

# A Glimpse of Physical Layer Decision Mechanisms: Facts, Challenges, and Remedies

SELEN GEGGEL CETIN<sup>1</sup> (Student Member, IEEE), CANER GOZTEPE<sup>1</sup>,  
GUNES KARABULUT KURT<sup>2</sup> (Senior Member, IEEE), AND HALIM YANIKOMERGLU<sup>3</sup> (Fellow, IEEE)

<sup>1</sup>Department of Electronic and Communications Engineering, Istanbul Technical University, 34469 Istanbul, Turkey

<sup>2</sup>Poly-Grames Research Center, Department of Electrical Engineering, Polytechnique Montreal, Montreal, QC H3T 1J4, Canada

<sup>3</sup>Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada

CORRESPONDING AUTHOR: S. G. CETIN (e-mail: gecgel16@itu.edu.tr)

**ABSTRACT** Communications are realized as a result of successive decisions at the physical layer, from modulation selection to multi-antenna strategy, and each decision affects the performance of the communication systems. Future communication systems must include extensive capabilities as they will encompass a wide variety of devices and applications. Conventional physical layer decision mechanisms may not meet these requirements, as they are often based on impractical and oversimplifying assumptions that result in a trade-off between complexity and efficiency. By leveraging past experiences, learning-driven designs are promising solutions to present a resilient decision mechanism and enable rapid response even under exceptional circumstances. The corresponding design solutions should evolve following the lines of learning-driven paradigms that offer more autonomy and robustness. This evolution must take place by considering the facts of real-world systems and without restraining assumptions. In this paper, the common assumptions in the physical layer are presented to highlight their discrepancies with practical systems. As a solution, learning algorithms are examined by considering the implementation steps and challenges. Furthermore, these issues are discussed through a real-time case study using software-defined radio nodes to demonstrate the potential performance improvement. A cyber-physical framework is presented to incorporate future remedies.

**INDEX TERMS** Cybertwin, decision mechanisms, learning-driven solutions, machine learning, physical layer, real-world impairments.

## I. INTRODUCTION

TOWARDS the sixth-generation (6G) networks, flexible and ubiquitous connectivity is expected, even under extraordinary conditions. Numerous technologies are envisioned to achieve this goal. On the other hand, heterogeneity of application domains for these technologies constitutes a significant need for more customized deployments. Therefore, stringent physical layer (PHY) requirements are emerging in terms of fast-responsiveness, reliability, latency, spectral efficiency, and security [1]–[3].

A typical digital communication system contains several signal processing blocks at the transmitter and receiver sides, such as equalization, bandpass signaling, channel coding, multiplexing and multiple access. Therefore, communication

systems at the PHY can be defined as consecutive decision mechanisms that separately handle a subproblem and find appropriate decisions for signal processing blocks. Each decision, from transmit power to coding rate, jointly ensures efficient transmission of information bits. Although optimal decisions are theoretically attainable for the most part of problems, they may not be available under real-world impairments such as correlations or delays, as demonstrated in Fig. 1. Theoretical implications are generally based on assumptions that benefit from the advantages of simplicity, such as effortlessness to achieve an understandable formulation, perceptibility, and interpretability. They allow us to discover the theoretical bounds of the proposed system and give rise to diversify sub-optimal methods by compromising

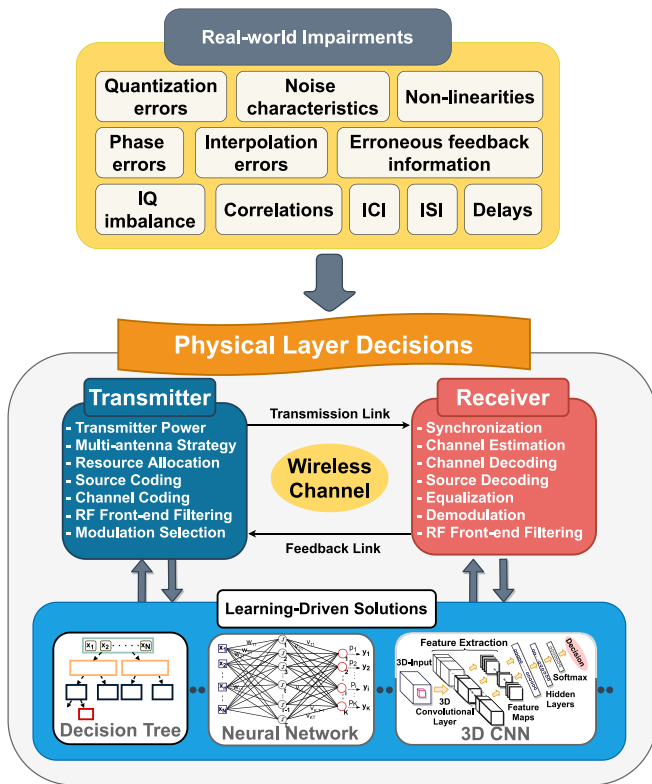


FIGURE 1. A demonstration of major decision steps in PHY layer and the interaction of learning-driven solutions.

on a number of necessity. On the contrary, modeling mismatches due to these assumptions distract the solutions from actual systems, especially in real-world implementations. Current decision mechanisms for wireless communications are largely modeled under oversimplifying assumptions, as listed in Table 1. However, future implementations will require re-evaluation of the solutions for PHY decision mechanisms to meet demanding targets.

Over the past decade, learning-driven approaches have been considered the leading candidates to achieve the ambitious goals of next-generation communication systems. Several studies are motivated with the advantages of learning-driven solutions, such as the elimination of human intervention and the use of big data stacks in PHY. In [4], three machine learning algorithms are separately applied for channel assignment problem to evade the high complexity of the convex optimization based algorithm. The results show that the time complexity can be reduced without making concessions from the prediction accuracy. The study in [5] compares the performance of machine learning assisted and the conventional threshold-based link adaptation schemes. The learning-driven scheme outperforms its counterpart by achieving higher throughput. In [6], a deep neural network is integrated into Viterbi algorithm instead of log-likelihood computations and it can be reliably employed in complex channel models thanks to its dynamic channel

tracking capability. The results indicate the robustness, success, and adaptability of the learning-driven solutions. In [7], power control, interference coordination, and beamforming are enabled by using deep reinforcement learning without the knowledge of channel state information (CSI). This study shows that learning algorithms can provide competent systems capable of making joint decisions with instantly responsive designs.

The aforementioned studies seem highly incentive for further researches, but offer no guarantee that the same results will be obtained under real conditions. The fact that learning-driven solutions do not outperform the optimal solution in the presence of ideal conditions should be kept in mind. However, instead of solutions based on impractical assumptions, they can provide solutions that go hand in hand with environmental changes and are based on real systems' attributes [8]–[12]. In [8], neural network topologies are examined for non-coherent demodulation, which is convenient but difficult to model optimally for practical wireless communication systems due to non-linearities, non-stationarity, and non-Gaussian noise. The authors of [9] sign the complex channel conditions without a mathematically tractable model and design the communication system as an autoencoder. Adaptive transmission and generalized data representation schemes are proposed to maximize the data rate under different channel conditions. Block error rate and minimum mean squared error performance of the schemes are inspiring at lower signal-to-noise ratio (SNR). The study in [10] compares learning-driven algorithms with the classical Euclidean distance-based method for the antenna selection problem considering channel imperfections and correlation. The results over a real-time test bed show that learning-driven decision mechanisms are also able to track correlation and deal with channel imperfections besides improving error performance. In [11], a PHY security problem is addressed, and neural networks acting as detector and identifier show high accuracy under actual conditions. The authors of [12] address the self-interference problem of flexible duplexing that take place in specifications. The study applies machine learning to the tuning process of radio frequency (RF) canceller and achieves the fastest convergence.

The accomplishments of above studies evince that learning algorithms are not only an alternative to their classical counterparts, but also very well suited to deal with PHY challenges. Most part of the studies analyze the deep-learning algorithms by presenting challenges and applications [3], [13], [14]. In this paper, an emerging necessity in PHY, a *changeover to learning-driven decision mechanisms* is highlighted. To this end, our major contributions are listed below:

- Design issues for PHY decision mechanisms are addressed under five aspects: Synchronization, channel estimation errors, erroneous feedback information, RF front-end impairments, and correlation. The main incompatibilities between the real systems and the

