functionalities are dynamically placed. This so-called continuum is seamlessly either horizontal (two-dimensional) or vertical (three-dimensional). Espe-

cially, the latter can involve different kinds of terrestrial and aerial platforms, satellite, and haps.

2.2. Architectural Characteristics of Current Communication Networks. A reference model gives a conceptual framework to abstract network functionalities. In communication networks, most of these models adopt a hierarchical layering approach. Layering in networks is similar to the concept of objects in software engineering, furnishing services, and hiding their implementation. In the hierarchical layering approach, layers are stacked one on top of the other. Each layer offers service to the layer immediately above and receiving service from the layer immediately below. A layer interface defines how the services can be accessed and what restricts the information that can be retrieved from a layer. The ongoing trend to have network layers represented in software gives room for innovation, and at the same time, hides to some degree the physical representation of the layer from the network engineer.

Similar entities (process, agents) at the same layer in different network nodes can communicate with each other by obeying a set of rules, called a protocol. There is a protocol or set of protocols for the provisioning of communication services at each layer, realized by the exchange of pdu. Pdus are composed of a header, payload, and a trailer also called sdu. A pdu header contains information for processing the pdu at a receiving device, and a trailer delimits the end of the pdu. Trailers are less adopted in different protocol pdus since most of the headers contain a field defining the PDU size. Two endpoints do not exchange pdus directly, but a pdu is passed to the layer immediately below until reaching the physical mèdium, where physical transmission effectively occurs. A layer-specific pdu is created at each layer by adding a header to the payload received from the layer above. The header is processed and removed at the corresponding layer at the receiving network node, and the resulting payload is passed to the layer immediately above.

Two standard reference models in communications networks are the OSI [11] and TCP/IP [12] reference models [13]. The Open System Interconnection, defined by the International Organization of Standardisation (ISO), is a de facto standard model commonly used to understand
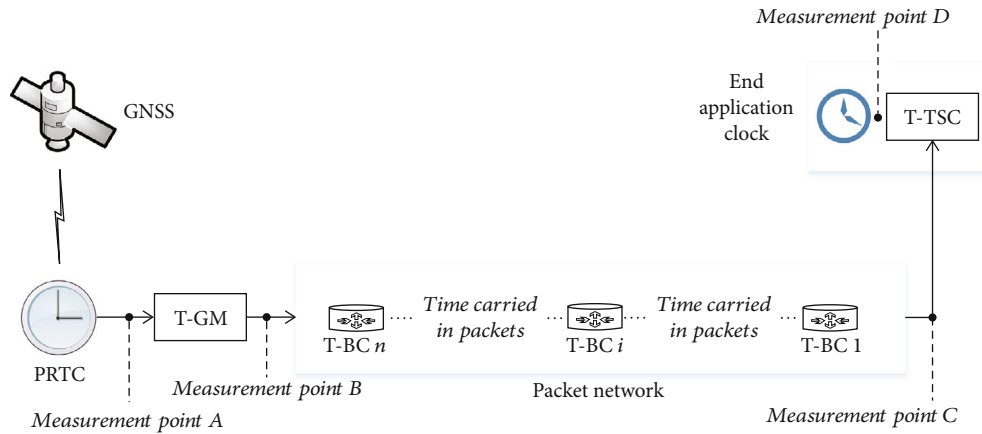
FIGURE 2: Time synchronisation model according to recommendation ITU G.8271.1 [6].
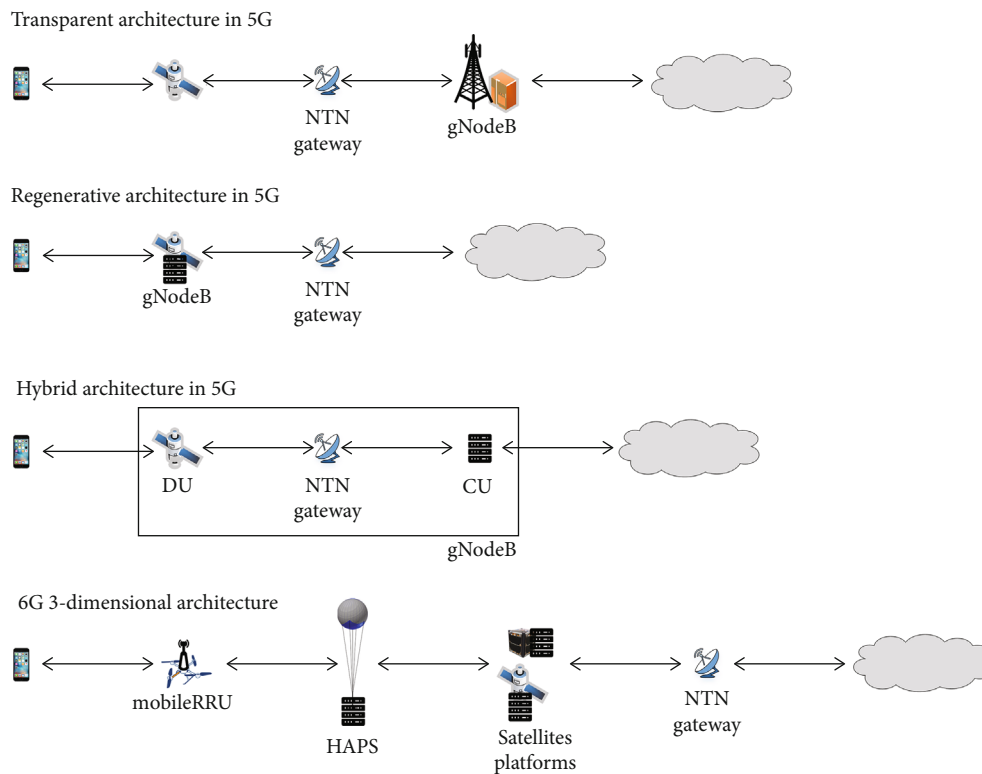

FIGURE 3: Satellite architectures proposed in 5G. The transparent payload acts as a relay since the gNodeB is on the earth surface. With the regenerative architecture, the base station functions are moved to the satellite. The hybrid architecture considers the potential softwarization and functional split of the functionalities of the base station. The last one is the 6G three-dimensional architecture, where softwarized functionalities of the base station can be activated dynamically at any aerial/satellite platform seamlessly.

new networking technologies and the relationship between different networking technologies [14, 15]. The OSI model defines seven layers: physical, data link, network, transport, session, presentation, and application layers, with functionalities described next.

In the physical layer, digital bits are transformed into electrical, radio, or optical signals, which are then transmitted on a physical communication channel. This layer's specifications include the definition of voltage levels and timing, data rates, maximum transmission distances, modulation scheme, and channel access schemes.

The data link layer gives the abstraction of a communication channel (link) for node-to-node data transfer. Error control and correction mechanisms support the creation of an error-free channel. Flow control, another common mechanism, attempts to avoid one node flooding another node with data. A multiple access control protocol needs to be defined in a broadcast type of link to avoid simultaneous

transmission of frames (data link pdus). Network nodes that realize connectivity and multiplexing of bits are called switches at the data link layer.

The network layer abstracts the subnet, that is, the set of routers and links. It provides the functionality of transferring packets from one node to another connected on the same network layer. Routing can be defined either as a predetermined fixed path or when the packet goes through a router (hop by hop). Since the network layer abstracts the subnet (network core), it is natural to place congestion control on this layer to avoid that the network enters a congestion state degrading the quality of the transport service substantially.

The transport layer abstracts the whole network between a pair of communicating processes in computers far apart. It is considered the first end-to-end layer connecting a data source and the destination. This layer defines the transport service seen by networked applications, which can be either connectionless or connection-oriented. A transport connection is a point-to-point reliable (error-free) channel that delivers pdus to the destination in the same order they were generated. The OSI model defines five classes of connections with distinct functionalities such as concatenation and separation, segmentation and reassembly of pdus, error recovery, reinitiation of connections, multiplexing/demultiplexing over a single fixed path, and flow control.

The session layer controls the dialogues between communicating endpoints and may include several transport connections. Session control also includes synchronisation (based on checkpoints) and token management (for access to critical regions).

The presentation layer is concerned with the syntax and semantics of the information transmitted. It makes possible communication between applications in computers with a different representation of information. It should define abstract data structures and the mapping of coded information to a standard abstract structure. The application layer contains the communication protocols used by applications. It hosts the various applications used by the end-users.

