future internet

MDPI

Article

Cross Layer Optimization Using AI/ML-Assisted Federated Edge Learning in 6G Networks

Spyridon Louvros [1,*], AnupKumar Pandey [2], Brijesh Shah [2] and Yashesh Buch [2]

[1] Jio Platforms Limited (JPL), Miiduranna Village, 74015 Viimsi Parish, Estonia
[2] Research & Development, Jio Platforms Limited, TC-22, 5th Floor, A Wing, Reliance Corporate Park, Thane Belapur Road, Ghansoli, Navi Mumbai 400701, Maharashtra, India; anupkumar.pandey@ril.com (A.P.); brijesh.i.shah@ril.com (B.S.); yashesh.buch@ril.com (Y.B.)
* Correspondence: spyridon.louvros@jio.eu; Tel.: +30-6977277076

Abstract

This paper introduces a novel methodology that integrates 6G wireless Federated Edge Learning (FEEL) frameworks with MAC to PHY cross layer optimization strategies. In the context of mobile edge computing, typically ensuring robust channel estimation within the 6G network use cases presents critical challenges, particularly in managing data retransmissions. Inaccurate updates from distributed 6G devices can undermine the reliability of federated learning, affecting its overall performance. To address this, rather than relying on direct evaluations of the objective function, we propose an AI/ML-assisted algorithm for global optimization based on radial basis functions (RBFs) decision-making process to assess learned preference options.

Keywords: 6G; HARQ retransmissions; federated (collaborative) learning (FL); Federated Edge Learning (FEEL); mobile edge computing (MEC); AI/ML; optimization; radial basis functions (RBF)

Check for updates

Academic Editors: Jianwu Zhang and Léo Mendiboure

Received: 7 November 2025
Revised: 12 January 2026
Accepted: 14 January 2026
Published: 30 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Federated learning (FL), as defined by 3GPP [1] and IEEE [2,3], represents an innovative paradigm in machine learning that addresses the challenges of decentralized data environments [4]. The integration of artificial intelligence (AI) into the 3GPP framework has reached a significant milestone, with ongoing research and specifications aimed at enhancing data utilization and predictive capabilities expected to be featured in 3GPP Release 19 and Release 20. This development underscores the organization's commitment to leveraging advanced analytical techniques for improving telecommunications standards. Within this context, both the Technical Specification Group for Radio Access Networks (TSG RAN) and the Technical Specification Group for Service and System Aspects (TSG SA) have outlined specific requirements to incorporate AI alongside its companion, machine learning (ML).

1.1. Background and Motivation

Notably, Working Group RAN3 has finalized Technical Report 37.817, which investigates enhancements for data collection relating to New Radio (NR) and Evolved Universal Terrestrial Radio Access (ENDC). This report highlights three primary use cases where AI and ML can provide meaningful solutions. Starting with Network Energy Savings, it addresses strategies such as traffic offloading, modifications to coverage, and cell deactivation to reduce energy consumption. Moreover, a further focus on Load Balancing

supports the implementation techniques to optimize load distribution across cells or groups of cells in multi-frequency and multi-radio access technology (RAT) environments, thereby improving overall network performance based on predictive analytics. Finally, a focus on Mobility Optimization will practically apply a robust network performance during user equipment (UE) mobility events by selecting optimal mobility targets grounded in predictive assessments of service delivery. These focal points illustrated a strategic approach towards evolving network management through intelligent data-driven decision-making.

The FL architecture is supported by the NWDAF in the crucial 3GPP standard method [5]. It efficiently collects data from user equipment, network functions, Operation and Maintenance (OAM) systems within the 5G Core, Cloud, and Edge networks. This wealth of data is then utilized to trigger powerful 5G analytics, enabling better insights and actions to enhance the overall end-user experience. In any case the Federated Edge Learning (FEEL) shall comply with the existing NWDAF framework and 5GS framework as specified in [6–9]. Artificial intelligence (AI) and machine learning (ML) over NWDAF have become pivotal technologies driving advancements in wireless communication networks. These cutting-edge tools offer innovative solutions to enhance the efficiency, scalability, and performance of modern network infrastructures. Recognizing their potential, 3GPP has integrated AI/ML technologies into the Radio Access Network (RAN) as part of Release 18, marking the beginning of 5G-Advanced [10]. The inclusion of AI/ML in Release 18 underscores its importance in advancing the capabilities of 5G networks.

The Management Data Analytics Function (MDAF) serves as a fundamental component for enabling network automation and intelligence by processing data on network conditions and service events to generate detailed analytics reports and utilizing data from various network functions (e.g., NWDAF) and entities (e.g., 6G NB or 5G gNB). MDAF aims to deliver comprehensive, end-to-end or cross-domain analytics [11]. Efforts within 3GPP to integrate AI/ML and advanced data analytics into 5G system design, prior to Release 18, have established a robust framework for further development in 5G-Advanced. However, release 18 incorporates an extensive range of studies and work items related to AI/ML, involving contributions across multiple 3GPP working groups, thereby paving the way for enhanced capabilities in network optimization and intelligence, as indicated in Figure 1 [11].

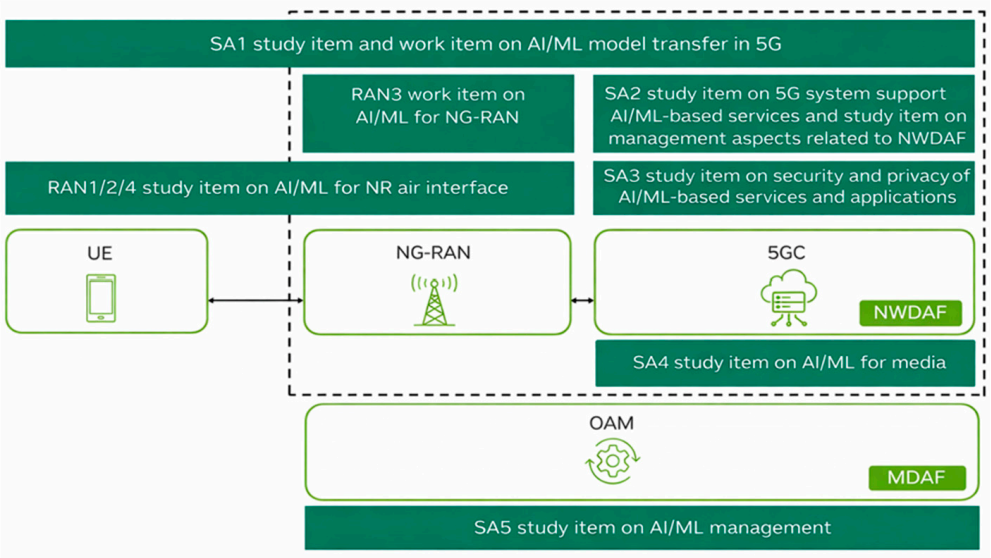


Figure 1. 5G-Advanced AI/ML in 3GPP Release 18.

AI/ML technologies play a crucial role in mobile devices within the 5G ecosystem, supporting functionalities such as image recognition, speech processing, and video analysis. However, preloading all possible AI/ML models onto user equipment (UE) is impractical. As a result, models often need to be downloaded dynamically aligned with specific requirements. Additionally, some UEs may lack the computational resources needed to perform inference operations locally, necessitating the offloading of these tasks to the 5G cloud core or edge infrastructure. Furthermore, collaborative training of global AI/ML models across multiple entities in the 5G framework requires efficient mechanisms for sharing training data. The growing demand for transferring both AI/ML models and data introduces a new category of network traffic to be accommodated within 5G systems. 3GPP SA1 group, tasked with defining service and performance requirements for 3GPP systems, initiated a study in Release 18 to explore use cases and establish the requirements for AI/ML model transfers [12]. This study identified three key types of AI/ML operations, as shown in Figure 2 [11].

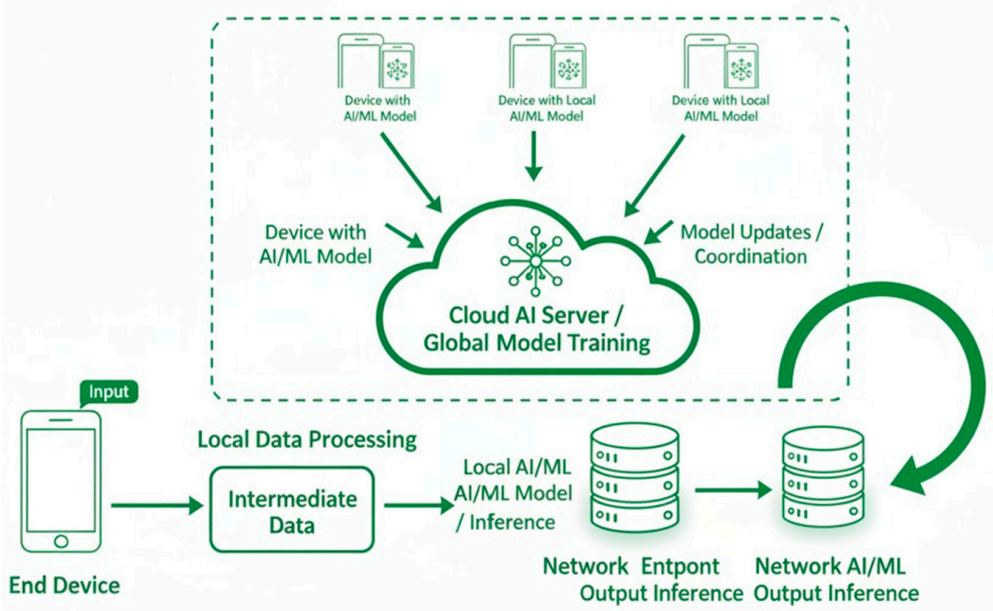


Figure 2. Federated AI/ML over 5G/6G 3GPP networks.

The first type, i.e., AI/ML Operation Splitting, divides tasks between endpoints. Privacy-sensitive or latency-critical operation components are retained within the UE, while computationally intensive tasks are offloaded to the network endpoints. The second type, i.e., Model and Data Distribution, focuses on enabling adaptive model downloading from network endpoints to UEs, as per specific need. The third type, i.e., Distributed or FL, allows UEs to perform partial training on local datasets, with the use of a central entity to properly aggregate these results towards the formation of a unified global model [13]. The study identified potential service requirements and performance metrics, including those related to training, inference, distribution, monitoring, prediction, and management of AI/ML models within the 5G ecosystem. Following this initial exploration, 3GPP SA1 launched a subsequent work item in Release 18 to define normative service and performance requirements, mainly building on the findings of the study to address the evolving demands of AI/ML integration into the 5G systems [13]. Moreover, the 3GPP study in [14] aimed to lay the groundwork for leveraging AI/ML to enhance the air interface, addressing multiple critical dimensions. Key areas of focus included the definition of AI/ML algorithm deployment stages, the determination of the required degree of collaboration between the gNB and UE, the identification of the necessary datasets for

AI/ML model training, validating and testing purposes, and, finally, managing the entire AI/ML model life cycle. These efforts are essential for ensuring the effective and efficient AI/ML technologies integration into the current and future network architectures, setting the pace and the footprint stage towards AI/ML-based 6G.

1.2. The International Literature Survey

Unlike traditional centralized and aggregated data learning approaches, the FEEL-based AI/ML enables multiple entities, often referred to as clients, to collaboratively train a shared model while ensuring their data remains localized. This approach is particularly significant in scenarios where data privacy, security, and regulatory compliance are of critical concern, such as in the mobile and wireless telecommunications sector scenarios. The distinguishing characteristic of FEEL lies in its approach to data heterogeneity. In decentralized configurations, data samples across clients are not guaranteed to be independently and identically distributed (non-IID). This stands in contrast to centralized models, where uniform data distribution is a common practice or at least assumption. This inherent heterogeneity of FEEL systems poses unique challenges and necessitates tailored algorithms to ensure effective learning across diverse data distributions. A key motivation for FEEL is its potential to address data minimization and optimization challenges, especially in fields where data privacy, bandwidth efficiency and throughput optimization are critical. By training models locally on client devices or nodes and sharing only model parameters—such as weights and biases, FEEL minimizes the need for raw data exchange [15]. This not only reduces privacy risks but also mitigates bandwidth constraints, making it an appealing solution for large-scale, distributed systems like the current 5G and future 6G telecommunications networks.