**TABLE 1.** An overview of the simplifying assumptions and challenging facts in PHY.

Technical Aspects	Issues	Assumptions	Facts
<b>Synchronization</b>	<i>Initial Acquisition</i>	- Prior probabilities and pre-determined threshold are always valid for hypothesis tests.	- False alarm can occur when there is no reference signal.
	<i>Carrier Phase</i>	- No ISI. - Channel provides the Nyquist criterion. - Estimated phase variance satisfies Cramer-Rao bound (then optimal).	- ISI can be realized in practice: ISI suppression should be added to approximate the assumption.
	<i>Carrier Frequency Offset</i>	- The carrier phase is constant or does not change significantly. - No Doppler shift. - Channel satisfies Nyquist criteria.	- The carrier phase can change in time.
	<i>Timing Offset</i>	- No ISI. - No sampling delay. - Channel conditions are known.	- ISI generally exists. - CSI is estimated with errors due to TO. - Delays can not be eliminated entirely.
<b>Channel Estimation Errors</b>	<i>Quantization Errors</i>	- The optimum threshold to quantize is known. - The quantization error is uniformly distributed. - ADCs and DACs operate ideally and do not introduce any non-linearities for channel estimation and signal detection.	- The converters in the transceiver may have different resolutions. - The threshold can change depending on the environmental conditions. - Quantization noise may have different statistical characteristics due to its time-varying nature.
	<i>Interpolation Errors</i>	- Perfect knowledge of signal autocorrelation and noise variance is available (then Wiener interpolator is optimal).	- Excellent knowledge about the signal and noise is not attainable.
<b>Erroneous Feedback Information</b>	<i>Feedback Errors</i>	- Reverse (feedback) channel is noise-free. - Feedback channel is not bandwidth limited. - Feedback overhead can be negligible. - Reverse link has the same statistical reciprocity. - Receiver knows perfectly the forward channel or has knowledge related to feedback errors.	- The equalizer may be unstable. - The channel coefficients can change in time. - The channel coefficients can be correlated in time, frequency, and space.
	<i>Feedback Delays</i>	- There are no feedback delays. - The forward channel and feedback channel delays are known.	- The instantaneous feedback information is not available.
<b>RF Front-end Impairments</b>	<i>IQ Imbalance</i>	- No imbalance between I and Q.	- IQ imbalance leads to performance degradations due to phase rotations and signal distortions.
	<i>Phase Noise</i>	- No frequency deviation at the output of the RF oscillator. - The higher carrier frequencies do not increase the phase noise.	- RF oscillators are low-cost devices and have instabilities. - Higher carrier frequencies lead to an increase in errors and residual terms.
	<i>Non-linearities</i>	- The electronic units work ideally.	- The perfect stability is unrealistic for active devices.
<b>Correlation</b>	<i>Time Correlation</i>	- Information sources are statistically stationary, wide-sense stationary, ergodic, or cyclostationary random processes. - The channel is merely time-invariant. - The position of the transmitter and the receiver do not change.	- The stationarity conditions can not always exist practically. - The channel characteristics can vary because of the external factors and mobility but show mnemonic relationships.
	<i>Spatial Correlation</i>	- The channel correlation matrix of the users is equal to zero. - If present, the correlation between MIMO channels is constant. - The receiver or transmitter moves slowly, and the received signals are not impressed by shadowing.	- Each channel varies instantaneously. - User channels may not be statistically independent. - The current scenarios include the high mobility and the ones will be in 3D.
	<i>Frequency Correlation</i>	- The frequency correlation of the channel is related to the power delay profile.	- The scattering is not always wide-sense stationary and uncorrelated.

underlying assumptions (used in the derivations and simulations) of the current theoretical implications are outlined for each aspect in Table 1 (Section II).

- We methodologically discuss the justifications to changeover from current methodologies to future methodologies and present the critical observations that must be taken into account for new solutions. The motivations why machine learning algorithms should be part of future decision mechanisms instead of some current approaches are addressed (Section III).
- We present an insightful roadmap that provides a perspective on the development of machine learning-driven PHY decision mechanisms for real communication systems (Section IV).
- In the spirit of fair analysis, we also discuss the main challenges for learning-driven solutions (Section V).
- The eligibility of learning-driven solutions for real systems is discussed through the problem of antenna selection over a test-bed of software defined radio (SDR) nodes. The discussions are extended for different analog-to-digital/digital-to-analog (ADC/DAC) resolutions (Section VI).
- Four learning algorithms are selected by considering the time and computational complexities. Performances of

decision tree (DTREE), multi-layer perceptron (MLP), random forest (RForest), and convolutional neural network (CNN) algorithms are investigated via the real-time SDR-based case study (Section VI).

- We present an intelligent cyber-physical framework towards tight operational requirements of 6G in the presence of stringent PHY constraints. We provide insight into how learning-driven solutions interoperate with the other remedies thanks to the proposed framework (Section VII).

The remainder of this article is organized as follows. Section II addresses the design issues under five aspects related to PHY decision mechanisms for real systems. Section III gives a methodological overview of current solutions and presents the reasons and motivations why future solutions should focus on machine learning algorithms. Section IV clarifies the roadmap for researchers to present more practical learning-driven solutions. Key challenges for learning-driven solutions are discussed in Section V. Section VI proposes four learning-driven solutions to the antenna selection problem and compares their performances with a classical approach by discussing the considerations from the previous sections. Section VII presents a novel framework that encompasses the 6G remedies and not only the learning-driven solutions. The paper concludes with Section VIII.

## II. FIVE DESIGN ASPECTS FOR PHY DECISION MECHANISMS

In this section, we present a projection of five compelling aspects to obtain superior PHY decisions considering real-world systems. Based on these aspects, simplifying assumptions and challenging facts of PHY are addressed and listed in Table 1.

### A. SYNCHRONIZATION

Synchronization is the major indispensability for the compatibility of the transmitter's and receiver's processes and for accurate transmission [15]. Focusing on PHY, the receiver must detect the suitable times to sample the transmitted signal and compensate the oscillators' phase and frequency errors depending on the received signal. A large number of studies assume that transmitter and receiver are perfectly synchronized and prove the success with numerical results. However, a residual synchronization error may remain in time and frequency due to the factual circumstances. Contrary to idealized systems, the real-world systems must include more qualifications due to the RF front-end impairments, mobility, variety of channel conditions, and delays [16], [17]. They can include several processes to reach a better synchronization, such as initial time synchronization between the transmitter and receiver, a robust lock mechanism, the estimation of carrier frequency offset (CFO) and timing offset (TO), algorithms for adaptable redundancy insert, and synchronization recovery.

Estimation plays a critical role in reaching tenable synchronization. Theoretically optimal data-assisted methods are presented, including the hypothesis test for the initial acquisition, maximum-likelihood based estimation algorithms for CFO, TO, and phase offset [16]–[29]. However, there may not be sufficient statistics for these methods without idealizing the systems. For example, the synchronization parameters may not be constant, or the variance of the synchronization error may not be exactly calculated due to random fluctuations caused by the phase noise of the oscillator. If we view other solutions, the majority of sub-optimal methods comprise moderately or based on the availability of the ideal CSI. In practice, propagation delay is generally unknown, intersymbol interference (ISI) or interchannel interference (ICI) may occur, and RF impairments may exist [30]–[32]. Therefore, the existing methods should be evolved or new approaches should be improved considering the facts of the actual systems.

### B. CHANNEL ESTIMATION ERRORS

Accurate channel estimation is critical to realize signal processing steps in PHY. The channel can be identified by correlation, maximum-likelihood, maximum a posteriori, or least-squares based estimators [33]–[36]. These methods hinge on several assumptions, as listed in Table 1, that are not always feasible from practical aspects. The acquisition of ideal CSI is not always available contrary to the common conjecture. For example, conventional systems assume

that ADCs have infinite resolution, and there is no quantization error. The real systems have to overcome estimation errors that cause round-off errors, unavailability of instantaneous feedback information, and interpolation errors. On the other hand, the impact of estimation errors on performance will be more critical due to massive antennas, 3D deployments, and high mobility scenarios because the channel conditions in future networks will change rapidly, and channel estimation errors must be eliminated for a superior performance [37]–[39]. Therefore, real-world impairments must be considered, and the solutions should be evaluated in terms of versatility.

### C. ERRONEOUS FEEDBACK INFORMATION

Feedback information is the primary requirement to coordinate the transmitter and receiver. It affects the overall performance. However, it is mostly erroneous by virtue of over-optimistic assumptions, as summarized in Table 1. The feedback information obtained via the forward link may not ideally represent the reverse link since they are not available simultaneously. Another issue is the outdated feedback caused by feedback delays [40]–[43]. Conversely, it must include all variations to ideally adjust the processes, such as antenna selection or beamforming. Consequently, the erroneous and delayed feedback information leads to a decrease in overall system performance [44], [45]. To feed the entire system with more accurate information, we have to develop a more intelligent and cautious solutions.

### D. RF FRONT-END IMPAIRMENTS

RF front-end impairments are apparent differences between real-world systems and simulated systems, and they are inevitable. Cost constraints and the fact that RF incompatibilities are hardware-based make them difficult to handle. Here, we addressed three main issues with respect to the assumptions in Table 1 that significantly affect overall system performance.

#### 1) IN-PHASE/QUADRATURE-PHASE (IQ) IMBALANCE

IQ imbalance is the mismatch of the amplitude and/or phase between in-phase (I) and quadrature-phase (Q) components. It stems from low-cost devices and causes a performance degradation; therefore, it must be considered in real systems [17], [46], [47]. IQ imbalance restrains the systems by sensitizing towards other impairments such as CFO. Estimation algorithms to compensate IQ imbalance based on least mean squares, maximum-likelihood, expectation-maximization, or iterations exist, but they introduce extra computation.

#### 2) PHASE NOISE

The phase noise in the oscillator results from the active circuit elements and makes frequency adjustments difficult. These fluctuations result in common phase error (CPE) and ICI. The real systems have to include a suppression process for ICI and an estimation algorithm for CPE mitigation [48].