The OSI reference model provides a framework to compare and understand different network technologies. One example of such understanding remotes the launch of the Asynchronous Transfer Mode (ATM) networks which had four layers, and the first three correspond to the data link layer of ISO reference model with very few functionalities that could be considered functionalities typical of the network layer [16, 17]. The network layer also supports variable size packets (datagram) from the Internet Protocol (IP). This type of adaptation may be essential for adopting new physical and data link layers such as those in quantum communications.

Network architectures derived from a reference model may add planes to the layering models. These planes host specific functions composing a layer functionality. Typical planes are the data (user), control, and management plane. The data plane transports only packets generated by the end-user (forwarding function); the control plane transport control (signaling) packets which carry information for dynamic set up of the network; and the management plane, which coordinates the other two planes. A typical example is the transport of information by virtual circuits, fixed paths, or the transport of packets by the mpls [18]. The control plane is responsible for the setting and tears down of the virtual circuits (mpls paths), while the data plane is responsible for the forwarding of the packets generated by the users. Another example is the data and control plane of software-defined networks in which controllers residing on the control plane determine the routing of flows.

The TCP/IP reference model's development took a different path than that taken by the OSI model. The protocols were defined first, and then the reference model was specified. Indeed, the TCP/IP reference model resembles more a protocol suite than a predefined architecture. The technical standards underlying the Internet Protocol suite are under the IETF. The TCP/IP reference model loosely defines four layers: link, internet, transport, and application layers.

The link layer is not a well-defined layer, and it specifies only an interface with links and devices on the same link layer. The link layer can be a single link or a whole network architecture. Indeed, anything below the internet layer is considered a link layer. Such definition reinforces the fact that the internet layer is independent of hardware implementations.

The internet layer solved the crucial issue of interconnecting incompatible networks by adding a layer on the top of all networks without needing translations and mappings between the connected networks. Connecting different networks, in other words, making different networks work together (internetworking), calls for the essential routing functionality, which defines the path packets (pdus) should take from a source to a destination node. Routing in the internet layer is carried out hop by hop, and decisions are made considering only the packet's destination address. Following the minimalist principle, the internet layer's delivery model is unreliable, which implies that packets can be dropped at network routers in case no space is available on router buffers to store them for later forwarding. Packets can also arrive out of order at the destination. Such type of service is known as best-effort service and is provided by the IP protocol, the only protocol employed to transport information on the Internet. Signaling on the Internet is in-band, contrary to other networks which have separated channels for data and signaling. Nonetheless, the forwarding of packets can be imagined belonging to a data plane while determining the next hop (routing) as residing in a control plane.

The transport layer offers two types of transport service for the applications: a connectionless one provided by the udp, which adds no functionality on the top of the internet layer, and a connection-oriented service provided by tcp. The application layer hosts all the communication protocols employed by the applications running on the Internet. These protocols use the transport layer protocols through the interface provided by sockets APIs. The application layer in the TCP/IP model is often compared to a combination of the session, presentation, and application layers of the OSI model. Figure 4 compares the OSI and TCP/IP reference models [19].
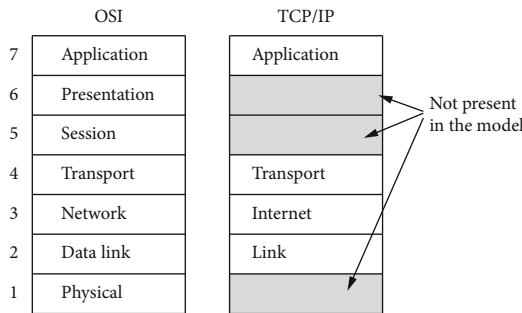
FIGURE 4: The TCP/IP reference model.

Although the layering principle is fundamental to network architectures and reference models, some functionalities are implemented in several layers and may require the interaction of mechanisms in different layers to realize the functionality needed effectively. An example of a cross-layer solution is error detection in tcp over wireless networks, where the link-layer hides and retransmits some lost packets to avoid unnecessary reduction of tcp window transmission. Such an approach is called cross-layer [20].

2.3. Future Softwarized Classical Networks. In Subsection 2.2, the protocol stack has been presented as a "monolithic" entity, consisting of layers interacting with each other via specific interfaces, the so-called sap. In such a structure, there is also the possibility of cross-layer solutions, but these are not as flexible as the software-based instances of network functions and operations previously described. In order to overcome this limitation, the concepts of pps and a wireless network operating system were proposed.

A programmable protocol stack is a software-based layered architecture, which can flexibly and adaptively manage protocols and network layers. The various entities in the virtualized protocol stack can reconfigure such as reassigning parameters, updating services, and replacing active functionalities, according to various conditions and requirements caused by users, network, and environment. This idea comes explicitly from the rise of applications for multimedia content distribution. Figure 5 depicts the logical structure of the two leading solutions proposed for pps, such as the Wireless Network Operating System and the Software-Defined Protocol.

First, the Wireless Network Operating System (WNOS) [21] exploits a network abstraction that targets a network control problem, given by the specific objectives of the hosted services. By characterising the network status via specific APIs, it is possible to adaptively optimise the KPIs like throughput and latency. In this scenario, the resources of the physical layer represent the available constraints. The PPS is included as a software-based pile, which adapts and configures according to the abstraction and the respective control problem to be addressed. The PPS also involves the physical layer since the adaptivity is possible via the deployment of reconfigurable radio technologies such as software-defined radios (SDRs).

Second, Software-Defined Protocol (SDP) system [22] consists of controllers and servers, which run specific blocks.

The SDP blocks perform packets' routing. The SDP controller sets up the protocol stack's functionalities and characteristics to adapt layering, ensuring the required QoS. The SDP controller also configures flow tables in the switches and within the blocks in SDP servers.

Management and orchestration represent a crucial functionality to enable proper control on the softwarization of network functions. The reference in this field is represented by the ETSI-MANO architecture (see Figure 6). The key of the architecture is the availability of the network function virtualization infrastructure (NFVI), which enables to virtualize the available computational, storage, and networking resources. The NFVI is controlled by the Virtualized Infrastructure Manager (VIM). Management and orchestration are implemented by means of the NFV (Network Function Virtualization) orchestrator that oversees the operation of the NFV manager.

Even if it cannot be considered an official standardisation effort, the uonos project represents an effort to extend the capabilities and characteristics of ETSI-MANO architecture based on SDN and NFV. It is led by the open source community hosted by The Linux Foundation. uonos specifically aims at proposing a standard architecture for distributed control plane [23]. The main idea is the realization of a new generation of the SDN control plane based on ONOS. The objective is the splitting of SDN controller's functions into a set of subfunctions or microservices. These functionalities are deployed as virtual containers and managed by the Kubernetes orchestrator.

The μONOS protocol stack employs a new generation control protocol such as P4/P4runtime [24], which guarantees greater flexibility compared to the original OpenFlow protocol. The communication between functionalities is via Google's gRPC-based protocols, including network management interface and network command operations. Currently, the μONOS effort is leading the research on the decomposition of the SDN controller. Since this has just started, there is still no available implementation to test its performance. Moreover, the communication protocol is based on gRPC, and not on the rest API. This implies some specific limitations. First, μONOS has a limited isolation mechanism, which means the core functions and applications share the same resources or processes. Second, μ ONOS cannot have on-platform tenant-specific applications but only tenant-aware ones: tenant-specific apps must be off-platform, and it should use rest APIs. Third, the on-platform applications are limited to Java-based languages: applications developed using other languages have to be off-platform and need to use rest APIs. Next, the horizontal service scaling is difficult. Finally, it has limited integration with and support for NFV that do not adhere to either an OpenFlow abstraction of that of a legacy network element.