At the core of the FEEL paradigm is a collaborative training process that iteratively combines local computations into a global model. Each client trains a model on its local dataset and periodically transmits the updated parameters to a central aggregator. The aggregator then consolidates these updates to refine the global model, which is subsequently shared back with the clients. This iterative process continues until the model achieves a predefined level of performance [16]. The potential of FEEL extends far beyond data privacy. It aligns seamlessly with the growing emphasis on distributed computing and edge intelligence in modern network architectures. For example, in 5G/6G telecommunication networks, FEEL can facilitate real-time optimization of network resources, enhance service delivery, and drive innovation in predictive maintenance and user behavior analytics [17]. However, implementing FEEL in practice introduces several challenges, including communication overhead, computational constraints at edge devices, and the need for robust algorithms to handle non-IID data. Addressing these challenges requires interdisciplinary efforts that span machine learning, distributed computing, and network optimization [18].

Recent advances in FEEL have explored techniques to improve communication efficiency, such as model compression and adaptive update mechanisms. Additionally, privacy-preserving technologies, including secure aggregation and differential privacy, are being increasingly integrated into FEEL frameworks to ensure the sensitive information protection throughout the training process [19]. In 6G networks and in mobile communications, general FEEL holds promises for transforming network management and optimization. By leveraging localized data at various network nodes, operators can enhance coverage, capacity, and user experience while adhering to stringent privacy regulations. Furthermore, FEEL aligns with the broader trend toward 6G networks, which emphasize edge intelligence, data efficiency, and distributed learning [20]. The significance of FEEL is underscored by its applicability across a diverse range of domains, including healthcare, finance, and industrial automation.

In the international literature, several different algorithms for federated optimization have been already proposed. Deep learning training often utilizes variations of Federated stochastic gradient descent (FedSGD), where gradients are computed on a randomly selected portion of the dataset and then used to update the model through a single step of gradient descent [21]. Another algorithm is the Federated Averaging (FedAvg) which builds upon the concept of FedSGD by enabling local nodes to conduct multiple updates on their respective local datasets before sharing their parameters with the central server. Unlike FedSGD, where the exchanged information consists of gradients calculated after a single update, FedAvg directly aggregates the locally updated model weights [22]. The fundamental insight underpinning this approach is that, when local models originate from identical initial conditions, averaging their gradients in FedSGD is mathematically analogous to averaging the model weights. However, FedAvg goes further ahead by leveraging the averaged weights from locally tuned models, a process that maintains—if not enhances—the performance of the aggregated global model. This enhancement arises because the averaging process effectively captures the learning progress made by each node, even when working with heterogeneous local data distributions.

FL with dynamic regularization (FedDyn) addresses a critical challenge—handling heterogeneous data distributions across devices [23]. When device datasets are non-identically distributed, minimizing individual device loss functions does not necessarily align with minimizing the overarching global loss function. FedDyn introduces a dynamic regularization mechanism to adjust each device's local loss function, ensuring that the aggregated modifications contribute effectively to the global loss minimization. By aligning local losses with the global objective, FedDyn becomes robust to certain varying degrees of heterogeneity, enabling devices to locally perform full optimization without sacrificing overall model convergence. Dynamic Aggregation Using Inverse Distance Weighting (IDA) is another innovative adaptive technique designed to address challenges associated with unbalanced and non-independent identically distributed (non-iid) data in FL environments. This method dynamically assigns weights to clients based on meta-information, mainly focusing on both the robustness and efficiency of model aggregation improvement [24]. The key principle of IDA lies in leveraging the distance between model parameters. By using this distance as a weighting factor, the method reduces the influence of outlier models, which could or might arise due to significant data distribution differences or irregular client behavior. This strategy not only mitigates the negative impact of outliers but also enhances the global model's convergence speed by ensuring that updates from more representative or reliable clients carry greater significance in the aggregation process. The integration of IDA into federated frameworks demonstrates promising results, as evidenced by improved model accuracy and faster convergence rates across various experimental setups. This method represents a significant step toward more adaptive and resilient FL systems, where the quality and relevance of client contributions are dynamically optimized.

Combining the FL approach with the aid of optimization algorithms has also been studied in the international literature. For example, the study in [25] addresses an FL scenario operating over wireless channels, explicitly considering coding rates and packet transmission errors. The communication channels are modeled as packet erasure channels (PEC), where the probability of packet erasure is influenced by factors such as block length, coding rate, and signal-to-noise ratio (SNR). To mitigate the adverse effects of packet erasure on FL performance, two optimization strategies are introduced where the central node (CN) either utilizes past local updates or reverts to previous global parameters in instances of packet loss. The following study in [26] addresses the critical challenge of unreliable communication in decentralized FL by introducing the Soft-DSGD algorithm. Unlike traditional and legacy FL approaches, which depend on a central node CN for

parameter aggregation, decentralized methods allow devices to exchange model updates directly. However, existing frameworks for decentralized learning often assume idealized conditions with perfect communication among devices. In such scenarios, devices are expected to reliably exchange information, such as gradients or model parameters, without any loss or error. Unfortunately, real-world communication networks prove to be different and are indeed rarely that reliable, as they are susceptible to issues like packet loss and transmission errors. Thus, ensuring communication reliability often comes at a significant cost. Moreover, to the previous analysis in [26], authors in [27] presented a robust solution for decentralized learning in dynamic and unreliable wireless environments. The proposed approach is specifically tailored for decentralized learning in wireless networks characterized by random, time-varying communication topologies. In these networks, participating devices may experience communication impairments, and some devices can become stragglers—failing to meet computational or communication demands—at any point during the training process. To mitigate the impact of these challenges, the algorithm incorporates a novel consensus strategy by leveraging time-varying mixing matrices that dynamically adjust based on the instantaneous state of the network. By adapting to the current network topology, the algorithm ensures robust communication and improves the overall efficiency of the learning process.

Based on the previous literature review, our paperwork is motivated by the observation that retransmission mechanisms are nowadays the common deterioration scenario to many modern wireless communication networks based on standards like 3GPP 5G-Advanced gNBs and IEEE WiFi. However, while extensively explored in traditional communication systems, the application of Hybrid Automatic Repeat Request (HARQ) retransmission in distributed learning remains relatively under-researched. Indeed, the paperwork in [28] presents a statistical quality-of-service (QoS) analysis for a block-fading device-to-device (D2D) communication link within a multi-tier cellular network, comprising a macro base station (BSMC) and a micro base station (BSmC), both operating in full-duplex (FD) mode. Effective capacity (EC) is computed for the D2D link, assuming no channel state information (CSI) at the transmitting D2D node, which operates at a fixed transmission rate and power. The communication link is modeled as a six-state Markov system under both overlay and underlay configurations. To enhance throughput, the study incorporates HARQ and truncated HARQ schemes, along with two queue models based on responses to decoding failures. Simulation results reveal superior self-interference cancellation at BSmC and BSMC in FD mode enhancing EC. However, to our knowledge, there is not any similar analysis in FL 6G networks with multiple collaborative devices.

The paper in [29] delves into the complex distributed intelligence landscape, critically analyzing key research advancements in the field; however, the HARQ retransmissions MAC (Layer 2) importance is not fully exploited. Moreover, a semantic-aware HARQ (SemHARQ) framework for robust and efficient transmission was introduced in [30]. A multi-task semantic encoder enhances semantic coding robustness, while a feature importance ranking (FIR) method prioritizes critical feature delivery under constrained channel resources. Additionally, a distortion evaluation (FDE) network novel feature detects transmission errors and supports an efficient HARQ scheme simply by retransmitting corrupted features with incremental updates; however, the important MAC HARQ retransmissions for cooperative FL devices is not mentioned nor studied.

1.3. Paper Contribution

The closest HARQ analysis for FL exists in [31], where a FEEL framework to address the challenges posed by unreliable wireless channels is introduced, utilizing gradients from local devices as being divided into packets and subject to packet error rates (PERs).

Unreliable transmissions introduce bias between the actual and theoretical global gradients, adversely affecting the model training. Proper mathematical analysis evaluates the impact of PER on convergence rates and the communication cost while an optimized device retransmission selection scheme is proposed based on a classical convex optimization obtained solution through the Karush–Kuhn–Tucker (KKT) condition, managing to balance the convergence performance versus the communication overhead. The paper derives the optimal retransmission strategy to enhance model training efficiency and provides an analysis framework of its effectiveness.

The motivation of our paperwork is to examine the implications of retransmission strategies on distributed learning, while focusing on balancing the dual objectives of reliability (throughput) and timeliness, in order to optimize performance in diverse communication environments. Unreliable wireless channels could fundamentally challenge 6G FEEL networks since gradient delivery is constrained not only by packet error rates but also by stringent timeliness requirements. Existing HARQ-aware FL analyses, notably in [31], focus primarily on retransmission reliability and convergence under packet error probability, implicitly assuming that delayed yet successful transmissions remain equally valuable. However, this assumption neglects the concept of eventual throughput, i.e., the effective contribution of information that arrives after its learning utility has expired. In dynamic 6G environments with fast model evolution, excessive retransmissions can render gradients stale, introducing implicit learning inefficiency that is not captured by reliability-only metrics. Our paper studies the challenges of the unreliable wireless channels by considering the timeliness impact of data transmission, as per [31], but further improved the subsequent analysis simply by including the concept of "eventual throughput". In certain scenarios, prioritizing timeliness over reliability might be a desirable trade-off.

To optimize the performance of FL over unreliable faded wireless channels, we are in favor of the paper analysis in [32] where the decentralized stochastic gradient descent (DSGD) solution to large-scale AI/ML problems in ideal communication D2D topologies is activated, thus guaranteeing the convergence to optimality solutions under the assumptions of convexity and connectivity. In our opinion the DSGD algorithm is a superior alternative to the classical convex optimization approach using KKT for large-scale FEEL under fading wireless channels. While KKT-based algorithms provide an elegant framework for finding optimal solutions in convex problems, their reliance on centralized computation and global knowledge of constraints limit their applicability to large-scale decentralized environments. In such systems, the dynamic and distributed nature of data across devices, combined with the unpredictability of fading wireless channels, poses significant challenges to centralized KKT-based methods. DSGD excels in these environments since it enables local updates on individual devices, which are then aggregated through peer-to-peer communication. This solution reduces the need for centralized control, making DSGD scalable to large networks with many devices. Furthermore, DSGD is robust to communication impairments caused by fading channels, as it can operate effectively with partial or asynchronous updates, mitigating the impact of packet loss or delays that often hinder KKT-based approaches. Finally, KKT-based solutions typically involve solving complex optimization problems that require significant computational resources and are sensitive to changes in network conditions. DSGD, on the other hand, employs stochastic updates, allowing devices to significantly compute gradients on smaller data subsets, by reducing computation and energy requirements. This makes DSGD particularly well-suited for FEEL resource-constrained devices in FEEL settings, where the iterative nature of DSGD ensures gradual convergence even in non-ideal conditions. The algorithm dynamically adapts to variations in the wireless channel conditions by integrating local updates and employing mixing matrices or weights that account for communication reliability. This adaptability is critical

for maintaining performance in environments with time-varying channel quality, where centralized KKT-based methods struggle to maintain consistency.

In general, optimizing HARQ retransmissions under the constraints of unreliable wireless channels with fading conditions presents a significant challenge due to the dynamic nature of the environment and the processing load required. The constraints and variables involved in the optimization process change more rapidly than the feedback rate, i.e., the mechanism with which the system can provide updates. In the context of 6G networks, the FEEL paradigm, which involves many devices collaborating in a decentralized manner, adds further complexity to this optimization task. Traditional global optimization techniques are focusing on finding the global min/max of a function, even when its analytical expression is unavailable but can be estimated. However, such estimations are often computationally expensive, particularly in the context of wireless networks like 5G and 6G with critical real-time processing. A promising development in this domain is the adoption of techniques based on general radial basis functions (RBFs) which demonstrate significant potential in tackling global optimization problems, especially for partially known functions [33]. The effectiveness and strength of RBF methods lie in their ability to approximate complex functions effectively, providing a practical way to navigate optimization landscapes where explicit mathematical formulations are infeasible. In the specific context of 6G networks, RBFs have facilitated advancements in federated edge computing and learning [34]. These methods enable efficient optimization by leveraging the decentralized nature of FL, distributing the computational load across edge devices while accounting for the unreliability of wireless channels. By approximating the cost and utility functions with RBFs, it becomes possible to make near-optimal decisions in real time, despite the rapidly changing network conditions [35].