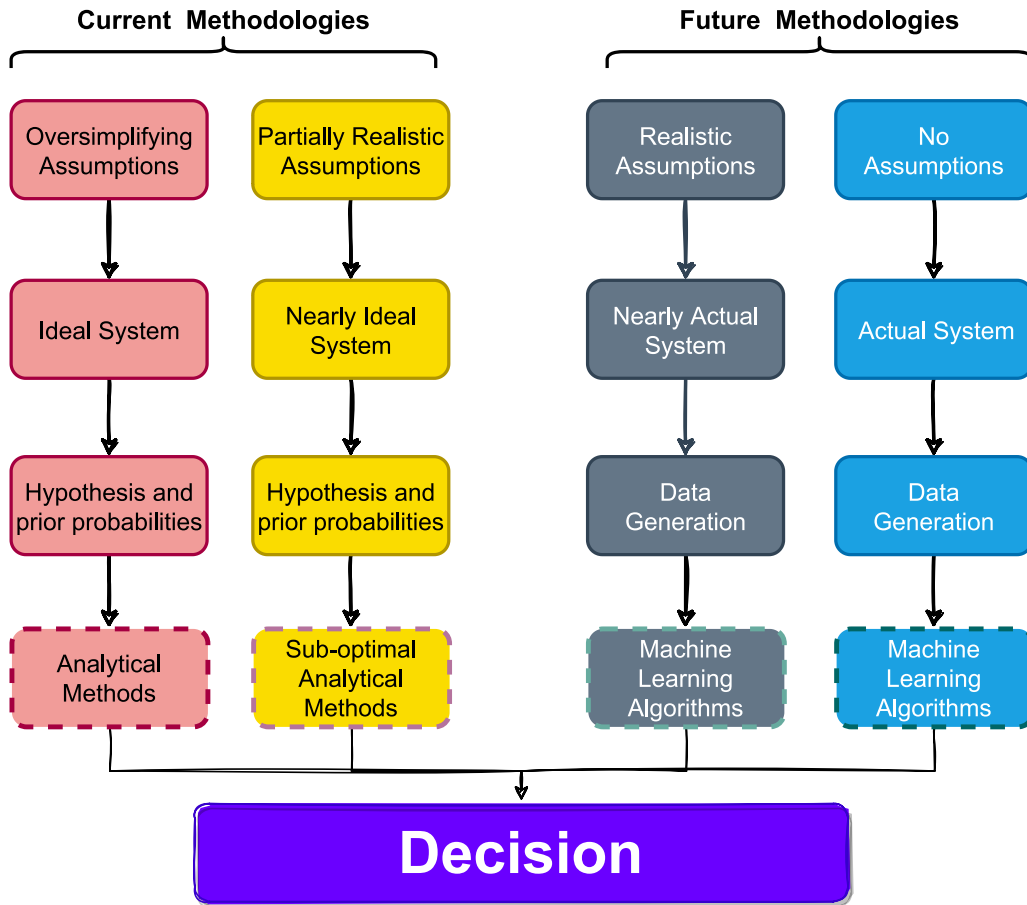


FIGURE 2. A Comparative Overview of Current and Future Methodologies for PHY Decision Mechanisms

### 3) NON-LINEARITIES

Non-linearities are realized by ADC/DAC, mixers, and amplifiers (the power amplifier in the transmitter and the low-noise amplifier in the receiver). The main part of the non-linearities in PHY is caused by power amplifiers. They are mostly ignored due to the difficulties of theoretically modeling.

### E. CORRELATION

Formulation of correlation is not straightforward due to its uncertainty. Making assumptions may cause a fallacious representation of the correlation properties, leading to an inaccurate model. Furthermore, describing correlation can require continuous pattern tracing in time, frequency, or spatial domains, which introduces additional processes and increases complexity. For example, auto-correlation of information sequence in time and frequency domains or cross-correlations between different channels introduce the requirement of new computational blocks in the transmitter and receiver.

### III. EVALUATIONS OF PHY DECISION MECHANISMS

The key point to obtain the optimal decision in PHY is structuring the problem in a linear model by assuming the

system is minimum variated and unbiased [49]. Current methodologies – as represented in Figure 2 – propose analytical methods by following this principle. The reason is that simplifying assumptions provide a reasonable and effortless means to examine the systems that are challenging to model. The majority of the solutions introduce oversimplifying assumptions – as listed in Table 1 – and idealize the system. Although some studies construct their solutions with partially realistic assumptions, they must make concessions in optimality or increase the complexity of the solution. After assuming that the system is ideal or nearly ideal, they define the problem with simplified possible conditions to constitute analytical methods. Consequently, the solution modeled by the current methodologies gradually deviates from the actual system step-by-step.

### A. CRITICAL OBSERVATIONS

Future methodologies must observe communication systems as in actual systems to make PHY decision mechanisms superior. The critical observations to consider when improving new PHY solutions are listed below:

- Actual systems do not operate ideally and struggle with synchronization problems, errors in feedback link,



channel estimation errors, correlation, and RF front-end impairments.

- Wireless communication environments and numerous parameters of a communication system change dynamically and may show unexpected variations.
- Analytical methods, such as the maximum a posteriori and maximum-likelihood based approaches, are employed by establishing hypotheses and assuming prior probabilities. However, these hypotheses and the corresponding probability distributions may not accurately represent realistic systems. For example, time variations or channel correlations can not be tracked accurately, and the corresponding assumptions become invalid. Furthermore, the inconsistencies in practice make it difficult to determine appropriate decision thresholds for the likelihood-based hypothesis tests.
- Towards 6G, each communication system will become unique by interlacing technologies, applications, and users. Until today, these methodologies have not affected the solutions' performance apparently. However, next-generation systems will face various exceptional conditions and suffer significantly from flawed system representations. Therefore, generalized solutions with over-simplifications will not meet the targeted performance requirements.

All these facts sidetrack or limit the presented solutions for PHY decision mechanisms. These limitations, arising from the assumptions, can be overcome with a changeover from the current methodologies to learning-driven methodologies.

### B. BENEFITS OF LEARNING-DRIVEN METHODOLOGIES

Learning-driven methodologies have a promising future for PHY decision mechanisms as well as in various parts of communication systems due to their attractive benefits. The inclusion of learning-driven methodologies in future solutions is motivated by the following benefits:

- The aforementioned problems due to confined system representations can be eliminated by providing data based on actual systems.
- A unique PHY decision mechanism can be built thanks to heuristic learning ability. If an inclusive dataset is available, learning algorithms can learn and decide based on the inherent characteristics of the system.
- The decision mechanism based on learning-driven methodologies can be capable of self-evolving and continuously improved. These capabilities make it more flexible and compatible with actual conditions.
- Autonomy can be increased by individual PHY decisions without human interference.

These motivations corroborate the idea that learning-driven methodologies get beyond the limits of the current methodologies. The next sections point out the main milestones and challenges to get the inference about learning-driven decision mechanisms.

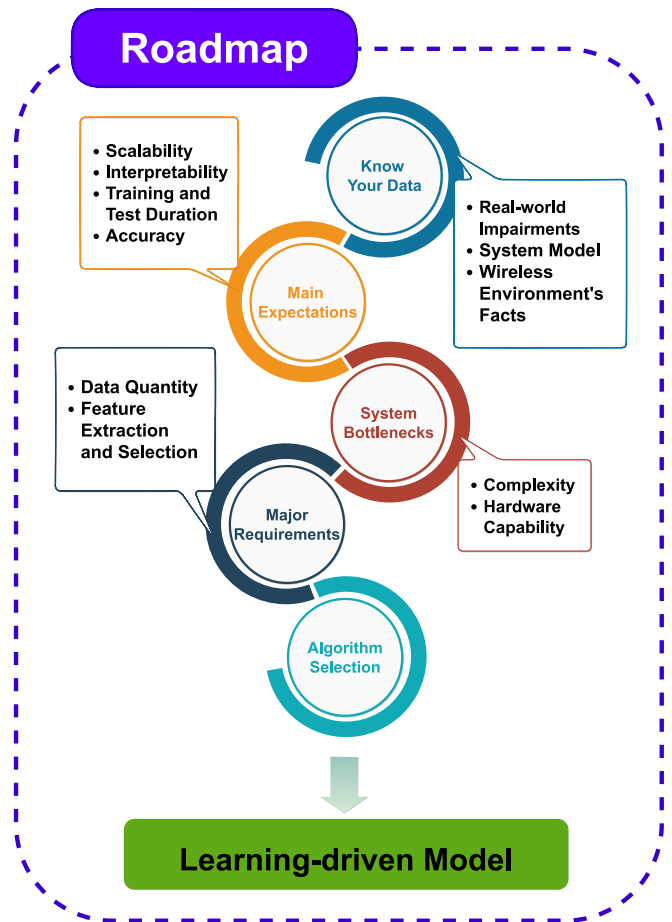


FIGURE 3. An illustration describing the introduction of the learning-driven solutions into PHY decision mechanisms in wireless communication systems.

### IV. A ROADMAP FOR LEARNING-DRIVEN SOLUTIONS

Learning algorithms from shallow to deep architectures have been advancing to comprehend the system facts. However, it is unclear how an algorithm should be sifted out from several algorithms for a solution and which steps should be considered primarily. The following roadmap, as visualized in Fig. 3, can be used to develop solutions for next-generation systems.

#### A. EXAMINE THE DATA SOURCE AND KNOW YOUR DATA

Learning algorithms do not magically provide answers to any system; they must be fed with suitable data to acquire the desired output [50]. It can be accomplished by understanding the problem and its source. Firstly, the system and environment that render the input data to output data must be analyzed correctly. Secondly, the input/output data variations should also be carefully observed, and their relations should be considered. These steps show what is expected to learn from machine learning. For example, in a communication system, let expect machine learning to gain insights about transmitted signals via received signals. Here, the input data may typify the received signal, whereas the output data may

refer to the transmitted signal. If only data based on transmitted and received signals are provided, the performance may not be sufficiently high. The reason for that the output data is formed with many factors during the transmission due to real-life impairments, and channel conditions. Then, the input data should include more information to obtain the correct output. If the knowledge about the data source exists, the input data can be enriched with more information. Additionally, it helps to find out reasons for some results during the machine learning model development stage.

### **B. DETERMINE THE MAIN EXPECTATIONS OF THE PROPOSED LEARNING-DRIVEN MODEL**

This step has a direct impact on learning-driven solution design. Performance targets such as accuracy, interpretability, scalability, training and prediction duration are determined depending on the proposed system [51]–[53]. Some learning models, mostly based on deep learning, are not obvious in terms of model transparency and functionality. However, interpretability and low-latency of the solution are indispensable, especially for mission-critical systems. Such learning models may not be suitable in PHY, despite their accuracy. Besides the interpretability, the scalability of the models is highly crucial because wireless communication devices and their capabilities diversify widely [54], [55].