From the architectural perspective, given the capabilities opened by network softwarization, ai is going to play a key role in the management and orchestration of future communication networks. In fact, 6G is planning the realization of in-network intelligence, so that ai becomes not only a service but fully an element of the network architecture [25]. First,
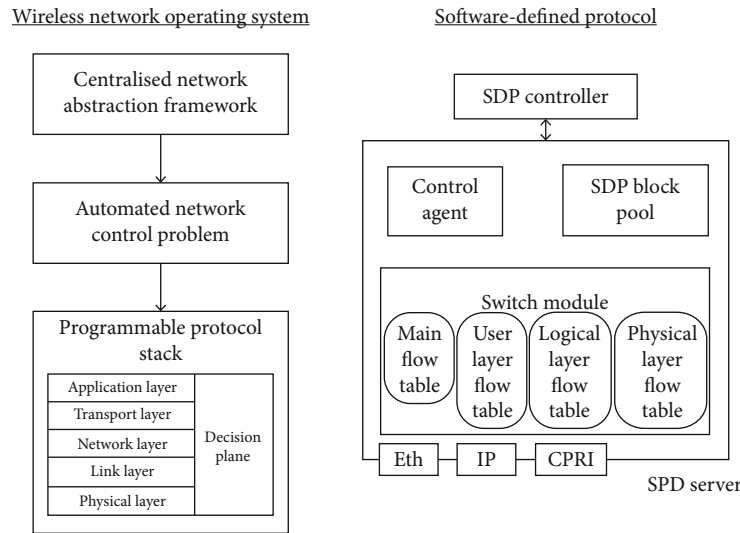
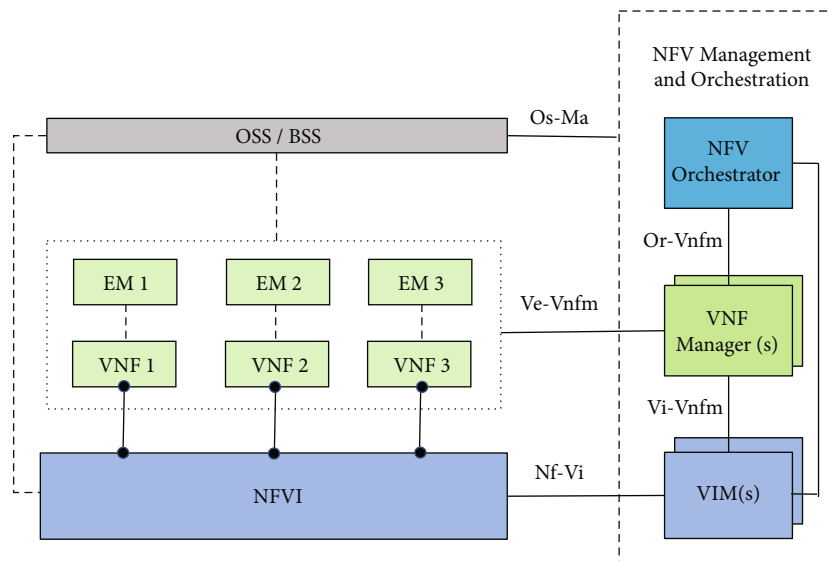FIGURE 5: Logical structure of the Wireless Network Operating System and the Software-Defined Protocol.

FIGURE 6: ETSI-MANO architecture.

the design of the AI-driven air interface of the ran will have a key role [25, 26]. In parallel, in-network learning methods will also be applied in the edge and core network for data, network, and users' management. As previously mentioned, in future 6G networks, each vnf will potentially be either a microservice or an intelligent agent. In the latter, intelligence will be integrated within several vnfs or sub-vnfs to realize multiagent systems, in which intelligent network entities collaborate to perform a specific network task [10].

Modern and upcoming communication networks are highly heterogeneous and complex ecosystems. Their stringent KPIs become necessary since the network design has been driven by new upcoming services such as the tactile internet, the industry X.0, and the internet-of-things. Moreover, future networks are going to fully employ network softwarization. This means that all the functions of the net-

work, which are implemented in dedicated hardware, will run in virtual environments. This complex communication context will require a significant improvement of the existing network synchronisation procedures.

Network time synchronisation is fundamental for secure and tactile network operations, which require precise synchronisation among the nodes of the network. There are two main approaches to time-synchronisation in networks: the deployment of independently synchronised clocks at each network node packet-based synchronisation of distributed clocks. In the former, each network device is equipped with an atomic clock. This is an expensive and generally impracticable solution due to its high cost. Normally, each network node has a clock, which is driven by an internal oscillator. In a networked system, where different nodes can also have different types of clocks; clocks are powered

by nonidentical oscillators. Thus, these oscillators have inconsistent behaviour in different conditions, which result in timing errors and make datagram-based synchronisation necessary. The most important standard is the IEEE 1588, which aims at transferring timing messages from a master reference clock through the communication network in order to synchronise slave clocks. Moreover, the synchronisation of virtual environments of softwarized networks adds additional synchronisation errors. These aspects make datagram-based methods not able to satisfy the increasing precision and security required by the critical services of future communication networks.

3. The Advent of Quantum Communication Networks

Future networks call for an increase in current storage and computing capacities, which also implies augmenting energy usage (for computing) and consumption (for communications). Additionally, in-network intelligence will demand a large number of resources for communication during data mining and distributed decision-making. Prediction of future network states will also bring computational overhead. All these aspects will require ultraprecise reliable synchronisation protocols, which must satisfy the low-latency KPIs.

Moreover, the targeted KPIs of the existing 6G proposal may arise contradicting objectives. For example, energy saving contradicts the massive amount of computing needed by in-network intelligence. Next, anticipatory networking sets a trade-off between low-latency and reliability. Moreover, low latency will be critical for the amount of computing, data mining, and quite-high data rates. In fact, increasing data rates and link usage will raise transmission and scheduling latency. These are only some of the several trade-offs and contradictions within the design and realization of future classical networks.

In order to exceed the intrinsic limitations imposed by the abovementioned issues, quantum-mechanical communications and computing have been considered to support envisioned future networks. By employing distributed quantum computing instead of classical computing, the exploitation of entangled qubits within several interconnected devices can achieve an exponential speed-up of the network computational capabilities with just a linear increase in physical resources. Thus, the limitations imposed by classical paradigms and "softwarization" can be solved by exploiting quantum-physical parallelism based on the concepts of quantum superposition, entanglement, and quantum measurement [2]. Next, an overview of the current preliminary status of the architectural design and standardisation of quantum communication networks will be provided.

3.1. Current Standardisation Procedures for Quantum Communication Networks.
The general prestandardisation focus group about quantum information technologies is the one belonging to ITU, called Focus Group on Quantum Information Technology for Networks (FG-QIT4N). This group was created in September 2019. The main objectives of the group are the study and definition of terminology and application for quantum information technologies, which can also open the way for a collaborative platform for designing future quantum communications with the contribution of industry, technical experts, scientists, and policy makers.

The vision of the quantum internet [27, 28] aims at designing and developing a quantum communication network, interconnecting quantum computers to target various quantum-enhanced network aspects such as security, synchronisation, and computing. The standardisation process of the quantum internet is going under the leadership of the IETF research group called qirg [4].

The standard starts with the definition of the atomic entity of information, the qubit, and subsequently the multiqubit systems. Entanglement between qubits is defined as the fundamental quantum resource for communication. However, quantum communications introduce some challenges such as those ones resulting from measurement, nocloning theorem, and the fidelity. Furthermore, the document [4] states the inadequacy of direct transmission since it requires expensive quantum error correction mechanisms to keep quantum errors at the minimum. At this point, an important claim of the draft standard [4] is that

[...] quantum error correction is not expected to be used until later generations of quantum networks.[...]

Then, the most efficient way of distributing entanglement remains the use of Bell pairs, which is the fundamental pillar of the basic quantum protocols of dense coding and teleportation.

Entanglement can be generated in three main ways: at midpoint, at source, and at both end-points. The first involves a third party, which distributes the entangled qubits via quantum channels to the communicating nodes. The second and the third only involve one or both the communicating nodes in the entanglement generation and distribution. Since entanglement is very sensitive to time and interactions with the environment, entanglement swapping is the procedure, which can be used to ensure distribution for distances greater than 150 km.

When drafting the architecture of the quantum network, the document states

[...] In a quantum network, the entangled pairs of qubits are the basic unit of networking. These qubits themselves do not carry any headers. Therefore, quantum networks will have to send all control information via separate classical channels which the repeaters will have to correlate with the qubits stored in their memory. [...]

From this quotation, it is important to make some initial architectural considerations. First, entangled pairs are lower-layers entities, mainly upper-physical and lower-link layers. However, the kind of correlation created by these basic units of networking has an inherent cross-layer nature so that they can affect the network layer via the output information of the specific sap. This will be considered in the proposed new architecture in Section 4.

Next, the draft standard [4] makes a distinction between control and data plane. This network abstraction is defined as fundamental to set for example forwarding rules of qubits.

The control plane should be similar to its classical counterpart, and it does not handle quantum data in general. However, some quantum control protocols might be defined like the quantum ping. Additionally, the document defines so-called control information messages, which aims at managing single entangled pairs. Nevertheless, the characteristics of control plane also in relation with data plane are claimed to be out of the scope of [4]. Regarding the data plane, the draft standard states the existence of two concurrent planes, classical and quantum, with their respective operations and protocols. In the document, the authors say that the design of the specific network abstractions remains an important open challenge for the realization of interoperable quantum network protocols.