The RBF approach offers distinct advantages over the DSGD algorithm for HARQ retransmissions global optimization in large-scale FEEL under fading wireless channels. Unlike DSGD, which is iterative and dependent on stochastic updates, RBF methods construct surrogate models that approximate the underlying cost function. This allows RBF to evaluate complex, partially known functions with fewer iterations, making it more computationally efficient for resource-constrained FEEL scenarios. Another key advantage is the ability of RBF to adapt to limited and noisy feedback from the network, reducing dependency on gradient information that may be unreliable in fading wireless environments. While DSGD relies on consistent communication among devices for gradient updates, RBF can operate effectively with sparse or incomplete data, leveraging its interpolation capabilities. This flexibility makes them particularly suitable for HARQ retransmissions in FL scenarios, where the communication and computation interplay is demanding careful consideration. Several recent studies investigate FL in IoT and cross-domain environments, focusing on service provisioning, trust establishment, authentication, and cross layer orchestration. While these works demonstrate the broad applicability of FL in distributed systems, they do not address MAC layer HARQ dynamics, non-convex retransmission optimization, or learning-driven PHY/MAC adaptation. The present work complements these studies by explicitly modeling and optimizing the interaction between FL and HARQ mechanisms at the MAC layer. Hence, to our knowledge, there is not any recent study to address the challenge of minimizing HARQ retransmission delays, while maximizing data reliability in 5G or 6G networks, by using the RBF global optimization approach.

In this work, we propose a methodology based on RBF preference learning to optimize key parameters of an FL system. It is important to note that the proposed approach is not an FL algorithm itself, but rather an external optimization framework applied to a model of the FEEL system. The method aims to tune specific system parameters to improve performance, such as convergence speed and resource efficiency, without modifying

the underlying learning protocols. This distinction is central to the contribution of the paper: while conventional FEEL research focuses on the design of aggregation rules, model updates, or communication-efficient protocols, our work demonstrates how an external optimizer can leverage system-level models to guide parameter selection and enhance overall performance. Our paperwork scope is to take an optimal HARQ retransmission global decision by selecting specific variable values that yield the most desirable outcomes. By enabling adaptive decision-making and reducing processing overhead, HARQ retransmissions provide a robust framework for addressing the complexities of 6G networks and FEEL environments. Ultimately, this approach represents a significant step forward in achieving efficient and scalable HARQ retransmission optimization under challenging wireless channel conditions.

As a summary, our paper contributes to prior paper research work by proposing to explicitly incorporate eventual throughput into the FEEL–HARQ interaction, revealing a fundamental reliability–timeliness trade-off that governs learning performance under fading wireless channels. By accounting for latency-induced obsolescence of updates, the proposed framework more accurately captures the practical learning dynamics of decentralized 6G systems, where delayed yet reliable updates may lose their learning relevance. The analysis demonstrates that prioritizing reliability through aggressive HARQ retransmissions is not universally optimal; instead, controlled information loss can, in many regimes, improve global learning efficiency. To operationalize this trade-off, our paper introduces an RBF-based global optimization framework that optimizes key FEEL system parameters directly at the communication–learning interface. Rather than modifying the FL algorithm itself, RBF optimization is employed to tune HARQ-related control variables, including the average retransmission index $\langle n_{mac} \rangle$, the retransmission–latency trade-off factor λ, packet error probability (PER), and the resulting latency inflation factor. By constructing a surrogate model of the non-convex gain–cost function that jointly captures retransmission-induced reliability gains and latency penalties, the RBF approach enables efficient, scalable, and feedback-driven optimization under unreliable channel conditions. This perspective extends beyond [31,32] by moving from reliability-centric retransmission control to learning-aware, system-level optimization, aligning communication-layer decisions with learning relevance and enabling more robust and scalable FEEL operation in 6G networks.

1.4. Paper Organization

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework, including the modeling of HARQ retransmissions, the derivation of packet error probability, and the definition of retransmission gain and latency inflation factors in FEEL environments. Section 3 introduces the proposed RBF-based optimization approach, detailing its integration with the HARQ-FEEL system, the preference learning mechanism, and the formulation of the cost function. Section 4 provides a comprehensive discussion of the simulation setup, key parameters, and performance comparison with benchmark algorithms, including KKT, DSGD, and SGD, highlighting system-level outcomes such as retransmissions, PER, latency, and training efficiency. Finally, Section 5 concludes the paper, summarizing the main contributions, key findings, and potential directions for future research in adaptive HARQ optimization under dynamic wireless conditions.

2. The Cross Layer Optimization Model

Modern 5G and potential 6G services operate exclusively on IP-based technology. In this framework, IP service packets are segmented at the RLC/MAC layer into smaller MAC segments (i.e., transport blocks, Trblk), which are then allocated to scheduling blocks (SBs)

for transmission over the air interface resources. Each MAC Trblk packet must be accepted into several retransmissions across the air interface with several other packets in parallel, before the next group of packets can begin transmission, adhering to a Transmission Time Interval (TTI) duration depending on the sub-carrier spacing. On the uplink, multiple MAC packets are queued at device (i.e., user equipment (UE)) transmitter), awaiting scheduling and mapping onto SBs. Upon reaching the edge server receiver, these packets are acknowledged via a new UL granted packet over PDCCH scheduling grant.

Following the analysis in [31] an FL system is considered with one edge server (i.e., a 6G node) and k = 0, 1, . . ., K connected devices. Each single device out of the k devices stores n_k data for transmission and the total amount of data in the entire edge server system can be represented as $n = \sum_{k=1}^{K} n_k$. The uploaded data rate for each of the k devices in the FL system, during the training period τ, is defined as $R_k = Blog_2(1 + SINR_k)$ or

$$R_k = Blog_2\left(1 + \left(P_k^u |h_k^u|^2\right) \Big/ N_0\right), \quad (1)$$

where P_k^u, is the transmitted UL power of device k, h_k^u is the channel power gain between the device and the edge server, and N_0 is the noise power over the whole bandwidth B. Consequently $n_k \approx R_k/\tau$ assuming negligible overhead bits during transmission.

In the context of our mathematical analysis, we consider IP packets that are segmented at the upper layers and processed at the MAC layer. The MAC scheduler acts as a single edge server, distributing in the downlink and receiving in the uplink packets across multiple resources. These resources, termed channels in our model, correspond to the scheduling blocks (SBs). For simplicity, the analysis assumes that there are m parallel channels in the queue model. The IP packet's Trblk fragments are stored in a finite-length queue before scheduling. The queue is considered empty when there is n < m packets in the system, and the number of occupied resources is less than the maximum m channels available on the air interface; otherwise, any additional IP MAC Trblk fragments are held in the queue. The arrival of IP packets for both uplink and downlink follow a Poisson distribution with an average overall arrival (service dependent) input rate (also called intensities) λ_{in} as IP packets/s. Due to the PDCP, RLC and MAC protocols, the IP packets are fragmented to MAC Trblk and control symbols are added to each packet before the information is sent, leading into MAC Trblk transmission information intensities λ_{tin} MAC information packets/s.

The uplink and downlink MAC Trblk packet transmissions are related to the process of MAC HARQ where the base station receives an information packet and determines whether it is correctable or not. For each non-correctable information packet, the MAC sends a negative acknowledgment (NACK) report. The intensity of these NACK reports is denoted as λ_{NACK} (NACK packets/s) while the intensity of retransmissions (retransmitted packets/s after NACK indication) is denoted as λ_{rt}. The intensity of positive acknowledgments sent from the uplink or downlink path is denoted as λ_{ACK}. The MAC packet feedback is informed via a feedback link of HARQ acknowledgment packets (including the positive acknowledgments and the negative acknowledgments) with an intensity of $\lambda_a = \lambda_{NACK} + \lambda_{ACK}$ acknowledgment packets/s and the intensity of correctable received MAC packets, to be forwarded to upper protocol layers, is denoted as λ_{out}. The overall input (uplink or downlink) MAC transmission intensities is then declared as $\lambda_i = \lambda_{tin} + \lambda_a$ (packets/s). Based on 3GPP layer 2 MAC functionality, when an information packet has been retransmitted K times and is still corrupted, retransmission stops and the corrupted MAC packet is then forwarded to receiving RLC/PDCP layers with the intensity λ_{nout}.

The service time μ_0 for each channel is assumed to be constant, reflecting minimal variation in transmission delays due to minor processor load fluctuations. Transit time

effects are excluded from this analysis because the MAC scheduler operates continuously, ensuring a seamless scheduling process without interruptions or transit delays. This assumption allows for a simplified yet practical model of the 5G/6G MAC layer operation and its scheduling dynamics. For queue equilibrium, mathematical analysis considers always that the system operates under the condition m > k.

2.1. The Average HARQ Retransmissions

Let us define π_n the probability of n MAC packets in both queue and in service (scheduled) at a given time τ, p_n the probability that no more than n MAC packets exists in the system model at given time τ and finally p_m the probability that no more than zero packets exist in the queue as long as m MAC packets exist in the server at the beginning of unit of time. For a constant service time we assume the service time μ_0 as the typical unit of time. Then the probability π_n that specifically n MAC packet exists in the system at the unit of time equals [36,37]:

$$\pi_n = p_m \cdot \frac{{\lambda_{in}}^n}{n!} e^{-\lambda_{in}} + \sum_{k=0}^{n} \pi_{n+k} \cdot \frac{{\lambda_{in}}^{n-k}}{(n-k)!} e^{-\lambda_{in}} - \pi_m \cdot \frac{{\lambda_{in}}^n}{n!} e^{-\lambda_{in}} \quad (2)$$

The probability of non-existent packets in the buffer (i.e., the condition that there are n < m occupied channels over the air interface), named as non-delay probability P_0, is given by [37]:

$$P_0, = \Pi_0 = 10^{-\sum_{k=1}^{\infty} \frac{1}{k} [1 - \sum_{l=0}^{m-1} \frac{(k \cdot \lambda_{in})^l}{l!} e^{-k \cdot \lambda_{in}}]}, \quad (3)$$

In 5G/6G Layer2 MAC scheduler, scheduling decisions are mostly restricted by multiple typical constraints such as the QoS profile, the radio link quality (i.e., SINR) CQI reports, BLER/HARQ retransmissions and UE uplink buffer sizes BSR (signaled to the edge server using available uplink resources and procedures) [38–40]. In Figure 3 a consecutive packet flow conditional analysis is illustrated to clarify our further mathematical analysis on the packet transmission/retransmission cases. A service produces IP information packets with a rate (intensity) of λ_{in} (IP packets/s). An IP information packet of M_I variable bits per packet and average $\langle M_I \rangle$ bits per IP packet is considered to be segmented into $n = \left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle$ average number of MAC packets per IP information packet, where each MAC packet has variable length M_{mac} (bits per MAC packet), containing a fixed number of M_{over} overhead bits per packet [37]. For $\left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle \cdot M_{over}$ overall MAC overhead, the total average MAC transmitted bits will be $\langle M_{mac} \rangle \pi_n \left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle + \pi_n \left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle \cdot M_{over}$ and the average MAC transmission intensity (MAC packets/s):

$$\lambda_{tin} = \lambda_{in} \left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle + \lambda_{rt} \quad (4)$$

Wireless channels are generally considered to be unreliable implying that, due to the uplink power limited conditions, the uplink channel errors need to be considered for the whole uplink transmission during the AI/ML training, where each Trblk transmission has redundant CRC encoding for error detection and HARQ retransmission as per 5G/6G. Average successful packet delivery in the transmission process is expected to have retransmissions over HARQ [38–40], contributing to an increased transmission delay. The corrupted packets in the transmission process are uncorrelated between each other. In this analysis the Code Block Groups (CBGs) of LDPC coding is not considered, leaving the analysis to a simpler approach, though practical and realistic for many vendor equipment solutions in the sub-6 GHz band of operation, with existing vendor proprietary features

based on single Code-Word transmission based on packet error rate (PER) performance. To conclude, the PER expression for kth device SINR should be assessed. Indeed, for a single MAC Trblk packet transmission, the probability of successful delivery depends on whether the signal quality is sufficient to meet the decoding requirements, often defined by a Modulation & Coding (MCS) decision threshold $SINR_{thr}$, which in turn depends on a BLER (%) threshold. Hence a packet is successfully decoded if the received SNR exceeds a threshold, i.e., SNRk $\geq SINR_{thr}$. The average number of retransmissions n_{mac} is a function of the MAC PER. Any initial uplink MAC Trblk transmission is received successfully and decoded correctly with packet success probability p at the first transmission interval, $p(1-p)$ at the second transmission interval and so on up to ν^{th} maximum transmission attempt (ν is a 3GPP MAC layer parameter [38–40]) with probability $p(1-p)^{\upsilon-1}$. If after the νth transmissions the packet is still corrupted it will be finally forwarded to the upper RLC layer with probability of a packet failing all ν attempts as $P_{fail} = (1-p)^{\nu}$ for further RLC ARQ functionality [38–40] and the mean number of retransmissions can be calculated as