### **C. BE AWARE OF THE SYSTEM BOTTLENECKS**

Employment of machine learning contains data processing, data storage, training and testing stages, and execution of generated machine learning models. All steps bring computational complexities along and require appropriate hardware capabilities. The systems may not meet high-level computational needs and constraint the solutions [56]–[59]. Therefore, the detection of system bottlenecks is critical for the selection of the proper machine learning algorithm. The number of conceivable options for learning algorithms may decrease together with main expectations from the learning model. If the remaining algorithms do not satisfy the demand, some alternative ways should be considered to work around the system constraints. Complexity can be reduced, for example, by converting non-numeric features into numeric features or by increasing the flexibility of the system architecture.

### **D. DETECT THE MAJOR REQUIREMENTS OF THE SOLUTION APPROACH**

This milestone refers to the data quantity and quality, or major requirements for qualified models. The data quantity and quality requirements may vary depending on the results of the previous milestones' outcomes. The amount of data for the remaining algorithms through the roadmap is the key part. If a sufficient amount of data is available, model quality is dictated by data quality matters such as completeness, consistency, veracity, validity, and timeliness of the data. However, not all essential data may be readily available, especially for wireless communications, which have various

destructive effects on data. Even if it is likely that the data will be collected in a timely manner, real-life impairments lead to inconsistencies, and the data will need to be processed to make it valid. Therefore, these constraints must be considered before choosing an algorithm [60]–[62].

### **E. CHOOSE AND OPTIMIZE THE SELECTED ALGORITHM**

Following these instructions, the most suitable learning algorithm is chosen. It is beneficial to keep in mind the considerations of the previous steps in order to prioritize and enable the iterative process of improving the learning model. For example, if the model is overfitted, the amount of data can be increased or the complexity of the model can be reduced. If these are not possible, regularization techniques or early stopping can be considered. Afterwards, the performance of the learning model must be optimized by tuning the hyper-parameters. Here, depending on feasibility, both informed search methods and uninformed search methods such as grid search or random search can be preferred [63]–[65]. After training with the chosen hyper-parameter settings, the acquired learning model can be employed in the proposed system.

## **V. CHALLENGES FOR LEARNING-DRIVEN SOLUTIONS**

### **A. DATA SCARCITY**

Data scarcity is the lack of data to generate a machine learning model. It poses a major obstacle to learning concepts based on the opinion, acquiring knowledge and insights through experiences. The main issue due to data scarcity is the overfitting of the model that leads to failure in real-time. Detecting outliers or noise in a sparse dataset is another issue. Basic solutions to avoid these problems are to favor simpler or linear techniques, and feature engineering. There are also further techniques for different data scarcity problems that can be divided into five categories: no data, rare data, small data, unlabeled data, and imbalanced data [66]–[74]. Here, no data, the absence of any data is the toughest but a possible condition due to privacy, security, and confidentiality concerns. However, if the problem is clearly defined and structured, several options can serve the purpose such as open-source datasets, encrypting or anonymizing data, and federated or online learning. As for the other scarcity problems, all techniques focus mainly on the following goals: increasing the amount of data or reducing the need for big data.

#### **1) INCREASING THE AMOUNT OF DATA**

Increasing the amount of data can be considered for small data, imbalanced data, and unlabeled data. A well-known solution to enlarge datasets is data augmentation, which is the acquisition of new data by diversifying existing data such as cropping, random insertion, modifying, and transforming. Synthetic data generation can be considered as another solution. Generative adversarial networks, simulation environments, and the synthetic minority over-sampling method

help to form realistic data. Additionally, self-supervised learning can overcome entirely the unlabeled data challenge. When there are some labeled data but the amount is not enough, semi-supervised learning and weak-supervision should be considered.

## 2) REDUCING THE NEED FOR BIG DATA

Reducing the need for big data can be a shortcut for small and rare data problems. Transfer learning, federated learning, and few-shot learning are prominent solutions. Transfer learning enables carrying insights based on a large dataset. On the other hand, federated learning aggregates multiple knowledge from decentralized machine learning models. The few-shot learning should be given much thought of rare datasets because it enables a generalization of less data information and prior knowledge.

### B. DATA ACQUISITION

Data acquisition has a significant importance to train learning algorithms successfully. Each data sample includes the information of a certain time interval and sometimes inaccurately represents the system due to the instantaneous changes in the environment. This issue can be handled by increasing the quantity of data. However, this is not sufficient if the collected data is not qualified. Generating a qualified dataset requires overcoming the following issues:

- The justifiability of the data-generation environment's convenience should be considered. For example, the preference of simulation-based datasets in a communication system can not be satisfactory to prove the learning-driven solution's performance due to the differences between simulations and the real-world. Real datasets can be provided at least as an explicit reference to confirm the consistency between the datasets from different sources.
- The generation of the accurate features included in the dataset is a crucial issue. Features should consist of the information individually or be jointly related to the problem and its solution. For example, features can be created depending on CSI to express estimation errors and correlation problems in PHY due to the impacts of channel conditions, as in [10].
- The correctness of the dataset can not be verified easily due to the lack of standardization for the labeling procedure.
- The trade-off between the size of data and the model performance should be addressed. Each additional feature expands the dataset dimensionally and leads to an increase in the training and prediction complexity. However, feature extraction or selection methods can be employed to overcome the problem, as described below.

### C. FEATURE EXTRACTION OR SELECTION

The extraction and selection of features are the main concepts to leverage the performance of learning algorithms,

and both of them improve data representation. The difference between them is the fact that feature extraction is more general and proposes to create useful features by utilizing the existing data while the feature selection targets increasing the relevancy by removing redundant features [50]. The selection of features provides a clarified representation of the data in lower-dimensional space by filtering. Therefore it is a highly effective approach, especially for problems that have a computationally intensive dataset. For instance, the analysis of the correlation and mutual information between the features, the determination of rating benchmark, and building a score function by weighting features are applied to detect the optimal feature subsets. The derivation/extraction of new features can be realized with linear and non-linear transformation methods such as principal component analysis, linear discriminant analysis, and autoencoder.

### D. COMPUTATIONAL COMPLEXITY

A learning process consists of two main stages: model training and prediction. Both bring a computational complexity to the system. The complexities of these processes vary due to several factors such as the selected learning algorithm and the parameter values for the model architecture's tuning and design. In the recent years, deep learning techniques have shown tremendous successes in broad application areas, and the trade-off between choosing less complex algorithms and attaining a higher accuracy is observed. The training complexity is overcome by the transfer learning, and the prediction complexity can be reduced with computational offloading or collaborative techniques [75]–[77].

### E. HARDWARE CAPABILITIES

Learning algorithms require hardware competence for data pre-processing, model training and testing. The utilization of advanced methods such as deep or ensemble learning techniques requires large quantities of the following resources: processing power, physical size, cost, and memory. The hardware should include sufficient memory to store variables, dataset, and the trained model besides computational power. At this point, the implementation of learning algorithms on computationally constrained devices becomes a complicated problem, especially in PHY. For example, many edge Internet of Things (IoT) devices in the industry are insufficient to realize the training process or store the dataset. Cloud platforms are a possible solution to store data, build models, or control the devices remotely. However, cloud-based solutions entail sturdy communication between the user and the cloud. This introduces another load in communication systems.

### F. SECURITY

The significant challenge is meeting the security necessities of data and the learning model to centralize the learning algorithm as a solution. Whilst robustness of the model is provided against the model's replicas, defense mechanisms should be investigated for data preservation and user privacy.





FIGURE 4. 4 × 1 test-bed which performs antenna selection at the transmitter.

Poisoning or evasion attacks threaten the data integrity and damage the tenacity of the model besides continuous changes in the channel and attack types. At this point, security maintenance and management for learning-driven solutions become more critical issues [78]–[80]. However, the capabilities of learning algorithms offer advantages to enhance the security against unpredictable dynamics and various attacks, including modification, denial-of-service, malware, and message replay.

## VI. SOFTWARE-DEFINED RADIO BASED CASE STUDY

We aim to show the capabilities of learning-driven approaches on real systems without theoretical approaches' common assumptions through a benchmark. In this section, the following considerations discussed in previous sections are highlighted with the proposed case study;

- A real system must overcome the front-end impairments of RF transceivers to sustain the efficacy such as IQ imbalance distortions or phase noise [81], (*issues about synchronization, RF front-end impairments*).
- The increase in the number of antennas brings new difficulties [82]. For example, obtaining perfect CSI is not possible at the transmitter, (*issues about channel estimation errors, correlation, and RF front-end impairments*).
- The multiple antennas require an larger number of RF chains, ADC/DAC and larger volatile memory.

However, the rise in the number of converters leads to the emergence of enormous data stacks in IQ planes, and the management of these data is another challenge. They can be minimized by reducing the bit-resolution of the conversion, but then quantization errors must be taken into account. The trade-off between the converter resolution and quantization errors is the another challenge, (*issues about hardware constraints, and RF front-end impairments*).

- Machine learning algorithms have different accuracy performances and complexities. Even though, we expect high performance which may require a complex architecture, some edge devices may not have suitable hardware, (*issues about hardware constraints, computational complexity, and accuracy performance*).
- Transmission conditions in a wireless environment may show variations due to interference, mobility, or noise, (*issues about data quantity, and feature extraction or selection*).
- The impact of the data amount on performance is important, especially some learning algorithms are more vulnerable. However, there may not be enough data available for the training process, (*issue about data scarcity and accuracy performance*).
- Even if data scarcity is not a problem, the time-complexity of the training process and the required storage space increase with larger data sets, (*issues about hardware constraints and computational complexity*).