Finally, the draft standard [4] proposes a possible employment of mpls in quantum networks. Since the distribution and maintenance of entanglement among network nodes is a stateful process, the use of connection-oriented solution is the one suggested. This implies the connection via virtual communication circuits among network nodes for quantum entanglement distribution. Next, [4] mentions that signaling functions are needed for setting up virtual circuits so that protocols like resource reservation protocol (RSVP) or OpenFlow can be employed. Additionally, the generalized mpls (GMPLS) is suggested as a good potential protocol to handle separate channels for control and data plane flows.

Quantum key distribution (QKD) is a security protocol, which provides information-theoretic security against a third party such as an adversary or eavesdropper. The protocol distributes quantum keys to network entities. The security of these keys is not ensured by the limited capabilities of the adversaries. On the other hand, the physical characteristics of quantum mechanics make these keys inaccessible. In fact, the security mainly comes from the no cloning theorem and an underlying information gain/disturbance trade-off. If an eavesdropper interacts with a shared entangled state, it introduces an irreversible disturbance that is proportional to the information that has been gained. Then, the communication parties can detect and quantify the presence of noise and abort when it reaches levels that reveal the attack. It is important to notice that QKD protocols assume the existence of an authenticated classical channel among the parties that have to share the keys.

3.2. Currently Proposed Architectures for Quantum Communication Networks.

The existing literature on quantum communication networks has focused on the design of network architectures and protocol stacks for quantum-only networks. This means that the combination between classical and quantum infrastructure has not been considered. The following overviews the main trends proposed for quantum communication network architectures and protocol stacks.

Figure 7 depicts the initially proposed protocol stack for quantum communications by [29]. In particular, it refers the physical and link layers. The main protocols, which have been proposed, are the MHP, the EGP, and the DCP.

The MHP is a control protocol, which was proposed for the upper-physical layer. This protocol should be implemented to comply with very stringent timing requirements because it is responsible for deciding the generation of entanglement. In that sense, it defines an MHP cycle, which granularity directly affects the communication performance (e.g., the throughput). The protocol uses time-division communication in which a timestamp and an ID set the detection window for each photon. The MHP was proposed based on two procedures: create-and-keep and create-and-measure. The former considers only quantum operations on photons (so-called quantum gates), while the latter allows for performing measurements. The results of these measurements are used by the EGP protocol.

The EGP is the core protocol of the architecture in [29]. The protocol exploits some assumed logical blocks such as a distributed queue, a qmm, a feu, and a scheduler. As mentioned above, the setup of entangled photons is helped by the MHP protocol. The EGP protocol maintains distributed queues, which schedule the requests of entangled particles. These queues can also employ different criteria of priority. The qmm logic block selects the specific photons that have to be entangled. A critical aspect is the "quality of the entanglement" or fidelity. The feu estimates the fidelity of the entangled photons, guaranteeing that this value is above the required minimum threshold. Next, the scheduler decides the serving policy of the queue.

The EGP protocol starts when a request for a number of entangled photons arrives from the above-two layers. At this point, the feu sets the specific requirements for fidelity and completion time of the process. Next, the request of entanglement is assigned to the distributed queue. The scheduler manages the status of the request so that finally, the qmm can successfully allocate the requested qubit. As mentioned above, the MHP allows the processing of the requests coming from the EGP. Each request in the queue has a unique identifier, which also helps the management of the respective qubits.

The DCP manages the distributed queue of requests of entanglement, coming from all network nodes. The protocol also stores the information about the requests, such as the creation time and the minimum time (that is introduced by the presence of a timeout cycle at the MHP).

In 2019, [30] provided a description of the protocol stack of quantum communication networks, from the physical to the network layer (see Figure 8). The physical layer has the same role and characteristics of its classical counterpart. In fact, it transmits/receives unstructured raw data via a physical transmission medium by converting the qubits into optical signals. Moreover, it converts the signals into different forms according to the transmission technology and frequency.

The second layer is a new layer between the physical and link layer, called connectivity layer. This layer is responsible for quantum error correction and setup of long-distance quantum communication links. In particular, the considered communications can be single-source unicast or multicast. The critical aspect of this layer is the distribution of
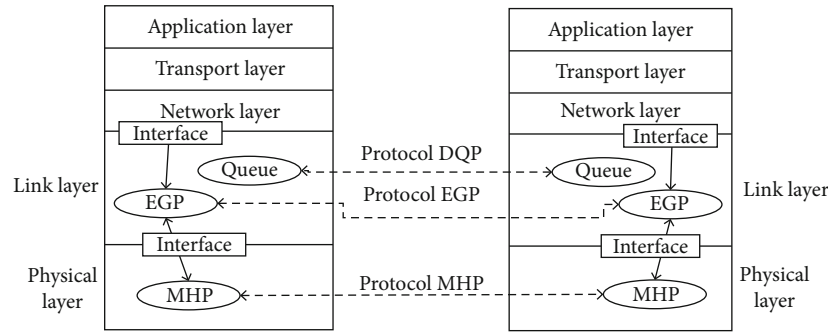
FIGURE 7: Initial protocol stack for quantum communications, proposed in [29].

FIGURE 8: High-level protocol stack for quantum communications, proposed in [30].

entanglement—either Bell pairs or Greenberger–Horne–Zeilinger (GHZ) states—among distant nodes, which is the pillar of quantum communications. An important aspect is that the operations within this novel layer are independent of the above link layer's protocols. The main objective is the decoupling between pure link layer and connectivity operations in order to simplify maintenance and network upgrades.

Next, the link layer hosts the protocols to create the network quantum state of arbitrary topology. The aim is to generate and distribute the entanglement to the device of the required quantum entanglement topology. This layer also manages entanglement distillation in order to ensure the specified level of fidelity. Entanglement swapping and merging quantum states are also performed at the link layer.

Finally, the network layer is responsible for manipulating entanglement and allowing routing at a network level. The primary devices involved are the quantum routers. The authors in [30] have proposed some preliminary quantum protocols, which are comparable to their classical counterparts.

3.2.1. Open Systems Interconnection Conformal Quantum Networks. In [31, 32], the authors consider a different route to the development of classical-quantum communication networks, following an approach of subsequent minimal changes to the existing network architecture rather than a radical replacement of the existing one. The work is focused on entanglement-assisted data transmission only, resulting in relatively minor changes to the existing architecture in the sense that only the physical and link layers are affected. The authors describe the Generate Entanglement When Idle (GEWI) principle [32] as a sender-side mechanism that starts generating and distributing entanglement as soon as there is no data in the sender-side data buffer. A correspond-

ing sliding window entanglement generation protocol is described in [31]. When there is data in the sender-side buffer but no entanglement stored between sender and receiver, the protocol transmits data without entanglement assistance. The recent development of entanglement-assisted communication techniques [33] makes the hybrid classical-quantum communication network structures an interesting proposal. Due to the need to distribute entanglement first, these proposals might suffer from the same problem that is inherent to quantum networks—namely, the dependence on a repeater architectures, where current achievements are promising but larger field trials have not yet been conducted. In addition, it is yet unclear in which data transmission scenario the entanglement-assisted communication schemes will bring the promised substantial benefits over nonassisted ones. While the focus on an end-to-end use of entanglement in networks is obvious from the literature, and current research has started to identify its potential [34] and use [33] for data transmission, the possibility of utilizing quantum technology in a localized fashion for the goal of faster data processing at lower energy consumption has been pointed out recently in [35]. The work [35] points out the potential of quantum signal processing where network functionalities are optimized using quantum techniques without any need to distribute entanglement over the network. Rather, the focus is on utilizing the enhanced sensitivity of quantum detectors and processing mechanisms. The research on the optimal data transmission methods has been pioneered in the fundamental works [36, 37]. A first concrete description for an implementation dates back to [38], follow-up works like [39, 40] described different variants and extensions of the proposed scheme. An excellent overview is given in [41].