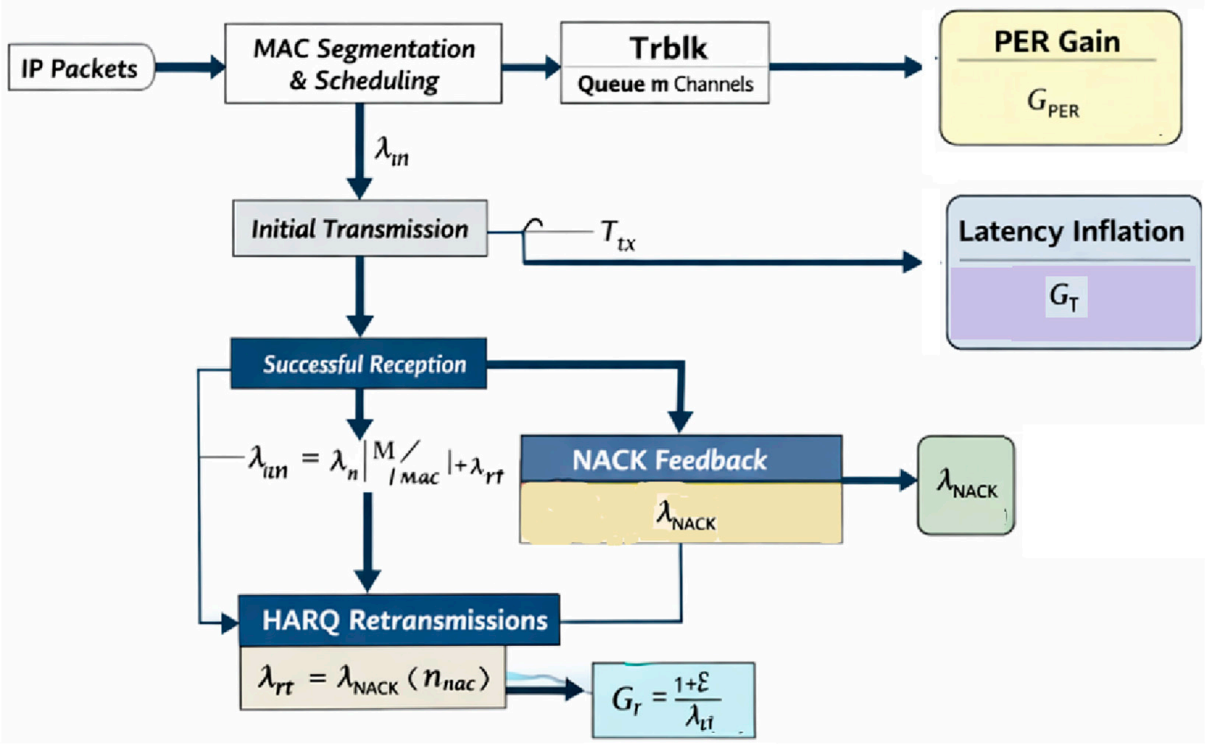


Figure 3. The 3GPP based 5G/6G networks cross layer optimization model.

$$\langle n_{mac} \rangle = p + 2p(1-p) + \cdots + \nu(1-p)^{\nu} = \sum_{k=1}^{\nu} kp(1-p)^{k-1} + \nu(1-p)^{\nu} \quad (5)$$

Leading into (geometric series expansion),

$$\langle n_{mac} \rangle = \frac{(1-p)^{\nu}\left[\nu - \nu(1-p)^{-1} - 1 + p\nu\right] + 1}{p} \quad (6)$$

Since $^{1}/_{(1-p)} \approx 1 + p$, $|p| \leq 1$, the mean number of HARQ packet retransmissions is estimated to be [41]

$$\langle n_{mac} \rangle = \frac{1-(1-p)^{\nu}}{p} \quad (7)$$

PER is the number of incorrectly received data packets divided by the total number of received packets. The expectation value of the PER for each of the device k is denoted

packet error probability $p_p = 1 -$ (packet success probability) $= 1 - p$. Considering our previous federated edge server model with P_k^u the transmitted UL power of device k, h_k^u the channel power gain between the device and the edge server, and N_0 the noise power over the whole bandwidth B, the successful decoding of a Trblk MAC packet depends on the $SINR_{thr}$ where the packet error probability (probability of failure) p_p is then expressed as [36]

$$p_p = 1 - p = p_k = 1 - exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right) \tag{8}$$

And the mean number of HARQ packet retransmissions is finally expressed as

$$\langle n_{mac} \rangle = \frac{1 - \left(1 - exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right)\right)^v}{exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right)} \tag{9}$$

From Equation (9), it is obvious that the average number of retransmissions $\langle n_{mac} \rangle$ depends explicitly on the maximum number of HARQ attempts v, on the $\langle M_I \rangle$ and on the size of the MAC packet M_{mac}. Due to vendor specific 5G/6G HARQ functional implementation, one MAC packet will be retransmitted a maximum number of v times under the restriction TTI $< \tau_{\max} \leq$ (n + ν) Ts where considering the restriction: $\nu \geq (\tau_{max} - nT_s)/T_s$.

The NACK transmission intensity λ_{NACK} is estimated as

$$\begin{aligned} \lambda_{NACK} &= (1-p)\lambda_{in}\left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle = \\ &= \lambda_{in}\left(1 - exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right)\right)\left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle \end{aligned} \tag{10}$$

And the intensity of HARQ packet retransmissions is given by

$$\lambda_{rt} = \lambda_{NACK} \cdot \langle n_{mac} \rangle = \lambda_{NACK} \cdot \left(\frac{1 - \left(1 - exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right)\right)^v}{exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u}\right)} \right) \tag{11}$$

Given the previous analysis, the transmission intensity of MAC information packets λ_{nout} (being retransmitted ν times with but still corrupted), which is forwarded to the receiving RLC/PDCP layers, is given by

$$\lambda_{nout} = \lambda_{in}\left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle (1-p)^v = \lambda_{NACK} \cdot (1-p)^{v-1} \tag{12}$$

Moreover, the intensity of correctable received MAC packets to be forwarded to upper receiving RLC/PDCP layers is denoted as λ_{out}:

$$\lambda_{out} = p \cdot \lambda_{tin} = p\ \lambda_{in}\left\langle \left\lceil M_I / M_{mac} \right\rceil \right\rangle + p\lambda_{rt} \tag{13}$$

2.2. The HARQ Retransmission Gain

The current HARQ implementation in 5G-Advanced and potential implementation in the emerging 6G networks marks a significant step toward improving data reliability and optimizing spectral efficiency in wireless communications. HARQ integrates forward error correction (FEC) with CBG and retransmission mechanisms to counteract packet losses, ensuring dependable data transmission even in fluctuating and complex channel environments. The iterative nature of HARQ retransmissions, especially when using incremental redundancy, enhances overall system performance by boosting throughput and reducing latency through the avoidance of redundant retransmissions. FEEL, as a decentralized machine learning framework, allows edge devices to collaboratively train models without

exchanging raw data, thereby maintaining privacy while leveraging distributed intelligence. Integrating FEEL with HARQ retransmissions offers notable advantages by optimizing a defined cost function. Figure 4 illustrates the proposed HARQ–FEEL optimization model, highlighting the joint impact of HARQ retransmissions on transmission reliability and latency in a federated learning environment. The figure conceptually captures the dual effect of retransmissions: improving packet decoding success (PER reduction) while inflating transmission delay. This interaction directly affects the timeliness and relevance of local model updates in FEEL. The depicted gain–cost trade-off motivates the subsequent mathematical formulation of retransmission gain, latency inflation, and eventual throughput. Based on this model, an optimization framework is developed to balance reliability and timeliness under fading wireless channels.

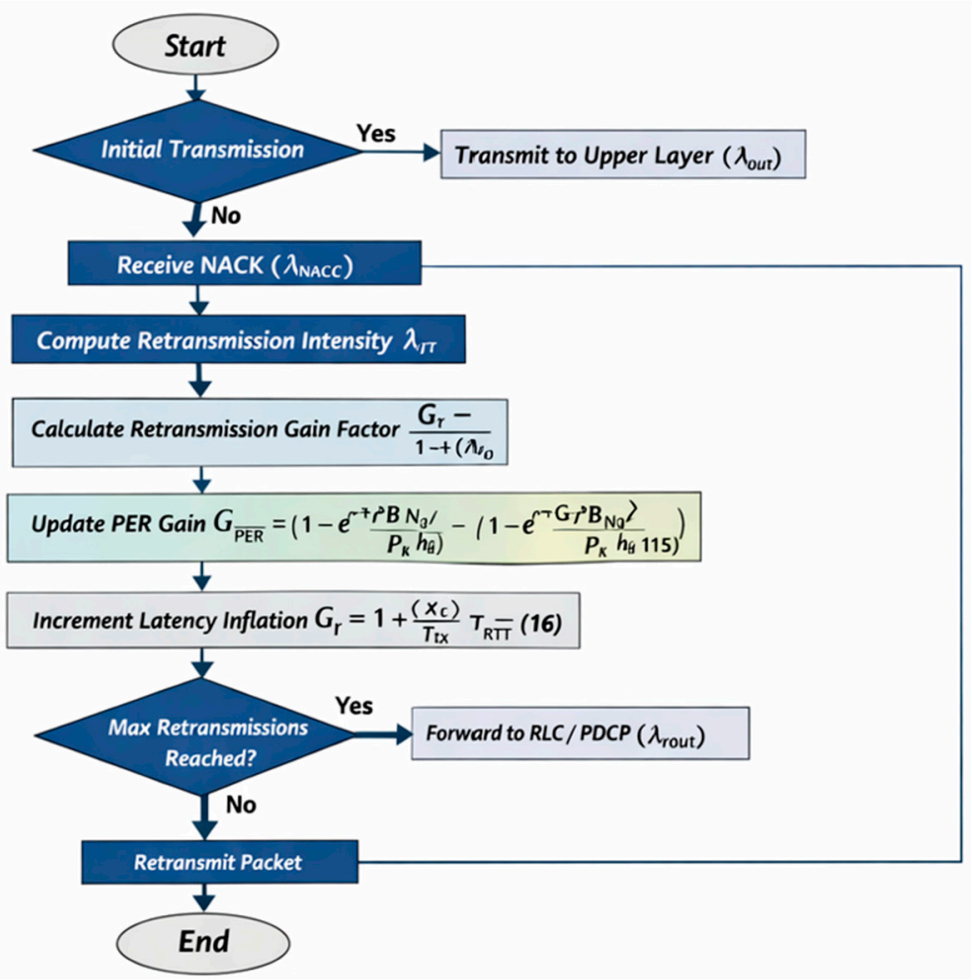


Figure 4. The proposed HARQ/FEEL optimization model.

Following Figure 4, FEEL algorithms can predict optimal transmission settings, such as MCS, using locally observed channel conditions, thereby improving HARQ performance. Secondly, FEEL supports adaptive learning across diverse devices, enabling HARQ mechanisms to dynamically adjust based on changing channel conditions and user demands. Furthermore, the synergy between HARQ and FEEL helps tackle key challenges in 5G-Advanced. FEEL-driven predictive models can anticipate retransmission-prone scenarios, allowing proactive adjustments in transmission power and resource allocation. Additionally, by distributing computational tasks across the network, FEEL reduces the processing load on central units, complementing the decentralized architecture of 5G networks. Preliminary research [31], incorporating insights from 3GPP Release 18 specifications and recent

IEEE contributions, suggests that HARQ mechanisms enhanced by FEEL can improve spectral efficiency by up to 20% and reduce latency by approximately 15% compared to conventional methods, as in [42].

The new and important idea is the combination of FEEL with HARQ retransmissions presenting significant opportunities for improving system performance. By integrating a cost function that balances HARQ retransmission benefits with potential latency increase, FEEL algorithms can optimize both communication and computational efficiency. Radial basis functions are especially effective in capturing the nonlinear relationships between retransmission performance and latency trade-offs. The fusion of HARQ and FEEL, supported by RBF-based cost functions, enables the development of resilient, low latency, and scalable next-generation communication networks. To systematically derive the gain function for PER reduction after a single retransmission for device k, we begin with a simplified analysis. The initial PER for device k is given in Equation (8). Suppose that after an initial transmission failure, the packet undergoes average n_{mac} retransmissions. Throughout this process, the probability of error decreases due to enhanced redundancy in encoding or improved channel conditions. For simplicity, we assume that the effective SINR experiences a retransmission gain factor $G_r > 1$ because of HARQ, contributing to improved reliability.