We proposed a learning-driven solution for antenna selection decision that is performed at the transmitter with the goal of reducing the bit error rate (BER). Antenna selection decision performance is separately investigated through 200 different cases (2 ADC/DAC resolutions, 4 SNR values, 5 training data amount, and 5 algorithms) with a test-bed design by considering aforementioned issues. Measurements are taken from a 4×1 multiple-input-single-output test-bed, which is constructed using SDR units. The measurement-based performance results allow us to observe the composite impact of these challenges in the BER results, as seen in Figure 5.

### A. TEST-BED DESIGN

The dataset is prepared via the test-bed demonstrated in Fig. 4. It is designed with five SDR units, USRP-2943Rs at the transmitter and receiver. Each one is used with two RF chains. To provide hardware synchronization: reference clock is generated and shared via all USRP-2943Rs by using CDA-2990 8 channel clock distribution accessory. The transmitter's operating frequency and bandwidth are tuned as 2.45 GHz and 1 MHz. As a single-carrier modulation method, BPSK is used, with a root-raised cosine filter of roll-off factor 0.5. The distance between the transmitter and the receiver is set to 1.5 meters. 32 symbols are used for the acquisition, and the data/pilot rate is selected as 5/1 at an IQ rate of 125 ksample/s. CSI feedback is obtained via time-division duplex feedback.

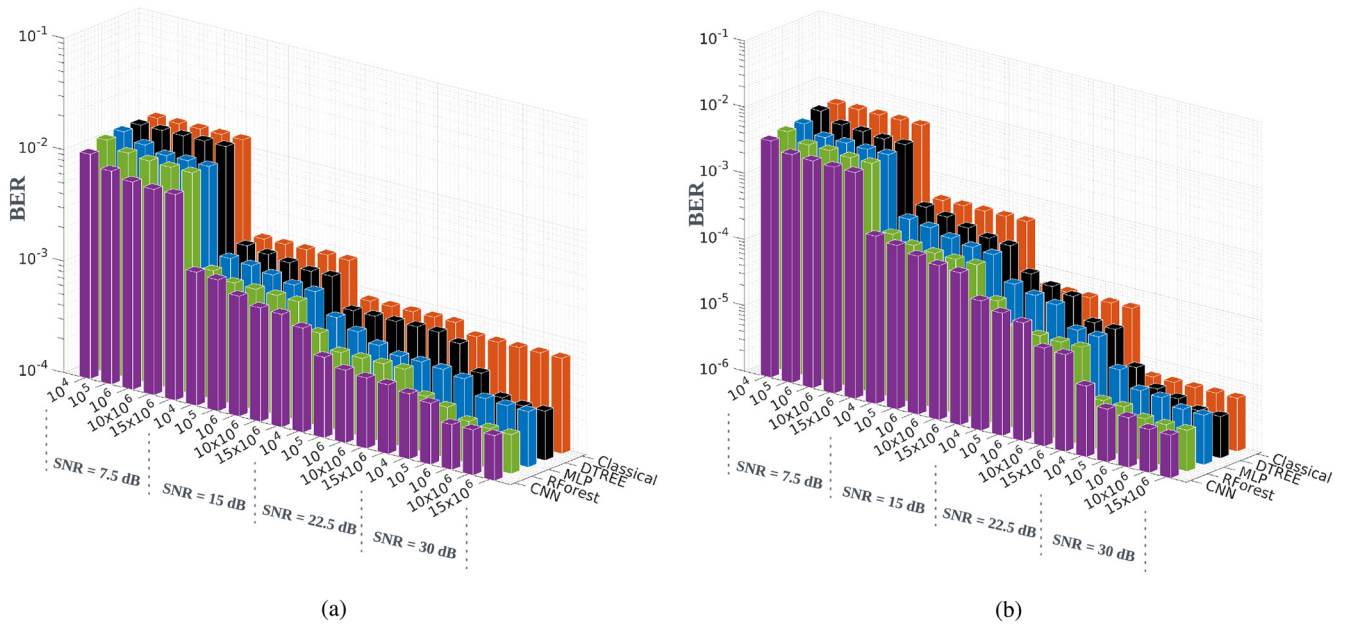


FIGURE 5. The comparative results of learning algorithms and maximum-likelihood (classical) based detector. a), and b) The results for 8-bit and 16-bit ADC/DAC, respectively.

**B. MACHINE LEARNING**

The proposed issue is defined as a classification problem. Four machine learning and deep learning algorithms are chosen: DTREE, MLP, RForest, CNN. The selection of algorithms is realized by considering the system constraints and expectations from learning models. The RForest’s depth, DTREE’s depth, MLP’s hidden layer number, and neurons at one layer are set as 25, 15, 2, 10, respectively. As a deep learning algorithm, 1D-CNN is chosen to present a fair comparison. The CNN architecture is designed with a one-dimensional convolution layer (256 filters) and two fully-connected layers (128 and 6 neurons, respectively). The batch size of CNN and MLP is selected as 1024. Training of the CNN is performed by utilizing the categorical-cross entropy loss function and the adaptive moment estimation function. The learning rate 0.001 and the exponential decay rates for the first and second moment estimates 0.99 and 0.999 are set.

**C. RESULTS**

The performances of RForest, DTREE, CNN, and MLP are compared with the conventional detector based on maximum-likelihood in terms of the BER measurements. Least-square-based channel estimation is performed in maximum-likelihood along with Moose’s algorithm adopted to a single carrier as described in [83]. A comparative illustration of their capabilities against the aforementioned practical challenges is shown in Fig. 5. The algorithms are trained with an offline learning mechanism by performing the same analysis for the different training data sizes. When the data size becomes larger, the learning algorithms achieve a higher performance than the conventional method because

of two reasons. The first reason is that numerous data samples carry the pervasive pattern of the system and represent conceivably. The second reason is the ability of heuristic learning. Moreover, these results prove the importance of data quantity and the performance flexibility of learning algorithms. Even though the necessary increase in data amount is not remarkable to reach a higher performance in this study, note that the required data amount depending on the problem can change to provide a desired leap of performance.

We can see that CNN and RForest outperform the other, less complex methods, while the conventional method generally shows a lower performance for each SNR and converter resolution. Although CNN performs better than RForest at lower SNR values for the 8-bit ADC/DAC resolution test-bed, RForest improves performance for the 16-bit ADC/DAC resolution test-bed. MLP and DTREE offer mediocre performance compared to their learning-driven and conventional counterparts. However, if the computational limitations are considered, their performances are substantially preferable. Furthermore, DTREE, after rising steeply, surpasses CNN and MLP and approaches RForest in the case of the 16-bit converter at a high SNR. Even if it offers a simpler structure, it does not provide the capability to capture non-linear relationships between data. Therefore, its performance is closer to maximum-likelihood detector than neural network based algorithms.

The decrease in the converters’ resolution leads to an explicit increment in BER, as shown in Figs. 5(a) and 5(b). BER differences increase, especially in high SNR. Although 16-bit ADC/DAC utilization is reasonable to sustain the reliability, this means two times IQ data, memory space, and complexity according to the number of antennas. The

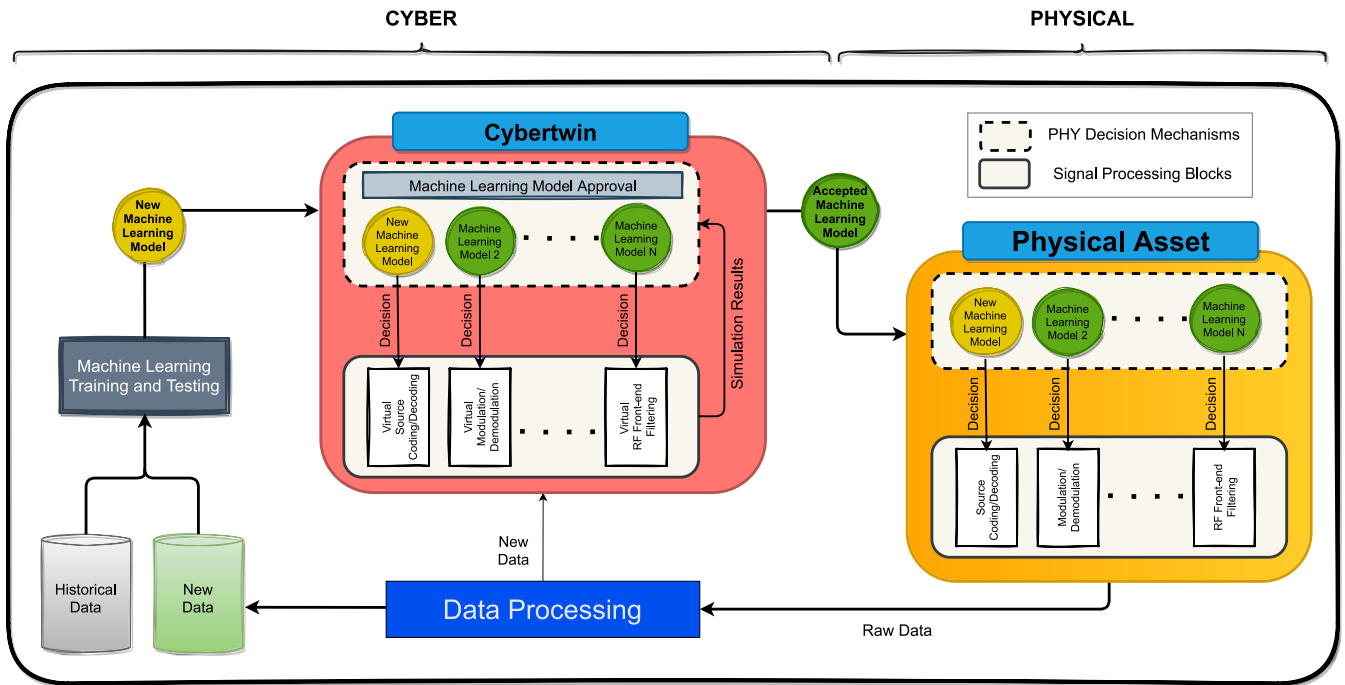


FIGURE 6. An intelligent cyber-physical framework for future remedies of next-generation communication systems.

enormous data stacks in the IQ plane can be reduced by including feature extraction. The results show that feature extraction is a favorable process to defeat the hardware constraints in PHY and take advantage of the reliability besides boosting the learning algorithms' performance.