3.3. Physical Layer Service Integration. Physical layer service integration (PLSI) is an approach that is emerging from a series of works analyzing communication tasks from an information-theoretic perspective, taking into account models going beyond the initial work by Shannon (sometimes these works are categorized as post-Shannon theory) in the sense of more accurate mathematical modelling of tasks, resources, and involved parties. PLSI identifies the principal resources available at the physical layer, the bottlenecks and vulnerabilities in typical communication scenarios, and the ability of the physical layer resources to solve

the identified problems. Plsi aims to rebalance the network architecture by adding critical functionalities to the physical layer, to accelerate their execution. Since the physical layer of every future communication network utilizing quantum technologies has to eventually address the question how to generate and distribute entangled states, plsi is important when thinking about quantum networks in general. Plsi can correct some of the drawbacks arising from softwarization by moving critical services towards the physical layer, thereby increasing its flexibility. While it is obvious that future networks will in addition to signal generation, transmission and detection also need to generate entanglement, a novel task which is being motivated in the post-Shannon context is also the generation and distribution of randomness. Fast random number generation can be used as an input to physical layer network coding [42]. The availability of entanglement and sometimes also shared randomness even allows the execution of novel security primitives such as oblivious transfer [43, 44], the origin and historical developments of which are well described in [45]. An oblivious transfer allows the secure computation of functions between network entities and reduces the need for a sharing of private data in such applications. With regards to the security assumptions, there are a variety of proposals. The information-theoretic perspective applied to the executions of oblivious transfer over single-hop classical network connections is described well in [46]. Recent work [47] claims a protocol for achieving positive oblivious transfer rates even in situations where the participating parties can be dishonest. Despite drawbacks, vulnerabilities, and the development of impossibility results [48], novel methods for the execution of oblivious transfer protocols have also been researched and proposed in quantum communication [49], such that the possibility of secure computation over communication networks remains.

Once the generation of distributed randomness (which can also be harvested from entanglement) is a physical layer service, the network nodes sharing this resource benefit from increased robustness against jamming and Denial of Service (DoS) attacks. These effects have been highlighted in the quantum information-theoretic literature on arbitrarily varying (quantum) channels [50–52] and motivated research on plsi. The increased resilience is vital for wireless links of critical infrastructures, such as in-campus networks. A crucial assumption in the aforementioned line of arbitrarily varying channel models is the existence of shared randomness which is unknown to a potential attacker (secret, for short). A possible way to satisfying such assumptions is to distribute entanglement through the network. The value of distributed shared (secret) randomness is best understood from the recent work [53]. In [53], the problem of detecting DoS attacks, for example, on wireless networks, has been formulated and analysed using the formalism of (classical) arbitrarily varying channels. As it turns out, deciding whether a DoS attack is possible on a given wireless link is not possible in general. Still, scenarios where a DoS attack is not possible can be detected. If shared secret randomness is available for a given system, its capacity can be computed using standard techniques. Thus, the question of how to

assure the quality of secret shared randomness is vital for resilient communication.

Following from the work on OSI network conformal quantum networks as described in Subsection 3.2.1, physical layer service integration also encompasses efficient error correction in the transmission of classical data. To build on the concrete example given in [35] where quantum communication techniques were proven to reduce energy consumption in long-haul fiber transmission, an obvious and simple example of physical layer service integration is fully optical error correction in optical fiber transmission, where the task of error correction is handled by the physical layer. This approach ultimately allows building receivers attaining the data transmission capacity of any physical medium. Depending on the boundary conditions in terms of bandwidth, signal energy, or transmission range, quantum receivers can beat their classical counterparts by orders of magnitude.

In order to finally give a concrete historical example of successful plsi, we point to all-optical networks. In optical long-haul fiber transmission, a choice exists regarding the method of signal regeneration. In particular, a design choice can be made between opto-electronic conversion including error correction and feedback methods between nodes placed along a link connecting a sender end receiver, or fully optical amplification. The vision of all-optical networks [54], built on the latter approach, corrects the signal to noise ratio while leaving the signal in the optical domain and without applying error correction steps. In this sense, plsi is a design choice that with examples of successful application existing already today.

3.4. Spatial Structure of Current and Future Networks.

A final important architectural aspect to be discussed is the highly-growing interest for aerial and satellite platforms' integration into the terrestrial quantum communication networks. This is mainly due to one of the major issues that are affecting the research on and the design of the quantum internet: the need for quantum repeaters. Quantum communications in fibres can reach a distance of about 100-150 km. Then, in order to maintain the fidelity of entanglement and to avoid decoherence, devices like quantum repeaters have to be employed. However, these devices are still under research, and they are highly complex and expensive. Around 2014, the research started to focus on qkd and entanglement distribution via satellites in order to achieve intercontinental communications more easily. A recent record was set when entangled photon pairs were distributed via two bidirectional downlinks from the Micius satellite to two ground observatories in Delingha and Nanshan in China [55]. In the last years, the scope has been moving to ensuring miniaturisation, lower costs, and lower orbits [56]. These advantages are necessary for a subsequent and seamless integration with the terrestrial network. However, the realization of entangled quantum systems and their distribution to any node on earth via nanosatellites (e.g., Cube-Sats) is still an open research challenge. Distributing entanglement via nanosatellites would significantly reduce the need and the cost of repeaters. That is why the three-

dimensional networking has a crucial role also in quantum communication networks.

3.5. Key Performance Indicators. In the following, we give a short overview over the most relevant Key Performance Indicators of quantum channels. As most systems are in a technically premature stage, we focus on those derived from information theory. In contrast to classical communication systems, where the Shannon capacity of a channel was the dominant metric over several decades, research on quantum communication systems identified three different capacities early on. In the context of post-Shannon literature, several additional KPIs were discovered both for quantum and for classical channels. The simplest KPI for a quantum channel is its message transmission capacity [36, 37]. A second KPI is the entanglement-assisted message transmission capacity which is based on the idea of dense coding [57]. For this type of data transmission, it is assumed that entanglement has been established between sender and receiver which is available to both of them at the time where the data needs to be transmitted. The third fundamental KPI of a quantum channel is its capacity for transmitting entanglement [58–60]. This latter capacity is also referred to as the quantum capacity of a quantum channel. It is an important open problem to derive exact formulas for the latter capacity, which is not even known for very simple transmission systems yet. In addition, the quantum capacities of quantum channels have been derived under the assumption of DoS attacks in [50] and under the assumption of incomplete information regarding critical system parameters in [61]. Channels with memory have been studied in [?], and the second-order behaviour which is relevant to the performance with finite blocklength in works such as [62, 63]. The effect of fading on quantum channels has been studied in [64, 65]. Finally, the identification capacity of a classical channel has been derived in [66], and formulas for several quantum communication systems have been proven [67, 68]. The identification capacity is a KPI from the domain of post-Shannon information theory, which describes the number of messages per channel use that can be achieved in situations where the receiver is only interested in the question whether or not the incoming message was intended for him. It should be noted that a direct technological comparison of the different techniques that are theoretically available from quantum technology development with those existing as a state-of-the-art is typically not available in a systematic fashion due to a lack of readily available components on the quantum technology side.

4. An Architecture for Future Classical-Quantum Communication Networks

In Section 3.2, the legacy quantum network architectures have been outlined. The question arises how to reconcile these architectures with current trends. Of particular importance is the decoupling between network functionalities and hardware through softwarization, which has opened the way to extensive employment of in-network intelligence, easier network management and upgrading, and to efficient and effective management of multitenancy.

The novel architecture this work proposes leverages the idea of network softwarization realized via SDN and NFV, in order to achieve a better integration of quantum communication networks with upcoming and future generation networks. Furthermore, it also inherit the pros of network softwarization that have been just mentioned, opening the way for a more flexible and advanced quantum-classical network management and operations. Figure 9 depicts the logic architecture that this work is going to propose. The network infrastructure consists of a hybrid quantum-classical network, combining 4G and 5G technologies, and the three-dimensional 6G communication networks with a quantum physical layer. As seen in the previous sections, the three-dimensionality is pivotal both for quantum and classical communication networks. Next, the end-to-end management and orchestration has to handle a hybrid three-dimensional infrastructure considering data and control planes with quantum capabilities as well.

As mentioned previously, the Internet is evolving towards a SDN approach, as modern networks are incrementally deploying SDN to facilitate configuration, operation, and automated management. In this framework, an interesting direction might be represented by introducing the detachment between control and data planes at the basis of the SDN paradigm and extend it to quantum communications. This integrated classical-quantum Internet architecture would enable the deployment of classical as well as quantum link level communication technologies and devices that will constitute the data plane of the converged infrastructure, while maintaining a "traditional" control plane functionality in the form of an evolved SDN controller—be that centralized or distributed. The relevant advantage of this solution, which in our considerations would be preferable, is the convergence of quantum-plus-traditional Internet as a single-integrated and single-managed entity that would enable faster integration of quantum technologies within modern networks.