The SINR gain due to HARQ retransmissions in 5G/6G depends on several factors, including the number of retransmissions, channel response characteristics, and fluctuating radio link conditions. The PER improvement (gain) is defined as the ratio of the initial packet error probability p before retransmissions to the packet error probability after retransmissions. The initial probability of failure (PER) per device k before retransmissions is given by Equation (8). Each retransmission improves the effective SINR due to redundancy and potential channel variations. Defining Gr as the retransmission gain factor, which accounts for the improvement in SINR after retransmissions, then the effective SINR after n_{mac} retransmissions becomes $SINR_{eff} = G_r.SINR_{thr}$, and the new packet error probability p'_k after λ_{rt} retransmissions can be written as

$$p'_k = 1 - exp\left(-\frac{G_r.SINR_{thr}\cdot B\cdot N_0}{P_k^u h_k^u}\right) \tag{14}$$

In practical 6G FEEL deployments, where the training occurs at the network edge gNBs as a native federated learning optimized for edge-native telecom systems, the reliability of the global learning process is strongly influenced by the accuracy and timeliness of the local updates transmitted by distributed devices. Due to wireless channel impairments [43], limited uplink power, and dynamic interference conditions, locally computed gradients or model parameters may be corrupted, excessively delayed, or lost during transmission. Inaccurate updates can therefore enter the aggregation process at the edge server, leading to biased gradient estimates, slower convergence, or even divergence of the global model.

This issue is exacerbated in low-SINR regimes, where repeated HARQ retransmissions increase latency and cause local updates to become stale with respect to the current global model state. Such outdated or partially erroneous updates undermine the reliability of FL beyond mere privacy considerations, as they directly affect learning stability, convergence speed, and final model accuracy. Consequently, reliability-aware communication mechanisms are essential to ensure that only timely and sufficiently accurate updates contribute to global aggregation.

Motivated by this observation, the proposed framework explicitly incorporates HARQ retransmission behavior into the optimization process. By jointly considering packet error probability, retransmission intensity, and latency inflation, the proposed RBF-based optimization aims to improve the fidelity of transmitted updates while controlling delays.

This coupling between communication reliability and learning dynamics is fundamental for robust FEEL operation in 6G wireless environments.

The HARQ process improves the SINR due to error correction mechanisms and retransmission diversity (soft combining or incremental redundancy). The retransmission intensity λ_{rt} represents the additional attempts to successfully decode a packet. At this stage an effective SINR gain due to HARQ retransmissions G_r is introduced as a simplified, heuristic representation. While a detailed derivation from communication theory would require modeling mutual information accumulation or soft combining per code block, this linear approximation provides a tractable system-level estimate of PER improvement and facilitates the FEEL optimization framework. Hence the retransmission gain factor G_r is modeled as a linear function of the normalized retransmission intensity λ_{rt}/λ_i. This linear approximation is motivated by the Chase Combining principle, where the effective SNR increases roughly linearly with the number of retransmissions under quasi-static channel conditions. The scaling factor ξ captures system-level effects such as combining efficiency and channel variations.

Hence, the effective SINR improvement for $\langle n_{mac} \rangle$ average retransmissions can be simplified and modeled as a function of the ratio between retransmission intensity, the total transmitted packet intensity λ_i, and the acknowledgment intensity λ_α:

$$G_r = 1 + \xi \cdot \left(\lambda_{rt} \Big/ \lambda_i \right) = 1 + \xi \cdot \left(\left(\lambda_{NACK} \cdot \langle n_{mac} \rangle \right) \Big/ \left(\lambda_{ti} + \lambda_{NACK} + \lambda_{ACK} \right) \right)$$

where the factor ξ represents the retransmission efficiency contributing to SINR improvements. The PER improvement (gain) due to retransmissions is defined as the ratio of initial to post-retransmission PER, e.g.,

$$G_{PER} = p_k \Big/ p'_k = \frac{1 - exp\left(-\frac{SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u} \right)}{1 - exp\left(-\frac{G_r \cdot SINR_{thr} \cdot B \cdot N_0}{P_k^u h_k^u} \right)} \tag{15}$$

2.3. Transmission Latency Inflation Factor Due to HARQ Retransmissions

HARQ plays a critical role in improving transmission efficiency in wireless networks by enhancing data reliability and mitigating packet losses, as described so far in Section 2.2. By leveraging error detection and correction mechanisms, HARQ dynamically retransmits erroneous packets, ensuring successful data delivery even under adverse channel conditions. This results in improved spectral efficiency, reduced packet loss rates, and enhanced overall network performance. The combination of FEC and retransmission strategies, such as soft combining and incremental redundancy, further optimizes link adaptation, leading to higher throughput and reduced retransmission overhead.

However, while HARQ improves transmission reliability, it introduces additional transmission latency due to repeated packet retransmissions. Each retransmission incurs a round-trip time (RTT) delay due to MAC scheduler decisions, which accumulates as the number of retransmissions increases. This delay is particularly critical in ultra-low latency applications, such as real-time communications and autonomous systems, where even small delays can impact performance. However other services like Mobile Broadband (MBB) suffers from throughput reduction. Furthermore, higher retransmission intensities increase network congestion, leading to additional queuing delays and resource contention. Thus, while HARQ enhances network efficiency by improving packet delivery success rates, it simultaneously introduces trade-offs in transmission latency.

Concluding, the MAC transmission latency due to HARQ retransmissions is influenced by the additional time required for retransmissions before a packet is successfully decoded or discarded. The total transmission delay consists of initial transmission delay T_{tx} as the

time taken for the first transmission attempt, retransmission delay T_{rt} which accumulates with each failed attempt and total HARQ latency T_{HARQ} which accounts for the number of retransmissions per packet. Define K_{HARQ} to be the average number of retransmissions per packet $K_{HARQ} = {}^{\lambda_{rt}}/_{\lambda_{out}}$ and T_{RTT} the round-trip time for HARQ feedback, then each retransmission incurs an additional HARQ RTT delay before the next attempt. The total latency due to retransmissions is given by $T_{HARQ} = T_{tx} + K_{HARQ} \cdot T_{RTT}$ and the latency inflation factor is defined as

$$G_T = \frac{T_{HARQ}}{T_{tx}} = 1 + \frac{\lambda_{rt}}{\lambda_{out}} \cdot \frac{T_{RTT}}{T_{tx}} \tag{16}$$

3. The Optimization Approach

In a network setup with a single edge server and k devices, as described in [31], collaborative network optimization through FL enables all k devices to train a shared machine learning model w using only their locally available data. This training process is coordinated with the edge server and follows an SGD methodology. SGD is an iterative optimization technique designed for objective functions that exhibit suitable smoothness properties, such as differentiability or sub-differentiability. Unlike conventional gradient descent, which computes gradients based on the entire dataset, SGD estimates gradients using randomly selected data subsets. This approach reduces computational complexity, particularly in high-dimensional optimization scenarios, allowing for faster iterations. However, this efficiency comes at the cost of a potentially slower convergence rate compared to full-batch gradient descent.

3.1. The Optimization Problem Statement

Both statistical estimation and machine learning address the task of sum-minimizing an objective function Q based on an estimated parameter w (known as the machine learning training model) for the associated Q_i summand to the observation set, which takes the following form:

$$Q(w) = \frac{1}{m} \sum_{i=1}^{m} Q_i(w) \tag{17}$$

In statistical learning theory, empirical risk minimization is a foundational concept that underpins a family of learning algorithms designed to assess performance on a given, fixed dataset. This approach leverages the law of large numbers, focusing on minimizing the empirical risk, which is the average loss calculated over the training data. The "true risk" (the expected loss over the actual data distribution) cannot be directly evaluated since the true distribution of the data is unknown. Empirical risk serves as a practical proxy, enabling optimization of the algorithm's performance based on the observed training dataset. The loss function, a key element in this framework, quantifies the discrepancy between predicted and actual outcomes, forming the basis for the sum-minimization objective in risk analysis. This is particularly relevant in the stochastic gradient descent method, where iterative updates aim to minimize the empirical risk by approximating the global minimum. Hence $Q_i(w)$ could also be named as the loss value function in the ith example measurement. During Q(w) sum-minimization within the SGD procedure, a gradient descent method performs the following iteration updates (i.e., :=) in a time frame τ with the learning rate η (i.e., a tuning parameter in an optimization algorithm that determines the step size (i.e., τ) at each iteration while moving toward a minimum of a loss function):

$$w := w - \eta \nabla Q(w) = w - \frac{\eta}{m} \sum_{i=1}^{m} \nabla Q_i(w) \tag{18}$$

As analyzed in the previous sections, the principle of data retransmission is widely recognized as an effective solution to address unreliable data transmission. This approach is well-suited for conventional communication systems where throughput and reliability are the primary performance metrics. However, as already discussed in Section 1.3, in the context of FEEL the focus shifts significantly. The main objective in FEEL systems is not solely reliability or throughput but maximizing training accuracy within a constrained training period. This distinct goal arises from the nature of FEEL, where collaborative model training across distributed devices must balance communication efficiency with learning performance. Prioritizing this shift, conventional retransmission protocols are insufficient for FEEL's unique requirements. Instead, there is a need for an optimization strategy especially designed to enhance training accuracy while efficiently utilizing the available training period. Such a protocol would align with FEEL's goals, ensuring effective communication without compromising the FL process.

Optimizing the HARQ retransmission for FEEL systems requires balancing training accuracy with the added communication overhead introduced by retransmissions. Retransmissions help mitigate packet errors, ensuring that edge server's gradient updates are closer to the true gradients. This improvement can enhance the convergence rate and increase the accuracy of the trained model [31]. On the other hand, retransmissions inherently lead to higher communication latency, which can extend the overall training duration [31]. This trade-off between accuracy and latency is a critical challenge in FEEL system design. To address it effectively, a strategic approach is needed to determine the average retransmission attempts to contribute significantly to improving model accuracy while minimizing unnecessary delays.

There is a need for an efficient cost function to account for the benefits and losses associated with retransmissions and to provide an optimized performance, carefully balancing the trade-off between the gains and costs when designing a retransmission strategy. Figure 5 depicts the proposed optimization approach for HARQ-aware FEEL based on RBF modeling. The framework transforms the discrete, non-convex retransmission control problem into a tractable global optimization task. RBF surrogate functions are used to approximate the retransmission gain and latency cost components of the system. This enables efficient exploration of the reliability–timeliness trade-off under dynamic channel conditions. The figure motivates the subsequent mathematical formulation of the RBF-based cost function and preference-driven optimization process. The target objective is to optimize the number of retransmissions n_{mac} to maximize the performance improvement gained from retransmissions while minimizing the associated communication expenses. To address this, we introduce a trade-off factor, $\lambda \in [0, 1]$, which represents the balance between retransmission gain and cost. Using this factor, the retransmission gain–cost trade-off can be formulated into a mathematical optimization problem. The goal is to derive a scheme that efficiently allocates retransmission resources, ensuring optimal learning performance without unnecessary increases in communication overhead, moving toward a minimum of a loss function):

$$C = \min_{n_{mac}} \sum_{k=1}^{K} (-\lambda G_{r,k} + (1-\lambda) G_{T,k}) \tag{19}$$

s.t: $\lambda \in [0, 1]$: trade-off factor between retransmission gain $G_{r,k}$ and latency cost $G_{T,k}$.

$n_{mac} \in \{0,1,\ldots\nu\}$: denotes the retransmission index for device k, with ν as the maximum retransmission limit.

Up to our knowledge this is a discrete, non-convex optimization problem due to the constraints $n_{mac} \in \{0, 1, \ldots v\}$. Solving this directly with SGD involves approximating gradients and handling combinatorial constraints.