## VII. FUTURE REMEDIES IN AN EMBRACING FRAMEWORK

The presented SDR-based case study substantiates the influences of the oversimplifying assumptions in Section II and the challenges in Section V. Machine learning can boost the system by eliminating oversimplifying assumptions and including reasoning. However, it does not become the sole remedy to meet the requirements of next-generation systems that are the compound of various cyber and physical processes. Considering these, several remedial solutions and concepts expand into the literature besides machine learning such as integrated sensing and communications [84]–[86], IoT, distributed ledgers [87], [88], edge-fog-cloud computing [89]–[91], and signal processing techniques [92]–[98]. In this section, we aim to provide an embracing framework for future remedies at cyber and physical levels.

The aforementioned and much more remedies have led up to an innovative view: cyber-physical systems (CPS) that are found on five fundamental aspects; sensing, communication, computing, control (monitoring and reasoning), and actuation. In the literature, CPS is mostly associated with IoT-based solutions for smart cities, Industry 4.0, or vehicular technologies [99]–[106]. Beyond these solutions, communication systems at all scales and levels need to

be formed into a cyber-physical framework to take advantage of upcoming remedies. They should also be equipped with learning-driven decision mechanisms and semantically upgraded. Therefore, we propose an intelligent and embracing cyber-physical framework for communication systems as demonstrated in Fig. 6. The proposed framework consists of three main operational parts: physical asset, machine learning, and cybertwin.

### A. PHYSICAL ASSET

The physical asset represents a tangible “thing” of the real-world communication systems. From the PHY perspective, it includes consecutive decision steps related to various transmission parameters such as coding rate, modulation type, constellation size, number of active antennas, transmit power, CFO, TO, channel estimation errors, or correlations. All these parameters feed signal processing blocks and must be adaptable to the variable conditions of the actual systems. Therefore, decision mechanisms are formed as machine learning models to provide more rational decisions. Each signal processing block is configured depending on the decision of the associated machine learning model, which is computed and controlled using historical running data. Here, secure data migration from the physical asset still needs to be investigated for different scenarios of future systems where the physical asset will not only be subject to different constraints but also to diverse set of applications.

### B. MACHINE LEARNING

The deep development part of the proposed framework comprises data processing and storage, as well as machine



learning training and testing. Thanks to numerous advances, this part of the framework can be flexibly designed for future communication scenarios. The roadmap, as explained in Section III, sets the focus of these operations, such as running fast algorithms, making approximate or exact predictions, and tolerance to dynamism or time lags in the data stream. A lot of remedial solutions from training strategies to state-of-the-art learning methods can be executed depending on the focus [107]–[112]. For example, when the focus is on accuracy, data processing may be more sophisticated and deep learning algorithms may be preferred. If avoiding the time lags in the data stream is the criterion, machine learning computations may take place at the edge rather than in the fog or in the cloud. Differently, federated learning can be applied as a training strategy to strengthen the dynamism, or transfer learning for latency vulnerability. However, data compression and reconstruction techniques still need to be explored to facilitate machine learning part for different scenarios and massive data sets. Although machine learning techniques can be equipped with numerous advances in the design phase, they do not guarantee that the last model generated will be better than the previous one. Future research studies need to focus on the robustness of learning models against model instability [113]. In addition, there is a mandatory examination before the model and the physical asset interaction but accomplishing this with human control seems unfeasible. Automated monitoring and developmental control of the model and the physical asset are crucial.

### C. CYBERTWIN

Cybertwin is a bridge between the physical asset and machine learning, and a virtual representation of the physical asset in cyberspace. It provides automated monitoring, control, and improvement by enabling simulations through the data [114]–[116]. Cybertwin is continuously updated with a new machine learning model and real data. If a new machine learning model meets the acceptance criteria after the simulations, it is implemented on the physical asset. It also checks the association of multiple machine learning models that are responsible for different signal processing blocks. These capabilities improve the flexibility, fidelity, and reliability of machine learning-driven decision mechanisms and leverage the overall performance. However, a lightweight and secure communication protocol for interaction between the cybertwin and other operational parts should be improved to take advantage of cybertwin [117].

### VIII. CONCLUSION

This article presents a comprehensive overview of the wireless communication systems' prominent conjectures in PHY and how they may introduce modeling errors in real-world systems. From this aspect, learning algorithms are addressed with an elucidating guideline to defeat their classical counterparts and eliminate the need for oversimplifying and impractical assumptions. Additionally, the challenges to employ learning-driven solutions in PHY are presented

for a holistic view. The listed considerations are supported by a real-time SDR-based case study. The results prove that learning algorithms will take a role as a key technology for future avenues of wireless communication. At this stage, future studies should focus on how learning algorithms can be implemented with other technologies. In the last section, a holistic cyber-physical framework is proposed for joint implementation with future remedies. Another direction of future studies will be the challenges of learning-driven solutions, especially trustworthiness and interpretability [118].

### REFERENCES

- [1] L. Bariah *et al.*, "A prospective look: Key enabling technologies, applications and open research topics in 6G networks," *IEEE Access*, vol. 8, pp. 174792–174820, 2020.
- [2] Z. Zhang *et al.*, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, Sep. 2019.
- [3] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.
- [4] G. Jia, Z. Yang, H.-K. Lam, J. Shi, and M. Shikh-Bahaei, "Channel assignment in uplink wireless communication using machine learning approach," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 787–791, Apr. 2020.
- [5] H. Liu, S. Lu, M. El-Hajjar, and L.-L. Yang, "Machine learning assisted adaptive index modulation for mmWave communications," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1425–1441, 2020.
- [6] N. Shlezinger, N. Farsad, Y. C. Eldar, and A. J. Goldsmith, "ViterbiNet: A deep learning based Viterbi algorithm for symbol detection," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3319–3331, Sep. 2020.
- [7] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5G networks: Joint Beamforming, power control, and interference coordination," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1581–1592, Mar. 2020.
- [8] P. E. Gorday, N. Erdöl, and H. Zhuang, "Complex-valued neural networks for Noncoherent demodulation," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 217–225, 2020.
- [9] X. Chen, J. Cheng, Z. Zhang, L. Wu, J. Dang, and J. Wang, "Data-rate driven transmission strategies for deep learning-based communication systems," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2129–2142, Apr. 2020.
- [10] S. Gecgel, C. Goztepe, and G. K. Kurt, "Transmit antenna selection for large-scale MIMO GSM with machine learning," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 113–116, Jan. 2020.
- [11] S. Gecgel, C. Goztepe, and G. K. Kurt, "Jammer detection based on artificial neural networks: A measurement study," in *Proc. ACM Workshop Wireless Security Mach. Learn.*, 2019, pp. 43–48.
- [12] K. E. Kolodziej, A. U. Cookson, and B. T. Perry, "RF Cancellor tuning acceleration using neural network machine learning for in-band full-duplex systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1158–1170, 2021.
- [13] H. Huang *et al.*, "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 214–222, Mar. 2020.
- [14] F. Restuccia and T. Melodia, "Deep learning at the physical layer: System challenges and applications to 5G and beyond," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 58–64, Oct. 2020.
- [15] F. Ling and J. Proakis, *Synchronization in Digital Communication Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
- [16] S. K. Sharma, T. E. Bogale, S. Chatzinotas, B. Ottersten, L. B. Le, and X. Wang, "Cognitive radio techniques under practical imperfections: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1858–1884, 4th Quart., 2015.
- [17] G. Xing, M. Shen, and H. Liu, "Frequency offset and i/Q imbalance compensation for direct-conversion receivers," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 673–680, Mar. 2005.