Then, it is pivotal to maintain the current separation between data and control plane, envisioned by SDN. In this sense, the control plane will not only manage the classical protocol stack but also the quantum physical-link layer resources. This will also advance the hybrid quantum-classical protocol stack towards the idea of programmability, which means the capability of adapting the hybrid protocol stack according to network conditions with the possibility of slicing quantum communication resources as well. This will highly enforce the coexistence and the progressive introduction of quantum technologies within future communication networks. An important consideration refers to the design of the interfaces between the upper softwarized layers and the hybrid physical classical-quantum layer. Especially southbound, but also the northbound interface, will require programming that takes into account the different algorithmic implementations that quantum resources demand. Furthermore, the inherent cross-layer characteristics of entanglement, mentioned in Section 3.1, require a specific
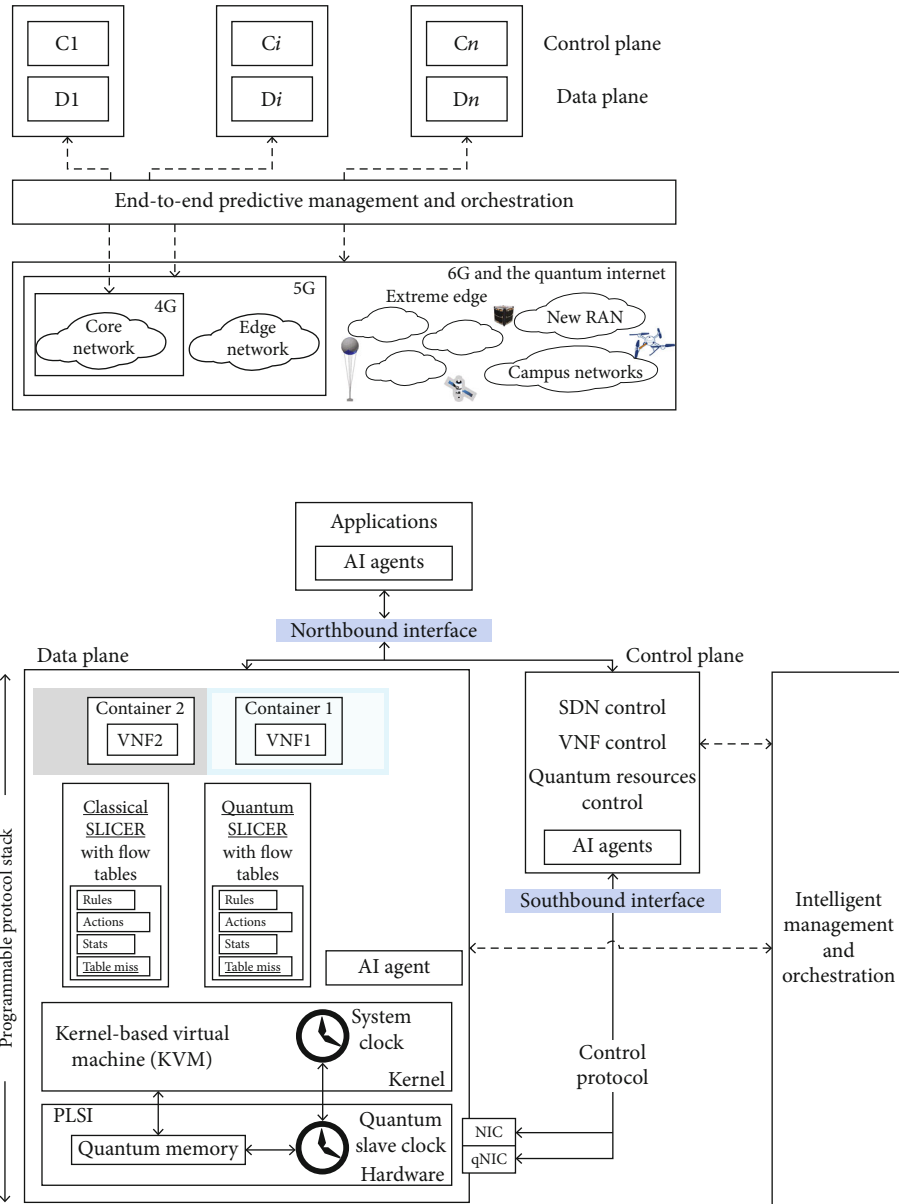
FIGURE 9: High-level schematic of the proposed quantum-classical network architecture, where control and data plane reside in all the devices. The network infrastructure is considered a continuum by the softwarized layers and by the management and orchestration, which can also employ intelligence (above). Logic architecture of quantum-classical communication network that has been proposed (below). The qNIC is the quantum network interface. The slave clock is the only one represented since the master clock can be either placed in another network node or in the controller according to the specific characteristics of the infrastructure.

design not only of link-layer but also of network-layer protocols and saps.

A distributed control plane can be preferable in respect of a centralized one since procedures like entanglement swapping for longer-distance entanglement distribution will imply significant communication control overhead. Then, controllers can be moved close to areas (for example geographically bounded campus networks or low latency verticals like the tactile internet) or routes where entanglement distribution is needed most. It should be taken into account that the controller will be a virtual function dynamically placed

in data centers. Thus, in our proposed logic architecture, computing and the role of data centers is pivotal for quantum communications. Entanglement creates a physically distributed state among network nodes, and it is not a stateless operation [4]. This means that the architectural choice of a control plane, either centralized or distributed, can be very effective to manage these aspects and to keep spatial-temporal track of quantum network states.

Next, the control protocol and the interfaces between control and data plane have a key role. This can be a flow-oriented protocol and considering ports abstraction for

managing the quantum and classical data flows. This flow-management is possible by setting up the flow tables and their internal rules. These flow tables refer to both quantum and classical resources' assignment and allow for potential slicing and multitenancy. This novel proposed architectural element also has the capability of concurrently programming multiple network devices. Thus, in respect of other proposed architectures in the literature, this has a "network viewpoint," which can enable a more efficient and effective entanglement distribution and quality-assurance, in parallel to a more effective control of network resources and paths (also leveraging the programmability for future employment of intelligence).

The interfaces that come into our quantum-classical architectural view and that are missed in existing quantum network architectures are southbound, east/westbound, and northbound. First, the southbound control interface on quantum communication devices, which should integrate a proper interface to enable the control plane to issue commands and provide status information. Second, the update of the SDN southbound interface. The communication between the control and data plane is implemented using the southbound interface. In the most common cases, such interface might be provided by protocols such as OpenFlow or languages such as P4. Those technologies should be enhanced to support quantum-specific operation and to enable devices to advertise their quantum or traditional features.

Regarding the east/westbound interface, these are internal to the control plane, and they become pivotal when the control plane is distributed. In this latter case, it is necessary another protocol to manage the distributed control of quantum-classical resources and data plane in general. Finally, the northbound interface has to be included (and its classical version modified) so that classical software applications can leverage the quantum effects at the physical and link layers.

Backward compatibility could be enabled by introducing small modifications to the features' discovery process of SDN. Indeed, features' discovery represents a mandatory process to enable SDN controllers to learn the specific characteristics of the network devices. For example, in the case of OpenFlow, after the initial handshake implemented over TCP/SSH using the HELLO messages, the SDN controller will issue a FEATURES_REQUEST message in order to acquire the functionalities supported by the OpenFlow switch. To support quantum communications, the FEATURES_REQUEST and FEATURES_REPLY messages should be enhanced to support quantum-specific operation.

As it is possible to see from Section 3.1, the definition of an architecture and a protocol stack for quantum communication networks is currently neglecting a major aspect of future 6G networks: ai. However, some works have been underlining the importance of quantum communication networks for achieving efficiently in-network intelligence [69], together with the new potentials of opened by quantum machine learning for the management and orchestration of future hybrid quantum-classical communication networks [70]. Then, Figure 9 also tries to envision the placement of

hybrid intelligence in the quantum-classical network architecture. The architectural integration of classical and quantum communication technologies implies the "hybrid" collaboration of classical and quantum intelligent agents. Since a great part of the control of the quantum networks will be performed classically, classical in-network intelligence will play a key role for management and orchestration. However, in this context, the decision-making and the prediction can also be based on potential new intelligent algorithms based on quantum machine learning, running in quantum data centers. Next, within the control plane and the data plane nodes, ai agents can perform specific tasks, exploiting either the classical or the quantum physical and link layers. This mainly depends on the technological advancements that will be achieved by quantum computing for communication nodes in the next ten to fifteen years. If the miniaturisation and costs of quantum computing will not reach reasonable thresholds, the role of quantum machine learning in 6G will be limited to centralized management and orchestration. Finally, intelligent agents can also be realized at the application layers, for paradigms like ai-as-a-service. In this context, the employment of "quantum" intelligent agents is going to take longer time since the maturity of "quantum" optimized software for fully-quantum hardware represents a long-term objective.