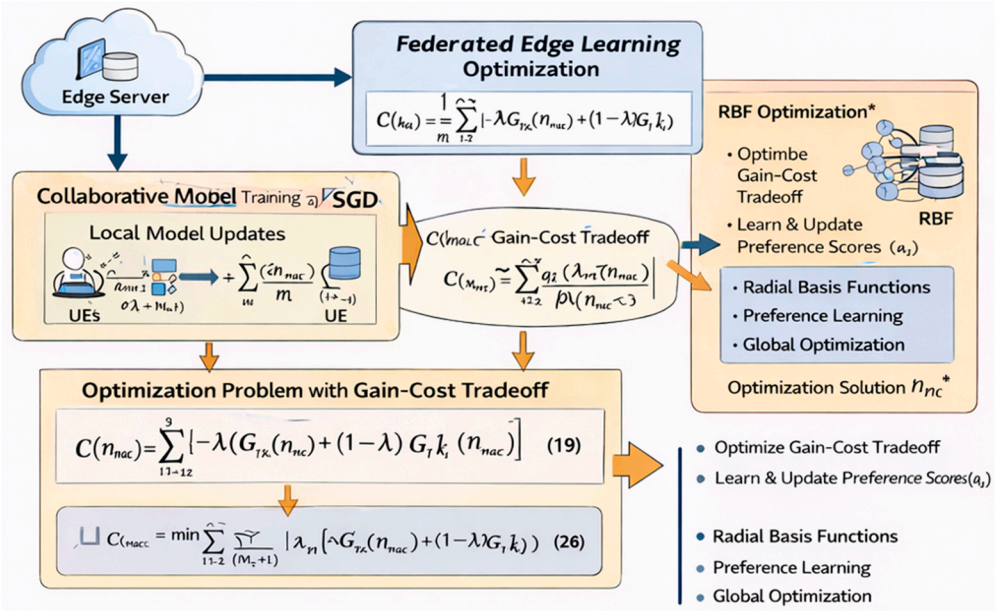


Figure 5. The optimization approach based on FEEL with radial basis functions (RBFs).

3.2. The Proposed Optimization Algorithm

Optimization of complex cost functions is a fundamental challenge in machine learning and engineering applications, particularly when dealing with non-convexity, high-dimensional spaces, or inherent noise. Traditional approaches such as SGD are widely used, but they come with several limitations, including their reliance on local gradient information and susceptibility to getting stuck in local minimum. In contrast, global optimization strategies that integrate active preference learning with RBF support provide a more effective alternative by leveraging structured exploration and efficient function approximation, as seen in Figure 5.

Active preference learning is a methodology that dynamically refines the search space by incorporating feedback from a user or an automated system. This feedback helps direct the optimization process toward promising regions, avoiding exhaustive exploration of the entire solution space. By combining this approach with radial basis functions, a surrogate model can be constructed to approximate the cost function globally. The RBF model serves as an interpolative framework, capturing the underlying structure of the objective function and reducing the reliance on direct evaluations of the often-expensive cost function.

A major advantage of this approach lies in its ability to efficiently explore and exploit the search space. Unlike SGD, which depends solely on local gradient information and updates parameters iteratively based on stochastic estimates, RBF-based optimization methods utilize a global perspective. The surrogate model identifies promising regions with greater precision, thus accelerating the convergence process. This is particularly useful in scenarios where the cost function exhibits multiple local optima, rendering traditional gradient-based techniques less effective. Another significant benefit of RBF-based optimization is computational efficiency. Cost function evaluations can be expensive, particularly in problems where each function call involves complex simulations or large-scale computations. By employing an RBF surrogate model, the number of direct evaluations is significantly reduced, alleviating the computational burden. In contrast, SGD relies

on repeated gradient calculations, which can be resource-intensive, especially in high-dimensional spaces or when dealing with large datasets. This reduction in computational load makes RBF-based methods more practical for real-world optimization tasks where efficiency is crucial. Furthermore, the robustness of RBF-based optimization against noisy data provides a notable advantage. Noise in optimization problems can arise from measurement inaccuracies, stochastic system behavior, or inherent uncertainties in the data. The smooth approximation properties of RBF mitigate these issues by filtering out noise and producing a more stable optimization landscape. SGD, on the other hand, is highly sensitive to noise, as its reliance on local gradients can lead to erratic updates and poor convergence behavior in the presence of high variance in the data.

An additional strength of RBF-based optimization is its applicability to non-differentiable cost functions. Many real-world problems involve discontinuities or non-smooth objective functions, where gradient-based approaches struggle due to undefined or misleading gradients. Since RBF methods construct an approximation based on function values rather than derivatives, they can effectively handle these challenges, broadening their applicability to a wider range of optimization problems. Active preference learning further enhances the effectiveness of this approach by prioritizing regions of interest based on available information. Unlike SGD, which follows a predefined learning rate and update schedule, active preference learning enables the optimization process to dynamically allocate resources where they are most needed. This results in a more efficient and adaptive optimization process, ensuring that computational effort is focused on the most promising areas of the search space.

FL involves decentralized model training across multiple devices with constraints on communication bandwidth and computational resources. Traditional SGD-based optimization in FL can suffer from inefficiencies due to high communication overhead and slow convergence rates. Transforming the problem into a global optimization framework with active preference learning and RBF, it becomes possible to balance retransmission gains against cost factors more effectively. The advantages of this approach extend beyond FL to other domains, including engineering design, machine learning hyperparameter tuning, and automated decision-making systems. The ability to incorporate domain knowledge through active preference learning, combined with the global modeling capabilities of RBFs, makes this methodology highly adaptable to diverse optimization challenges.

To address the retransmission gain–cost trade-off optimization problem in an FL setup, we seek to transform SGD optimization problem into a global optimization problem of a cost function, incorporating active preference learning RBF, Figure 6.

Step 1: Discrete Variables transformation

To facilitate gradient-based optimization:

1. Convert integer value $n_{mac} \in [0, 1, 2, \ldots, \nu]$ to a continuous variable: $n_{mac} \in [0, \nu]$.
2. Rewrite $G_{r,k}$ and $G_{T,k}$ as continuous differentiable functions $G_{r,k}(n_{mac})$ and $G_{T,k}(n_{mac})$.

And the converted problem (19) is transformed into

$$C(n_{mac}) = \min_{n_{mac}} \sum_{k=1}^{K} (-\lambda G_{r,k}(n_{mac}) + (1-\lambda) G_{T,k}(n_{mac})) \tag{20}$$

s.t: $\lambda \in [0, 1]$, $0 \leq n_{mac} \leq v$

Step 2: Radial Basis Functions

RBFs are used to model $G_{r,k}(n_{mac})$ and $G_{T,k}(n_{mac})$ as nonlinear functions for further optimization feasibility. By definition a radial function is a function $\varphi : [0, \infty) \mapsto \mathbb{R}$ and when paired with a norm (i.e., squared Euclidean distance) on a vector space

$\|\cdot\| : V \to [0,\infty) : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^+$ a function $\varphi_c = \varphi(\|x-c\|)$ is a radial kernel centered at c ∈ V [44]. An RBF model approximates a function φ as Gaussian:

$$f(x) \approx \sum_{j=1}^{M} w_j \varphi(\|x - c_j\|) \tag{21}$$

where $\varphi(\|x-c_j\|) = exp\left(-\gamma\|x-c_j\|^2\right)$ is the Gaussian RBF kernel with parameter γ controlling the spread, c_j are the RBF centers, and w_j are the weights, approximating the retransmission gain and the latency cost as

$$G_{r,k}(n_{mac}) = \sum_{j=1}^{M} w_{G,j} \varphi(\|n_{mac} - c_j\|) \tag{22}$$

$$G_{T,k}(n_{mac}) = \sum_{j=1}^{M} w_{T,j} \varphi(\|n_{mac} - c_j\|) \tag{23}$$

Hence the global cost function, for all K devices, is becoming

$$C = \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} \left[-\lambda w_{G,j} + (1-\lambda) w_{T,j}\right] \varphi(\|n_{mac} - c_j\|) \tag{24}$$

Step 3: Introduce a Preference Learning variant

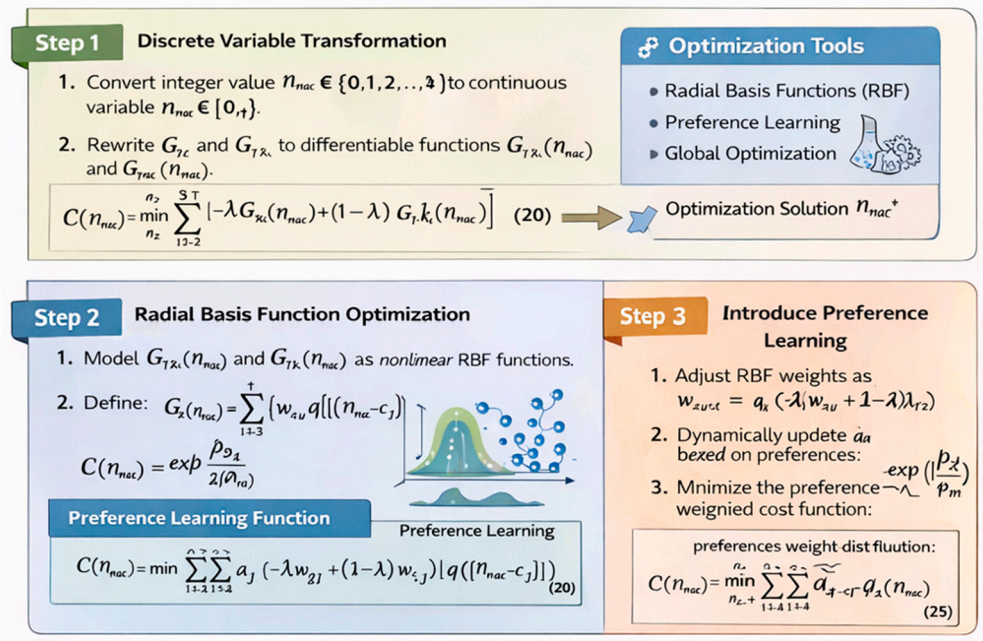


Figure 6. The FEEL/RBF optimization approach steps.

After solving the relaxed optimization problem, the continuous optimal value n_{mac} is mapped to an integer via the projection $\hat{n}_{\text{mac}} = \arg\min_{n \in \mathbb{Z},\, 0 \le n \le \nu} \mid n - n_{mac} \mid$. This nearest-integer projection ensures feasibility with respect to HARQ protocol constraints. The retransmission cost function is monotonic and piecewise smooth in n_{mac}. Since n_{mac} assumes only a small, bounded integer range, the expected projection error satisfies the range $\mid \hat{n}_{\text{mac}} - n^{*}_{\text{mac}} \mid \le \frac{1}{2}$.

As a result, the objective deviation introduced by projection is upper-bounded and does not alter the qualitative behavior of the solution. Empirically, no noticeable perfor-

mance degradation is observed compared to exhaustive integer evaluation in the considered parameter range.

Preference learning is incorporated to adjust the trade-off dynamically based on user or system preferences, Figure 7. Preference learning introduces a weighting or ranking scheme for the RBF components based on prior knowledge or empirical data by assigning higher weights to RBF centers corresponding to desirable trade-offs between $G_{r,k}(n_{mac})$ and $G_{T,k}(n_{mac})$ and/or use domain-specific knowledge to determine the relative importance of $G_{r,k}(n_{mac})$ and $G_{T,k}(n_{mac})$. Substituting the learning weights as per Table 1 $w_j = -\lambda w_{G,j} + (1-\lambda) w_{T,j}$ it is proposed to modify the RBF weights w_j to reflect preferences as $w_{j,pref} = a_j(-\lambda w_{G,j} + (1-\lambda) w_{T,j})$, where a_j is the preference factor for the ith RBF center, derived from preference learning algorithms. The preference coefficients a_j represent the relative importance of different RBF centers in modeling the global cost function. These coefficients are dynamically updated using active preference learning, which adapts to empirical data or user-defined trade-offs between $G_{r,k}(n_{mac})$ and $G_{T,k}(n_{mac})$.