- [18] H. Mercier, V. K. Bhargava, and V. Tarokh, "A survey of error-correcting codes for channels with symbol synchronization errors," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 1, pp. 87–96, 1st Quart., 2010.
- [19] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.
- [20] Z. Zhong, P. Chen, and T. He, "On-demand time synchronization with predictable accuracy," in *Proc. IEEE INFOCOM*, 2011, pp. 2480–2488.
- [21] A. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer–Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1391–1399, Feb. 1994.
- [22] J. J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.
- [23] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM. I," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668–1677, Nov. 1999.
- [24] M. Luise and R. Reggiannini, "Carrier frequency recovery in all-digital modems for burst-mode transmissions," *IEEE Trans. Commun.*, vol. 43, nos. 2–4, pp. 1169–1178, Jul. 1995.
- [25] W. G. Cowley and L. P. Sabel, "The performance of two symbol timing recovery algorithms for PSK demodulators," *IEEE Trans. Commun.*, vol. 42, no. 6, pp. 2345–2355, Jun. 1994.
- [26] B. Yang, K. Letaief, R. Cheng, and Z. Cao, "Timing recovery for OFDM transmission," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2278–2291, Nov. 2000.
- [27] J. Liu and J. Li, "Parameter estimation and error reduction for OFDM-based WLANs," *IEEE Trans. Mobile Comput.*, vol. 3, no. 2, pp. 152–163, Apr.–Jun. 2004.
- [28] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers, Volume 2: Synchronization, Channel Estimation, and Signal Processing* (Wiley Series in Telecommunications and Signal Processing). New York, NY, USA: Wiley, 1997.
- [29] U. Mengali and A. D'Andrea, *Synchronization Techniques for Digital Receivers* (Applications of Communications Theory). Boston, MA, USA: Springer, 1997.
- [30] A. Nasir, S. Durrani, R. Kennedy, H. Mehrpouyan, and S. Blostein, "Timing and carrier synchronization in wireless communication systems: A survey and classification of research in the last 5 years," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, pp. 1–38, 2016.
- [31] J. He, G. Gu, and Z. Wu, "MMSE interference suppression in MIMO frequency selective and time-varying fading channels," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3638–3651, Aug. 2008.
- [32] N. Beaulieu, "The evaluation of error probabilities for intersymbol and cochannel interference," *IEEE Trans. Commun.*, vol. 39, no. 12, pp. 1740–1749, Dec. 1991.
- [33] J. Proakis, *Digital Communications*, 5th ed. London, U.K.: McGraw Hill, 2007.
- [34] G. Stuber, J. Barry, S. McLaughlin, Y. Li, M. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.
- [35] Y.-S. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.
- [36] J.-J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *Proc. Veh. Technol. Conf. Countdown Wireless 21st Century*, vol. 2, 1995, pp. 815–819.
- [37] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.
- [38] Y. Liao, G. Sun, Z. Cai, X. Shen, and Z. Huang, "Nonlinear Kalman filter-based robust channel estimation for high mobility OFDM systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 7219–7231, Nov. 2021.
- [39] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A survey of channel modeling for UAV communications," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2804–2821, 4th Quart., 2018.
- [40] H. Zhang, S. Wei, G. Ananthaswamy, and D. L. Goeckel, "Adaptive signaling based on statistical Characterizations of outdated feedback in wireless communications," *Proc. IEEE*, vol. 95, no. 12, pp. 2337–2353, Dec. 2007.
- [41] Q. Ma and C. Tepedelenioglu, "Practical multiuser diversity with outdated channel feedback," *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1334–1345, Jul. 2005.
- [42] H. A. Suraweera, M. Soysa, C. Tellambura, and H. K. Garg, "Performance analysis of partial relay selection with feedback delay," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 531–534, Jun. 2010.
- [43] Y. Teng, M. Liu, and M. Song, "Effect of outdated CSI on handover decisions in dense networks," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2238–2241, Oct. 2017.
- [44] C. Jonietz, W. H. Gerstacker, and R. Schober, "Robust transmit processing for frequency-selective fading channels with imperfect channel feedback," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5356–5368, Dec. 2008.
- [45] Y. Shi, M. Badi, D. Rajan, and J. Camp, "Channel reciprocity analysis and feedback mechanism design for mobile beamforming systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6029–6043, Jun. 2021.
- [46] A. Tarighat, R. Bagheri, and A. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3257–3268, Aug. 2005.
- [47] N. Kolomvakis, M. Matthaiou, and M. Coldrey, "IQ imbalance in multiuser systems: Channel estimation and compensation," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 3039–3051, Jul. 2016.
- [48] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Heidelberg, Germany: Springer, 2008.
- [49] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1997.
- [50] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
- [51] S. Sinha, P. Franciosa, and D. Ceglarek, "Building a scalable and interpretable Bayesian deep learning framework for quality control of free form surfaces," *IEEE Access*, vol. 9, pp. 50188–50208, 2021.
- [52] K. Chen *et al.*, "Recent advances in data-driven wireless communication using Gaussian processes: A comprehensive survey," *China Commun.*, vol. 19, no. 1, pp. 218–237, 2022.
- [53] Z. Ye, K. Wang, Y. Chen, X. Jiang, and G. Song, "Multi-UAV navigation for partially observable communication coverage by graph reinforcement learning," *IEEE Trans. Mobile Comput.*, early access, Jan. 31, 2022, doi: [10.1109/TMC.2022.3146881](https://doi.org/10.1109/TMC.2022.3146881).
- [54] Y. Xu, F. Yin, W. Xu, C.-H. Lee, J. Lin, and S. Cui, "Scalable learning paradigms for data-driven wireless communication," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 81–87, Oct. 2020.
- [55] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1759–1799, 3rd Quart., 2021.
- [56] J. Borrego-Carazo, D. Castells-Rufas, E. Biempica, and J. Carrabina, "Resource-constrained machine learning for ADAS: A systematic review," *IEEE Access*, vol. 8, pp. 40573–40598, 2020.
- [57] H. Liu, W. Zipping, H. Zhang, B. Li, and C. Zhao, "Tiny machine learning (Tiny-ML) for efficient channel estimation and signal detection," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6795–6800, Jun. 2022.
- [58] B. Sudharsan *et al.*, "Toward distributed, global, deep learning using IoT devices," *IEEE Internet Comput.*, vol. 25, no. 3, pp. 6–12, May/Jun. 2021.
- [59] S. Gollapudi, *Practical Machine Learning*. London, U.K.: Packt, 2016.
- [60] D. Liu, G. Zhu, Q. Zeng, J. Zhang, and K. Huang, "Wireless data acquisition for edge learning: Data-importance aware retransmission," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 406–420, Jan. 2021.
- [61] Y. Liu, X. Guan, Y. Peng, H. Chen, T. Ohtsuki, and Z. Han, "Blockchain-based task offloading for edge computing on low-quality data via distributed learning in the Internet of Energy," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 2, pp. 657–676, Feb. 2022.
- [62] W. Yu, Y. Liu, T. Dillon, W. Rahayu, and F. Mostafa, "An integrated framework for health state monitoring in a smart factory employing IoT and big data techniques," *IEEE Internet Things J.*, vol. 9, no. 3, pp. 2443–2454, Feb. 2022.

- [63] G. Luo, "A review of automatic selection methods for machine learning algorithms and hyper-parameter values," *Netw. Model. Anal. Health Inform. Bioinf.*, vol. 5, no. 1, pp. 1–16, 2016.
- [64] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 2, p. 112, 2012.
- [65] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 24, 2011, p. 9.
- [66] V. S. Sheng, F. Provost, and P. G. Ipeirotis, "Get another label? Improving data quality and data mining using multiple, noisy Labelers," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Min. (KDD)*, 2008, pp. 614–622.
- [67] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.
- [68] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: A big data—AI integration perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 4, pp. 1328–1347, Apr. 2021.
- [69] K. Emam, L. Mosquera, and R. Hoptroff, *Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data*. London, U.K.: O'Reilly Media, 2020.
- [70] Z. Meng, X. Guo, Z. Pan, D. Sun, and S. Liu, "Data segmentation and augmentation methods based on raw data using deep neural networks approach for rotating machinery fault diagnosis," *IEEE Access*, vol. 7, pp. 79510–79522, 2019.
- [71] N.-T. Tran, V.-H. Tran, N.-B. Nguyen, T.-K. Nguyen, and N.-M. Cheung, "On data augmentation for GAN training," *IEEE Trans. Image Process.*, vol. 30, pp. 1882–1897, 2021.
- [72] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Int. Res.*, vol. 16, no. 1, pp. 321–357, Jun. 2002.
- [73] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of classification methods on unbalanced data sets," *IEEE Access*, vol. 9, pp. 64606–64628, 2021.
- [74] C. Geng, S.-J. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3614–3631, Oct. 2021.
- [75] M. Khayyat, I. A. Elgendy, A. Muthanna, A. S. Alshahrani, S. Alharbi, and A. Koucheryavy, "Advanced deep learning-based computational offloading for multilevel vehicular edge-cloud computing networks," *IEEE Access*, vol. 8, pp. 137052–137062, 2020.
- [76] B. Maschler and M. Weyrich, "Deep transfer learning for industrial automation: A review and discussion of new techniques for data-driven machine learning," *IEEE Ind. Electron. Mag.*, vol. 15, no. 2, pp. 65–75, Jun. 2021.
- [77] H. Xiao, C. Xu, Y. Ma, S. Yang, L. Zhong, and G.-M. Muntean, "Edge intelligence: A computational task offloading scheme for dependent IoT application," *IEEE Trans. Wireless Commun.*, early access, Mar. 11, 2022, doi: [10.1109/TWC.2022.3156905](https://doi.org/10.1109/TWC.2022.3156905).
- [78] M. Xue, C. Yuan, H. Wu, Y. Zhang, and W. Liu, "Machine learning security: Threats, countermeasures, and evaluations," *IEEE Access*, vol. 8, pp. 74720–74742, 2020.
- [79] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, and V. C. M. Leung, "A survey on security threats and defensive techniques of machine learning: A data driven view," *IEEE Access*, vol. 6, pp. 12103–12117, 2018.
- [80] F. O. Olowononi, D. B. Rawat, and C. Liu, "Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for CPS," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 524–552, 2nd Quart., 2021.
- [81] V. K. V. Gottumukkala and H. Minn, "Capacity analysis and pilot-data power allocation for MIMO-OFDM with transmitter and receiver IQ imbalances and residual carrier frequency offset," *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 553–565, Feb. 2012.
- [82] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [83] R. Heath, *Introduction to Wireless Digital Communication: A Signal Processing Perspective*. Upper Saddle River, NJ, USA: Prentice Hall, 2017.
- [84] J. Mu, Y. Gong, F. Zhang, Y. Cui, F. Zheng, and X. Jing, "Integrated sensing and communication-enabled predictive beamforming with deep learning in vehicular networks," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3301–3304, Oct. 2021.
- [85] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Netw.*, vol. 35, no. 5, pp. 158–167, Sep./Oct. 2021.
- [86] D. K. P. Tan *et al.*, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," in *Proc. IEEE Int. Online Symp. Joint Commun. Sens.*, 2021, pp. 1–6.
- [87] L. D. Nguyen, A. E. Kalor, I. Leyva-Mayorga, and P. Popovski, "Trusted wireless monitoring based on distributed ledgers over NB-IoT connectivity," *IEEE Commun. Mag.*, vol. 58, no. 6, pp. 77–83, Jun. 2020.
- [88] P. Danzi *et al.*, "Communication aspects of the integration of wireless IoT devices with distributed ledger technology," *IEEE Netw.*, vol. 34, no. 1, pp. 47–53, Jan./Feb. 2020.
- [89] D. Kimovski, R. Mathá, J. Hammer, N. Mehran, H. Hellwagner, and R. Prodan, "Cloud, fog, or edge: Where to compute?" *IEEE Internet Comput.*, vol. 25, no. 4, pp. 30–36, Jul./Aug. 2021.
- [90] Y. Zhou, L. Tian, L. Liu, and Y. Qi, "Fog computing enabled future mobile communication networks: A convergence of communication and computing," *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 20–27, May 2019.
- [91] Q.-V. Pham *et al.*, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.
- [92] A. I. Perez-Neira, M. A. Vazquez, M. B. Shankar, S. Maleki, and S. Chatzinotas, "Signal processing for high-throughput satellites: Challenges in new interference-limited scenarios," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 112–131, Jul. 2019.
- [93] E. Bjornson and P. Giselsson, "Two applications of deep learning in the physical layer of communication systems [lecture notes]," *IEEE Signal Process. Mag.*, vol. 37, no. 5, pp. 134–140, Sep. 2020.
- [94] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, May 2013.
- [95] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM? An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 80–93, Nov. 2014.
- [96] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [97] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Mar. 2008.
- [98] G. Altun and I. A. Arslan, "Joint transmit and receive antenna selection for spatial modulation systems using deep learning," *IEEE Commun. Lett.*, early access, Jun. 10, 2022, doi: [10.1109/LCOMM.2022.3182213](https://doi.org/10.1109/LCOMM.2022.3182213).
- [99] A. K. Tripathy, P. K. Tripathy, A. G. Mohapatra, N. K. Ray, and S. P. Mohanty, "WeDoShare: A Ridesharing framework in transportation Cyber-physical system for sustainable mobility in smart cities," *IEEE Consum. Electron. Mag.*, vol. 9, no. 4, pp. 41–48, Jul. 2020.
- [100] N. Jazdi, "Cyber physical systems in the context of industry 4.0," in *Proc. IEEE Int. Conf. Autom. Qual. Test. Robot.*, 2014, pp. 1–4.
- [101] Z. You and L. Feng, "Integration of industry 4.0 related technologies in construction industry: A framework of cyber-physical system," *IEEE Access*, vol. 8, pp. 122908–122922, 2020.
- [102] A. J. C. Trappey, C. V. Trappey, U. H. Govindarajan, J. J. Sun, and A. C. Chuang, "A review of technology standards and patent portfolios for enabling cyber-physical systems in advanced manufacturing," *IEEE Access*, vol. 4, pp. 7356–7382, 2016.
- [103] Y. Feng, B. Hu, H. Hao, Y. Gao, Z. Li, and J. Tan, "Design of distributed cyber-physical systems for connected and automated vehicles with implementing methodologies," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 4200–4211, Sep. 2018.
- [104] F. Amin and G. S. Choi, "Hotspots analysis using cyber-physical-social system for a smart city," *IEEE Access*, vol. 8, pp. 122197–122209, 2020.
- [105] R. Rodulfo, "Smart city case study: City of coral gables leverages the Internet of Things to improve quality of life," *IEEE Internet Things Mag.*, vol. 3, no. 2, pp. 74–81, Jun. 2020.