In legacy proposed architectures, a quantum link layer manages the mapping of entangled photons to entangled qubits, and it guarantees the overall quality of the communication. One possibility to achieve integration between quantum and traditional link layer technologies might be to consider quantum as another layer 1/layer 2 technology and to achieve loose integration by exploiting the interoperability provided by the Internet Protocol. Such a solution would require adapting TCP/IP lower layers to the needs of quantum communications. This mean introducing a proper version of the Address Resolution Protocol (ARP) as well as additional functionalities between IP and the quantum layers in order to support entanglement and channel setup functionalities.

An alternative might be to enable the integration of quantum-powered autonomous systems within the overall Internet topology by exploiting IP tunnelling functionalities. This alternative might represent a mature solution, considering the diffusion of tunnelling protocols for Virtual Private Networks (VPNs) and IPv6-IPv4 tunnelling. Quantum technologies would represent a specific physical-link layer technology operating under a network operator, connected to the rest of the Internet via tunnel end-points that would hide the specific characteristics of the quantum internals. Tighter integration might consider the possibility of enabling classical and quantum operation on the same physical link. This would require adaptation of classical mac protocols to this novel functionality or the introduction of an "upper" mac scheduler capable of facilitating the coexistence.

The advantage of both the above solutions would be to enable quantum communications to freely evolve as a separate standard providing methods for interoperating with other standards at the higher network layers without specific constraints from the rest of the public Internet while

incrementally enabling its implementation in parallel to innovations in Internet-related technologies.

In this new architecture, plsi as described in Subsection 3.2 can be used to meet specific demands in peak data rates, resilient low-latency communication, or secure network function computation.

Another service that can be considered as plsi is the network time synchronisation. Classically, this task is performed at layer 2 and layer 3 (as previously described). From synchronisation viewpoint, the architecture considers the presence of a quantum master clock, which transmits ultraprecise (in the order of picoseconds or below) timing information to its quantum slave clocks. The choice of master clocks and their placement is out of the scope of this work. However, it is important to notice that an example scenario may consider the placement of a master quantum clock for local area networks of slaves directly connected with the master.

Depending on the desired network hardware and structure, not all tasks can be considered in plsi. For example, the gains from entanglement-assisted data transmission depend both on entanglement storage times and the communication medium, while network time synchronisation is affected most by the topology and to some degree on storage times.

5. Conclusion

This article has provided a novel architecture for future quantum-classical networks, which leverages the trend of evolution of upcoming and future generation networks. While the legacy proposed architectures for quantum communications are limited to some perspectives, the one envisioned in this work considers the lessons learnt from in-network computing, network virtualization, and programmable stacks. Moreover, it considers the employment of quantum communications for network synchronisation operations, which are pillars of communications functionalities of the whole protocol stack. In order to justify our proposal and to compare it to the ones in the state-of-the-art, an introductory part on the important and related aspects of classical networks was provided. Next, the discussion also embraces research and standardisation status in order to show the differences between classical and quantum communications research and standardisation efforts. This is also important to show how our architecture brings together the different research trends and works by the different standardisation bodies.

From the explanation of the proposed architecture, it is possible to highlight important fields of research and standardisation, both in academia and industry. For example, the design, standardisation, and development of control protocols and architectural interfaces as have been done for SDN and NFV for classical networks. Next, the study and realization of quantum-classical network slicing and multitenancy. This will be fundamental in future scenarios with coexistence of heterogeneous services managed by multiple operators on the same network infrastructure.

An important role in such proposed architecture and in general in future quantum-classical networks will have data centers, both in the edge and in the cloud. Even the research of where and how to place the control of quantum network functions in data centers' network is still unknown. In particular, this will have to consider that in some cases (e.g., the tactile internet), entanglement generation and distribution will have to satisfy low latency, thus, being performed at the edge or access networks.

Next, the detailed design, analysis, and realization of centralized and distributed control planes are still in its infancy. This is also true for the various instances that may arise for the realization of quantum-classical pps, and, more generally, for quantum-classical network operating systems and software-defined protocols.

These are just few open fundamental challenges that our novel architecture opens to guarantee a seamless integration, operation, and management of future quantum-classical networks.

Data Availability

Data available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] F. H. P. Fitzek, F. Granelli, and P. Seeling, *Computing in Communication Networks - From Theory to Practice*, vol. 1Elsevier, 1st ed., ser. 1 edition, 2020.

[2] R. Bassoli, H. Boche, C. Deppe et al., *Quantum Communication Networks*, Springer, 1st ed. edition, 2021.

[3] R. Bassoli, F. H. Fitzek, and E. Calvanese Strinati, "Why do we need 6G?," *ITU Journal on Future and Evolving Technologies*, vol. 2, no. 6, pp. 1–31, 2021.

[4] W. Kozlowski, S. Wehner, R. Van Meter et al., *IETF quantum internet research group (qirg): Architectural Principles for a Quantum Internet*, International Telecommunication Union, Geneva, 2020.

[5] M. Lévesque and D. Tipper, "A survey of clock synchronization over packet-switched networks," *IEEE Communication Surveys and Tutorials*, vol. 18, no. 4, pp. 2926–2947, 2016.

[6] ITU-T, "G.8271.1/Y.1366.1 – series G: transmission systems and media, digital systems and networks. Packet over transport aspects – synchronization, quality and availability targets," 2020, https://www.itu.int/rec/T-REC-G.8271.1/en.

[7] "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," in *IEEE Std 1588-2008 (Revision of IEEE Std 1588-2002)*, pp. 1–269, 2008.

[8] A. S. Abdalla and V. Marojevic, "Communications standards for unmanned aircraft systems: the 3GPP perspective and research drivers," *IEEE Communications Standards Magazine*, vol. 5, no. 1, pp. 70–77, 2021.

[9] 3GPP, "3GPP TR 22.829 – 3rd generation partnership project; technical specification group services and system aspects; enhancement for unmanned aerial vehicles; stage 1 (release 17)," 2019, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?speci_cationId=3557.

[10] Hexa-X, "D5.1 – initial 6G architectural components and enablers," 2021, https://hexa-x.eu/wp-content/uploads/2022/01/Hexa-X.

[11] H. Zimmermann, "OSI reference model–the ISO model of architecture for open systems interconnection," *IEEE Transactions on Communications*, vol. 28, no. 4, pp. 425–432, 1980.

[12] B. E. Carpenter, *Architectural Principles of the Internet*, vol. 1958 of Internet Requests for Comments, , RFC, 1996.

[13] D. Meyer and G. Zobrist, "TCP/IP versus OSI," *IEEE Potentials*, vol. 9, no. 1, pp. 16–19, 1990.

[14] J. D. Day and H. Zimmermann, "The OSI reference model," *Proceedings of the IEEE*, vol. 71, no. 12, pp. 1334–1340, 1983.

[15] F. M. Burg and N. Di Iorio, "Networking of networks: interworking according to OSI," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 7, pp. 1131–1142, 1989.

[16] M. P. Clark, "The ATM protocol reference model," in *ATM Networks*, Verlag, 1996.

[17] L. Staalhagen, "A comparison between the OSI reference model and the B-ISDN protocol reference model," *IEEE Network*, vol. 10, no. 1, pp. 24–33, 1996.

[18] A. Viswanathan, E. C. Rosen, and R. Callon, *Multiprotocol Label Switching Architecture*, RFC 3031, 2001.

[19] A. S. Tanenbaum and D. Wetherall, *Computer Networks*, Pearson, 5th Edition. edition, 2011, https://www.worldcat.org/oclc/698581231.

[20] D. Kliazovich, F. Granelli, and N. L. S. Fonseca, "Architectures and cross- layer design for cognitive networks," in *Handbook on Sensor Networks*, Y. Xiao and H. Chen, Eds., pp. 3–24, Idea, 2010.

[21] Z. Guan, L. Bertizzolo, E. Demirors, and T. Melodia, "WNOS: an optimization-based wireless network operating system," *CoRR*, vol. abs/1712.08667, 2017http://arxiv.org/abs/1712.08667.

[22] R. Wen, G. Feng, W. Tan, R. Ni, S. Qin, and G. Wang, "Protocol function block mapping of software defined protocol for 5G Mobile networks," *IEEE Transactions on Mobile Computing*, vol. 17, no. 7, pp. 1651–1665, 2018.