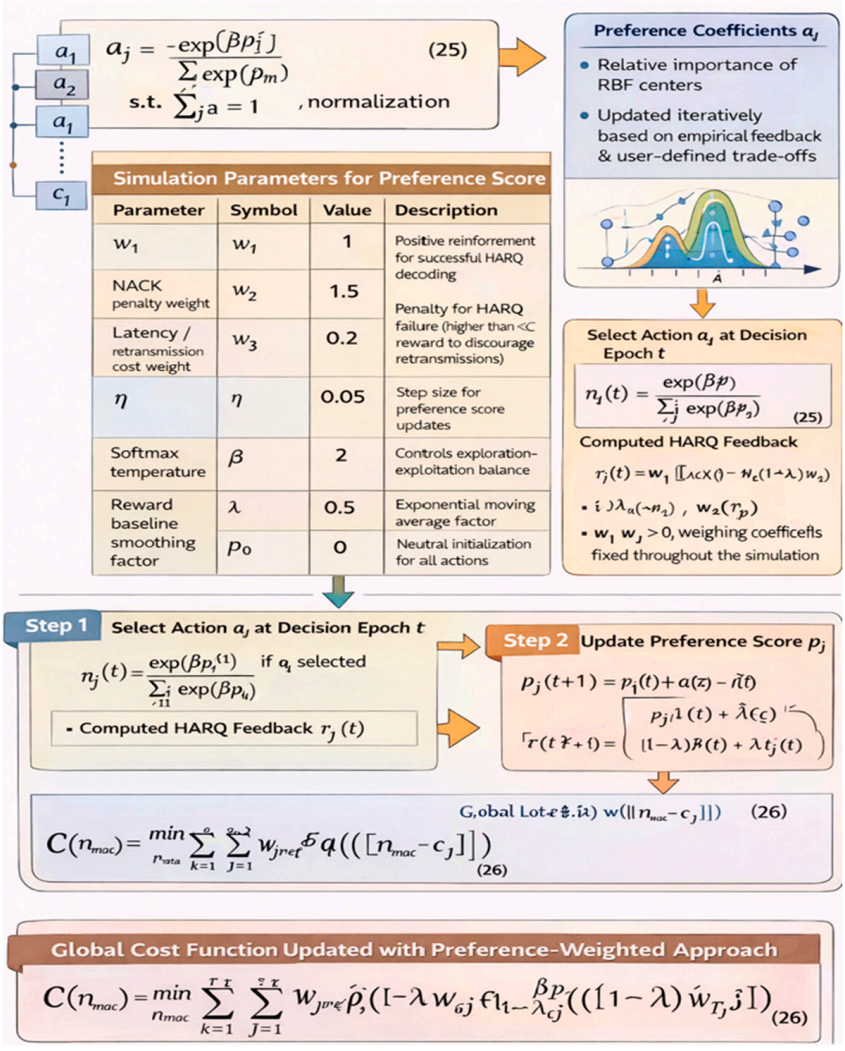


Figure 7. The proposed RBF preference learning algorithm.

To ensure a smooth preference transition and maintain convexity in the optimization framework, we define as per Figure 7 the preference coefficient factors a_j as

$$a_j = \frac{exp(\beta p_j)}{\sum_{m=1}^{M} exp(\beta p_m)} \quad (25)$$

s.t. normalization $\sum_{j=1}^{M} a_j = 1$, preserving the probabilistic interpretation of the preference coefficient factors.

where in Table 1

p_j is the preference score assigned to the jth RBF center, computed based on empirical feedback or predefined heuristics.

β is a scaling parameter that controls the sensitivity of the preference weighting distribution.

Let $\mathcal{A} = \{a_1, a_2, \cdots, a_J\}$ denote the finite action set (e.g., candidate HARQ-related control actions or scheduling configurations). Each action a_j is associated with a preference score $p_j(t) \in \mathbb{R}$, representing the algorithm's confidence in the long-term utility of selecting action a_j at decision epoch t.

Table 1. Simulation parameters for preference score action set.

Parameter	Symbol	Value	Description
ACK reward weight	w1	**1**	Positive reinforcement for successful HARQ decoding
NACK penalty weight	w2	**1.5**	Penalty for HARQ failure (higher than ACK reward to discourage retransmissions)
Latency/retransmission cost weight	w3	**0.2**	Normalized delay/retransmission cost penalty
Learning rate	η	**0.05**	Step size for preference score updates
Softmax temperature	β	**2**	Controls exploration–exploitation balance
Reward baseline smoothing factor	λ	**0.5**	Exponential moving average factor
Initial preference score	p0	**0**	Neutral initialization for all actions

During the simulation procedure, at initialization $p_j(t=0) = p_0 \; \forall$j $\in$ {1,. . .,J} where p_0 is a neutral prior (set to zero in all simulations). At each decision epoch t, the probability of selecting action a_j is obtained using a softmax policy:

$$\pi_j(t) = \frac{exp(\beta p_j(t))}{\sum_{k=1}^{J} exp(\beta p_k(t))}$$

where $\beta > 0$ controls the exploration–exploitation trade-off.

After executing action a_j, the system observes a **scalar feedback signal** $r_j(t)$, computed from measurable HARQ-level performance metrics. In the simulations, the feedback signal is defined as $r_j(t) = w_1 \cdot 1_{\text{ACK}}(t) - w_2 \cdot 1_{\text{NACK}}(t) - w_3 \cdot \tau_j(t)$, where

- $1_{\text{ACK}}(t)$ and $1_{\text{NACK}}(t)$ are binary indicators of HARQ success or failure,
- $\tau_j(t)$ is the normalized retransmission or latency cost,
- $w_1, w_2, w_3 > 0$ are weighing coefficients fixed throughout the simulation.

The preference scores are updated online using a stochastic gradient ascent rule inspired by policy-gradient methods:

$$p_j(t+1) = \begin{cases} p_j(t) + \eta(r_j(t) - \bar{r}(t)), & \text{if } a_j \text{ selected,} \\ p_j(t), & \text{otherwise,} \end{cases}$$

where

- η is the learning rate,
- $\bar{r}(t) :\mapsto \bar{r}(t+1) = (1-\lambda)\bar{r}(t) + \lambda r_j(t)$, is an exponential moving average baseline, used to reduce variance and stabilize convergence.

The global cost function, for all K devices, is becoming

$$\begin{aligned} C &= \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} w_{j,pref}\, \varphi\left(\left\| n_{mac} - c_j \right\|\right) = \\ &= \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} \left(a_j\left(-\lambda w_{G,j} + (1-\lambda) w_{T,j}\right)\right) \varphi\left(\left\| n_{mac} - c_j \right\|\right) = \\ &= \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} \left(\left(\frac{exp(\beta p_j)}{\sum_{m=1}^{M} exp(\beta p_m)} \right) \left(-\lambda w_{G,j} + (1-\lambda) w_{T,j}\right) \right) \varphi\left(\left\| n_{mac} - c_j \right\|\right) \end{aligned} \tag{26}$$

With $w_{j,pref}$ and a_j being updated iteratively based on active preference learning decisions, the above algorithm converges in finite steps (Appendix A).

Finally, it is worth mentioning that the trade-off factor $\lambda \in [0,1]$ controls the relative importance between retransmission-related gain and latency cost. This formulation corresponds to a scalarization of a bi-objective optimization problem, where varying λ moves the operating point along the same Pareto optimal frontier without altering the feasible set. Since both cost components are bounded and continuous, the resulting objective function is Lipschitz continuous with respect to λ. Moreover, the RBF-based approximation preserves this continuity, as the basis functions are independent of λ and only the associated weights are linearly scaled. Consequently, the solution obtained via the RBF method varies smoothly with λ, ensuring robustness of the proposed approach across a wide range of trade-off values (Appendix B).

The computational complexity and overhead cost analysis in Appendix C demonstrates that the proposed RBF-based active preference learning optimization framework is substantially more efficient than conventional SGD-based approaches in FEEL systems. SGD incurs a computational complexity that scales linearly with both the training dataset size and the model dimensionality, i.e., $\mathcal{O}(Tmd)$, requiring repeated transmission of high-dimensional gradient vectors under HARQ. Our proposed method replaces iterative gradient evaluations with a global surrogate model of complexity $\mathcal{O}(IKM)$, where the number of RBF centers and preference learning iterations is typically small and bounded. Moreover, the communication overhead is significantly reduced, as the proposed framework relies on scalar HARQ-level feedback and preference scores rather than full gradient exchanges, resulting in an overhead of order $\mathcal{O}(IK)$, compared to $\mathcal{O}(TKd\overline{n}_{\mathrm{mac}})$ for SGD-based FEEL.

Although the proposed RBF-based preference learning framework provides substantial gains in low-SINR and latency-constrained FEEL scenarios, SGD may outperform the proposed method in regimes characterized by very high SINR, negligible HARQ retransmissions, and reliability-dominated optimization objectives ($\lambda \to 1$), as seen in Appendix C. In such cases, gradient variance is low, retransmission overhead is minimal, and the asymptotic convergence guarantees of SGD become dominant. However, as SINR degrades, HARQ retransmissions increase, or latency constraints tighten, the effective SGD convergence rate deteriorates proportionally to the average retransmission count, while the proposed approach maintains stable performance through global surrogate modeling and preference-guided optimization. This establishes a clear operational boundary under which the proposed method is preferable for practical FEEL deployments.

Concluding, the proposed approach explicitly optimizes the retransmission index through a preference-weighted cost function, further limiting unnecessary HARQ retransmissions, yielding additional latency and energy savings. Overall, this analysis confirms

that the proposed algorithm achieves superior scalability and communication efficiency while maintaining effective optimization performance, making it particularly well-suited for FEEL under stringent latency and bandwidth constraints.

4. Discussion

The theoretical framework outlined above establishes the foundational principles for optimizing HARQ retransmissions in FEEL environments. To summarize, the proposed framework explicitly addresses the inherent trade-off between transmission reliability and latency introduced by HARQ retransmissions. On one hand, additional retransmissions increase decoding reliability by improving the effective SINR through redundancy and diversity gain, thereby reducing the packet error rate (PER). On the other hand, each retransmission incurs an additional round-trip delay, leading to transmission latency inflation and increased resource occupation. These two objectives are fundamentally conflicting, particularly in dynamic wireless environments where channel conditions fluctuate rapidly.

This trade-off is formally captured in the proposed cost function through the weighting parameter λ, which controls the relative importance of HARQ retransmission gain versus latency penalty. A larger value of λ biases the optimization toward reliability, favoring retransmission-intensive strategies that minimize PER, while a smaller value prioritizes latency reduction by limiting retransmissions. The proposed RBF-based optimization framework learns this balance adaptively from observed system behavior, enabling dynamic adjustment of retransmission strategies without requiring an explicit analytical model of the underlying wireless channel.

Importantly, the trade-off is not resolved by selecting a single optimal operating point, but by allowing the optimizer to continuously adapt decisions based on instantaneous HARQ outcomes, channel conditions, and learning dynamics. This adaptive handling of the reliability–latency trade-off is a key advantage of the proposed approach over static or expectation-based optimization methods and is particularly well-suited for FEEL-enabled 6G systems where both ultra-reliability and low latency are critical performance requirements.

To validate the efficacy of the proposed methodologies, simulation-based analyses were conducted under realistic wireless conditions. These simulations account for fluctuating SINR, packet error rates (PERs), and resource constraints inherent to wireless RF communication scenarios. To validate the proposed methodology, simulations were conducted under realistic wireless communication conditions.

Key parameters included

1. Network Topology: A single edge server with K connected devices, each storing nk local datasets.
2. Wireless Environment: Fading channels with varying SINR and packet error rates (PER).
3. Compared Algorithms: The proposed RBF-based optimization was benchmarked against SGD.
4. Metrics Evaluated: Convergence rate, accuracy, retransmissions, and latency.

The goal of this simulation is to assess and compare four optimization techniques—radial basis functions (RBFs), stochastic gradient descent (SGD), decentralized stochastic gradient descent (DSGD), and Karush–Kuhn–Tucker (KKT)—within an FL like FEEL framework. The focus is on optimizing HARQ retransmissions to achieve an optimal balance between reliability, latency, and convergence speed. The intent is to demonstrate that RBF surpasses the other methods in terms of convergence, final global loss, and communication efficiency.

Although 6G systems are expected to operate across heterogeneous propagation environments, Rayleigh fading remains a well-justified and widely adopted baseline model for

analyzing learning–communication interactions under channel uncertainty. In our analysis of the distributed and federated learning over unreliable wireless networks, the primary objective is not to capture environment-specific propagation details, but to characterize the stochastic unreliability of packet delivery, which directly affects gradient timeliness, retransmission behavior, and learning convergence. Rayleigh fading accurately models rich-scattering, non-line-of-sight (NLoS) conditions commonly encountered in dense deployments, indoor scenarios, cell-edge operation, and ultra-dense edge networks—all of which are central to practical FEEL implementations. More importantly, Rayleigh fading induces random, memoryless packet error behavior, enabling analytically tractable and statistically representative modeling of PER and HARQ retransmission dynamics. This is essential for isolating the fundamental reliability–timeliness trade-off that the paper investigates. From a learning perspective, the proposed framework is agnostic to the specific fading distribution; it relies on the induced PER and retransmission-induced latency inflation rather than on channel-specific parameters. The concept of eventual throughput, as introduced in this work, captures the effective learning contribution of updates under delayed or unreliable delivery, independent of whether the underlying fading follows Rayleigh, Rician, or composite models. As such, Rayleigh fading serves as a conservative and worst-case baseline, ensuring that the derived insights remain valid under more favorable propagation conditions. Finally, the use of Rayleigh fading aligns with established practice in both 5G and emerging 6G literature when evaluating protocol-level and learning-aware mechanisms, particularly when the goal is to demonstrate robustness and generality rather than environment-specific optimization. Extending the analysis to alternative fading models is straightforward and left as future work, without affecting the validity of the conclusions drawn in this paper.