- [106] D. Jia, K. Lu, J. Wang, X. Zhang, and X. Shen, "A survey on platoon-based vehicular cyber-physical systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 263–284, 1st Quart., 2016.
- [107] W. Y. B. Lim *et al.*, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.
- [108] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proc. IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021.
- [109] A. Feriani and E. Hossain, "Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 1226–1252, 2nd Quart., 2021.
- [110] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 1st Quart., 2019.
- [111] A. Steiner, A. Kolesnikov, X. Zhai, R. Wightman, J. Uszkoreit, and L. Beyer, "How to train your ViT? Data, augmentation, and Regularization in vision transformers," 2021, [arxiv.abs/2106.10270](https://arxiv.org/abs/2106.10270).
- [112] S. Jastrzebski *et al.*, "Catastrophic Fisher explosion: Early phase fisher matrix impacts generalization," 2020, [arxiv.abs/2012.14193](https://arxiv.org/abs/2012.14193).
- [113] M. Zhang, H. Zhang, Y. Fang, and D. Yuan, "Learning-based data transmissions for future 6G enabled industrial IoT: A data compression perspective," *IEEE Netw.*, early access, Mar. 30, 2021, doi: [10.1109/ACCESS.2021.3069707](https://doi.org/10.1109/ACCESS.2021.3069707).
- [114] S. Shen, C. Yu, K. Zhang, and S. Ci, "Adaptive artificial intelligence for resource-constrained connected vehicles in Cybertwin-driven 6G network," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16269–16278, Sep. 2021.
- [115] Y. Lu, S. Maharjan, and Y. Zhang, "Adaptive edge association for wireless digital twin networks in 6G," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16219–16230, Nov. 2021.
- [116] T. K. Rodrigues, J. Liu, and N. Kato, "Application of cybertwin for offloading in mobile multiaccess edge computing for 6G networks," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16231–16242, Nov. 2021.
- [117] G. Li, C. Lai, R. Lu, and D. Zheng, "SecCDV: A security reference architecture for Cybertwin-driven 6G V2X," *IEEE Trans. Veh. Technol.*, vol. 71, no. 5, pp. 4535–4550, May 2022.
- [118] A. Nascita, A. Montieri, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescapé, "XAI meets mobile traffic classification: Understanding and improving multimodal deep learning architectures," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 4, pp. 4225–4246, Dec. 2021.



**SELEN GEGGEL CETIN** (Student Member, IEEE) received the B.S. degree in electronics and communication engineering from Yildiz Technical University, Istanbul, Turkey, in 2016, and the M.S. degree in telecommunication engineering from the Istanbul Technical University (ITU), Istanbul, in 2019, where she is currently pursuing the Ph.D. degree. From 2019 to 2021, she was a Research Assistant with the Department of Electronics Engineering, Turkish Air Force Academy, NDU. She is a member of the ITU

Wireless Communication Research Laboratory, and her current research is interested in machine learning and wireless communications.



**CANER GOZTEPE** received the B.Sc. and M.Sc. degrees in telecommunication engineering from Istanbul Technical University, Turkey, in 2017 and 2019, respectively. He currently manages research and development studies on behalf of different companies in the USA, Canada, and Turkey. His main research interests are artificial intelligence, gamification, IoT, 5G+, physical layer security, and SDR applications beyond 5G physical layer schemes.



**GUNES KARABULUT KURT** (Senior Member, IEEE) received the B.S. degree (High Hons.) in electronics and electrical engineering from Bogazici University, Istanbul, Turkey, in 2000, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Ottawa, ON, Canada, in 2002 and 2006, respectively. From 2000 to 2005, she was a Research Assistant with the CASP Group, University of Ottawa. She was with TenXc Wireless, Canada, from 2005 to 2006,

Edgewater Computer Systems Inc., Canada, from 2006 to 2008, Turkcell Research and Development Applied Research and Technology, Istanbul, from 2008 to 2010, and Istanbul Technical University from 2010 to 2021. She is currently an Associate Professor of Electrical Engineering with Polytechnique Montréal, Montreal, QC, Canada. She is an Adjunct Research Professor with Carleton University. Her current research interests include space information networks, satellite networking, wireless network coding, wireless security, space security, and wireless testbeds. She is a Marie Curie Fellow and has received the Turkish Academy of Sciences Outstanding Young Scientist (TÜBA-GEBIP) Award in 2019. She is also currently serving as an Associate Technical Editor for the *IEEE Communications Magazine* and a member for the IEEE WCNC Steering Board. She is the Chair of the IEEE special interest group titled *Satellite Mega-Constellations: Communications and Networking*.



**HALIM YANIKOMEROĞLU** (Fellow, IEEE) is a Professor with the Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada. From 2012 to 2016, he led one of the largest academic-industrial collaborative research programs on pre-standards 5G wireless. In Summer 2019, he started a new large-scale project on the 6G non-terrestrial networks. His extensive collaboration with industry resulted in 39 granted patents. He has formally supervised or hosted at Carleton a total of 135 postgraduate researchers in all levels (Ph.D. and M.A.Sc. students, PDFs, and Professors). He has coauthored IEEE papers with faculty members in 80+ universities in 25 countries and industry researchers in ten countries. His primary research domain is wireless communications and networks. His research group has made substantial contributions to 4G and 5G wireless technologies.

Prof. Yanikomeroğlu received several awards for his research, teaching, and service, including the IEEE Communications Society Wireless Communications Technical Committee Recognition Award in 2018, the IEEE Vehicular Technology Society Stuart Meyer Memorial Award in 2020, and the IEEE Communications Society Fred W. Ellersick Prize in 2021. He is currently serving as the Chair of the IEEE Wireless Communications and Networking Conference (WCNC) Steering Committee. He was the Technical Program Chair/Co-Chair of WCNC 2004 (Atlanta), WCNC 2008 (Las Vegas), and WCNC 2014 (Istanbul). He was the General Chair of IEEE VTC 2010-Fall (Ottawa) and VTC 2017-Fall (Toronto). He also served as the Chair of the IEEE's Technical Committee on Personal Communications. He is a Fellow of Engineering Institute of Canada and Canadian Academy of Engineering and a Distinguished Speaker for IEEE Communications Society and IEEE Vehicular Technology Society.