[23] ONOS, "Micro ONOS," 2020, December 2020, https://docs.onosproject.org/.

[24] P4, "P4 consortium," 2020, December 2020, https://p4.org/.

[25] Hexa-X, "D4.1 – AI-driven communication & computation co-design: gap analysis and blueprint," 2021, https://hexa-x.eu/wp-content/uploads/2021/09/Hexa-X-D4.1_v1.0.pdf.

[26] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: the road to Pareto-optimal wireless networks," *IEEE Communication Surveys and Tutorials*, vol. 22, no. 3, pp. 1472–1514, 2020.

[27] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: a vision for the road ahead," *Science*, vol. 362, no. 6412, 2018https://science.sciencemag.org/content/362/6412/eaam9288.

[28] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, and G. Bianchi, "Quantum internet: networking challenges in distributed quantum computing," *IEEE Network*, vol. 34, no. 1, pp. 137–143, 2020.

[29] A. Dahlberg, M. Skrzypczyk, T. Coopmans et al., "A Link Layer Protocol for Quantum Networks," in *Proceedings of the ACM Special Interest Group on Data Communication, ser. SIGCOMM '19*, pp. 159–173, New York, NY, USA, 2019.

[30] A. Pirker and W. Dür, "A quantum network stack and protocols for reliable entanglement-based networks," *New Journal of Physics*, vol. 21, no. 3, article 033003, 2019.

[31] J. Nötzel and S. DiAdamo, "Entanglement-enhanced communication networks," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pp. 242–248, Denver, CO, USA, 2020.

[32] J. Nötzel and S. DiAdamo, "Entanglement-assisted data transmission as an enabling technology: a link-layer perspective," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1955–1960, Los Angeles, CA, USA, 2020.

[33] S. Guha, Q. Zhuang, and B. A. Bash, "Infinite-fold enhancement in communications capacity using pre-shared entanglement," in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1835–1839, Los Angeles, CA, USA, 2020.

[34] J. Nötzel, "Entanglement-enabled communication," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 2, pp. 401–415, 2020.

[35] J. Nötzel and M. Rosati, "Reducing energy consumption of fiber networks via quantum communication technology," 2022, https://arxiv.org/abs/2201.12397.

[36] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, vol. 9, no. 3, pp. 177–183, 1973, http://mi.mathnet.ru/ppi903.

[37] A. S. Holevo, "Coding theorems for quantum communication channels," in *Proceedings. 1998 IEEE International Symposium on Information Theory (Cat. No.98CH36252)*, p. 84, Cambridge, MA, USA, 1998.

[38] S. Guha, "Structured optical receivers to attain superadditive capacity and the Holevo limit," *Physical Review Letters*, vol. 106, no. 24, article 240502, 2011.

[39] M. Rosati, A. Mari, and V. Giovannetti, "Multiphase hadamard receivers for classical communication on lossy bosonic channels," *Physical Review A*, vol. 94, article 062325, 2016.

[40] K. Banaszek and M. Jachura, "Structured optical receivers for efficient deep-space communication," in *2017 IEEE International Conference on Space Optical Systems and Applications (ICSOS)*, pp. 34–37, Naha, Japan, 2017.

[41] K. Banaszek, L. Kunz, M. Jachura, and M. Jarzyna, "Quantum limits in optical communications," *Journal of Lightwave Technology*, vol. 38, no. 10, pp. 2741–2754, 2020.

[42] P. Chen, Z. Xie, Y. Fang, Z. Chen, S. Mumtaz, and J. J. P. C. Rodrigues, "Physical-layer network coding: an efficient technique for wireless communications," *IEEE Network*, vol. 34, no. 2, pp. 270–276, 2020.

[43] S. Wiesner, "Conjugate coding," *SIGACT News*, vol. 15, no. 1, pp. 78–88, 1983.

[44] M. Rabin, *How to exchange secrets by oblivious transfer*, Harvard University Technical Report 81, 1981.

[45] A. Broadbent and C. Schaffner, "Quantum cryptography beyond quantum key distribution," *Designs, Codes and Cryptography*, vol. 78, no. 1, pp. 351–382, 2016.

[46] A. C. A. Nascimento and A. Winter, "On the oblivious-transfer capacity of noisy resources," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2572–2581, 2008.

[47] F. Oggier and G. Zémor, "Coding constructions for efficient oblivious transfer from noisy channels," *IEEE Transactions on Information Theory*, vol. 68, no. 4, pp. 2719–2734, 2022.

[48] H.-K. Lo, "Insecurity of quantum secure computations," *Physical Review A*, vol. 56, no. 2, pp. 1154–1162, 1997.

[49] Y.-H. Chou, G.-J. Zeng, and S.-Y. Kuo, "One-out-of-two quantum oblivious transfer based on nonorthogonal states," *Scientific Reports*, vol. 8, no. 1, p. 15927, 2018.

[50] R. Ahlswede, I. Bjelaković, H. Boche, and J. Nötzel, "Quantum capacity under adversarial quantum noise: arbitrarily varying quantum channels," *Communications in Mathematical Physics*, vol. 317, no. 1, pp. 103–156, 2013.

[51] H. Boche and J. Notzel, "Arbitrarily small amounts of correlation for arbitrarily varying quantum channels," *Journal of Mathematical Physics*, vol. 54, pp. 734–738, 2013.

[52] H. Boche, C. Deppe, J. Nötzel, and A. Winter, "Fully quantum arbitrarily varying channels: random coding capacity and capacity dichotomy," in *2018 IEEE International Symposium on Information Theory (ISIT)*, pp. 2012–2016, Vail, CO, USA, 2018.

[53] H. Boche, R. F. Schaefer, and H. V. Poor, "Denial-of-service attacks on communication systems: detectability and jammer knowledge," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3754–3768, 2020.

[54] A. A. M. Saleh and J. M. Simmons, "All-optical networking—evolution, benefits, challenges, and future vision," *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1105–1117, 2012.

[55] J. Yin, Y.-H. Li, S.-K. Liao et al., "Entanglement-based secure quantum cryptography over 1,120 kilometres," *Nature*, vol. 582, no. 7813, pp. 501–505, 2020.

[56] D. K. Oi, A. Ling, G. Vallone et al., "CubeSat quantum communications mission," *EPJ Quantum Technology*, vol. 4, no. 1, 2017.

[57] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Physical Review Letters*, vol. 69, no. 20, pp. 2881–2884, 1992.

[58] S. Lloyd, "Capacity of the noisy quantum channel," *Physical Review A*, vol. 55, no. 3, pp. 1613–1622, 1997.

[59] P. W. Shor, "The quantum channel capacity and coherent information," in *Lecture notes, MSRI Workshop on Quantum Computation*, Springer, 2002.

[60] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 44–55, 2005.

[61] I. Bjelaković, H. Boche, and J. Nötzel, "Entanglement transmission and generation under channel uncertainty: universal quantum channel coding," *Communications in Mathematical Physics*, vol. 292, no. 1, pp. 55–97, 2009.

[62] W. Matthews and S. Wehner, "Finite blocklength converse bounds for quantum channels," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7317–7329, 2014.

[63] N. Datta, M. Tomamichel, and M. M. Wilde, "On the second-order asymptotics for entanglement-assisted communication," *Quantum Information Processing*, vol. 15, no. 6, pp. 2569–2591, 2016.

[64] D. Vasylyev, A. A. Semenov, W. Vogel et al., "Free-space quantum links under diverse weather conditions,," *Physical Review A*, vol. 96, article 043856, 2017.

[65] D. Dequal, L. T. Vidarte, V. R. Rodriguez, P. V. Giuseppe Vallone, A. Leverrier, and E. Diamanti, "Feasibility of satellite-to-ground continuous-variable quantum key distribution," *npj Quantum Information*, vol. 7, no. 3, 2021.

[66] R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 15–29, 1989.

[67] P. Löber, "Quantum channels and simultaneous ID coding," 1999, https://arxiv.org/abs/quant-ph/9907019.

[68] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 569–579, 2002.

[69] R. Ferrara, R. Bassoli, C. Deppe, F. H. P. Fitzek, and H. Boche, "The computational and latency advantage of quantum communication networks," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 132–137, 2021.

[70] S. J. Nawaz, S. K. Sharma, S. Wyne, M. N. Patwary, and M. Asaduzzaman, "Quantum machine learning for 6G communication networks: state-of the-art and vision for the future," *IEEE Access*, vol. 7, pp. 46317–46350, 2019.