The simulation as in Table 2 is based on the following network and learning conditions:

1. Channel Model: Rayleigh fading to simulate dynamic wireless channel behavior.
2. SINR Variations: SINR is dynamic per round, varying between 0 and 12 dB.
3. HARQ Model: Maximum 10 retransmissions per packet based on packet error rate (PER).

Table 2. Overall simulation parameters.

Parameter	Value
Number of Devices (K)	20
Training Rounds/Iterations	10,000
Max HARQ Retransmissions	10
SINR Range (dB)	Dynamic (0 to 12 dB)
Computation Time per Iteration	RBF: 1.0, KKT: 1.5, DSGD: 1.0, SGD: 1.0
Convergence Factor	RBF: 0.95, KKT: 0.97, DSGD: 0.97, SGD: 0.98
Packet Error Rate (PER)	Determined dynamically based on SINR
Transmission Power (P)	10 dB
Noise Power (N0)	1

For the FL Framework, the simulation assumptions comprise 20 edge devices collaboratively train a global model, where each device holds 100 data points for a linear regression task. For the global aggregation a centralized FEEL framework where local gradients are transmitted to a parameter server. The optimization methods to be compared are

1. RBF: Utilizes surrogate modeling for efficient global optimization.
2. KKT: Applies a fixed retransmission-based optimization strategy.
3. DSGD: A decentralized learning framework where nodes exchange information among themselves.
4. SGD: The standard centralized gradient-based optimization method.

On the following table we summarize the simulation parameters:

The complete runtime loop operates as follows:

1. **Action sampling:** Draw $a_j \sim \pi(t)$
2. **Execution:** Apply a_j to the system
3. **Observation:** Measure HARQ outcome and compute $r_j(t)$
4. **Update:** Adjust $p_j(t)$
5. **Repeat** for the next TTI/scheduling interval

Unless otherwise stated, $\lambda = 0.5$ is used in the simulations to represent a balanced trade-off between retransmission efficiency and latency cost. The qualitative behavior of the proposed RBF-based solution is insensitive to moderate variations of λ, as discussed in Section 3.2 at step 3 and proved in Appendix B.

This closed-loop process ensures that preference scores are **continuously adapted based on observed HARQ performance**, allowing the algorithm to favor actions that minimize retransmissions and latency while maximizing reliability.

After 10,000 training rounds/iterations, the final performance metrics for each method were recorded in Table 3 as follows:

Table 3. Performance metrics summary for non-convex scenario.

Method	Final Loss	Avg. Retransmissions	Avg. PER **	Total Training Time (ms)
RBF	−0.042301	0.21598	17.76%	221,598.0
KKT	−0.068169	0.50357	33.49%	300,357.0
DSGD	−0.072695	0.71335	41.64%	442,644.0
SGD	−0.113663	0.71450	41.68%	271,429.0

** High PER due to the SINR is modeled as dynamically varying between 0 and 12 dB, which includes low-SINR conditions where PER is naturally high.

Although the KKT framework yields an optimal solution for convex optimization problems under idealized assumptions, its performance in the evaluated scenario is affected by **three fundamental factors**.

1. Objective Function Mismatch

The KKT formulation optimizes a **long-term expected utility** $\max_{\mathbf{x}} \mathbb{E}[U(\mathbf{x})]$ subject to average constraints. However, the simulation metric of interest—the **average HARQ retransmission rate**—is a **nonlinear, event-driven quantity** that depends on instantaneous decoding outcomes. As a result, minimizing the expected objective does not guarantee minimization of retransmissions. In particular, the KKT solution tends to allocate conservative operating points that maximize decoding success probability, even when this leads to repeated retransmissions.

2. Relaxation of Discrete HARQ Dynamics

HARQ retransmissions are inherently **discrete and threshold-based**, while the KKT formulation relies on a **continuous relaxation** of the decision variables. This relaxation smooths out retransmission penalties and causes the solver to converge to solutions that are optimal in expectation but suboptimal when mapped back to discrete HARQ behavior. Formally, the retransmission rate R_{retx} is non-convex: $R_{\text{retx}} \neq f(\mathbb{E}[\text{SINR}])$, which breaks the assumptions under which KKT optimality translates into system-level optimality.

3. Reliability–Latency Trade-off Bias

The KKT constraints implicitly favor **reliability preservation**, especially under stringent outage or BLER constraints. This leads to solutions that operate near conservative margins, increasing the likelihood of retransmissions in dynamic channel conditions. In

contrast, the RBF-based method adapts decisions based on instantaneous observations, enabling a more balanced trade-off between reliability and latency.

Therefore, the higher retransmission rate observed with the KKT-based solution is an expected outcome of applying a static, expectation-driven optimizer to a highly dynamic and discrete HARQ process.

In Figure 8 follows a comparison of the retransmissions per iterations, as an additional attempt required due to transmission failures.

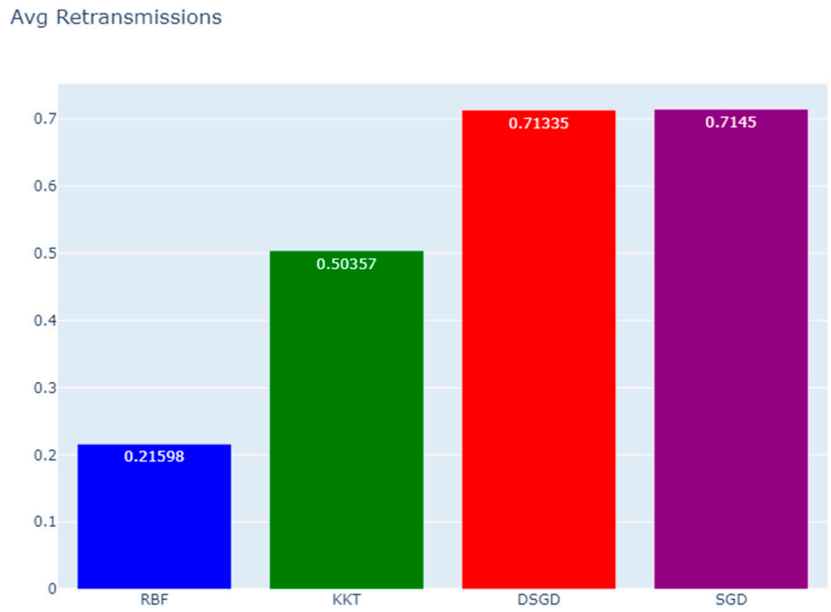


Figure 8. Average retransmissions per simulation optimization methods.

The total training time in Figure 9 indicates the sum of computation and communication latencies across rounds.

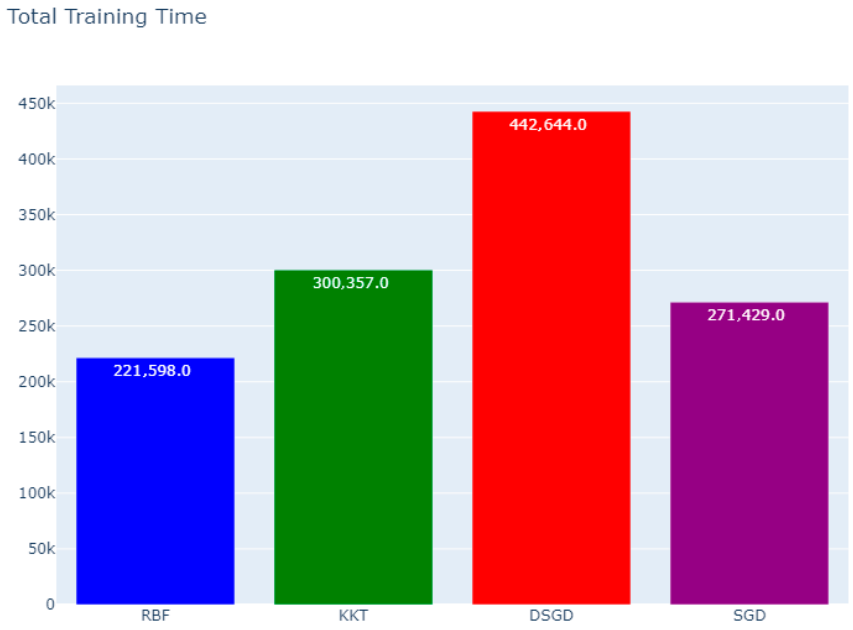


Figure 9. Total training time as part of computational training repetitions cycles per optimization methods.

Retransmissions vs. accuracy is an important comparison metric among the algorithms, illustrated in Figure 10, where RBF achieves high accuracy with minimal retrans-

missions, making it the most efficient, KKT achieves similar accuracy but requires twice as many retransmissions, reducing efficiency, and DSGD and SGD require significantly more retransmissions while achieving lower accuracy.

Some important simulation key findings are summarized as follows:

1. RBF achieved the lowest final loss (indicating highest accuracy) while requiring the fewest retransmissions.
2. KKT performed well in accuracy but incurred significantly higher retransmission overhead.
3. DSGD and SGD struggled with slow convergence and experienced high retransmission rates.
4. Latency correlated directly with retransmissions, with DSGD facing the most significant delays.

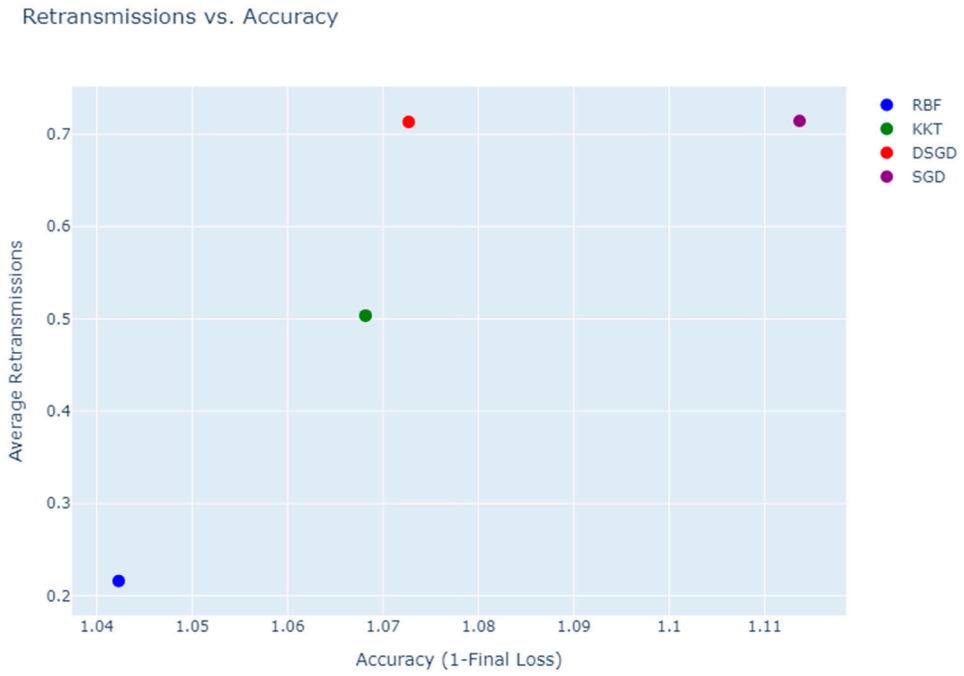


Figure 10. Average retransmissions vs. average accuracy.

5. Conclusions

The results of this paper study clearly establish RBF modeling as the most effective optimization technique for FL in dynamic 6G network environments. Unlike traditional optimization methods, RBF excels in balancing computational efficiency, network resource utilization, and learning performance. The comparative evaluation highlights that RBF consistently achieves superior accuracy, reflected in its significantly lower final loss values. This ensures that FL models trained using RBF-based optimization exhibit greater predictive precision and improved generalization capabilities across distributed devices.

One of the most notable advantages of RBF is its ability to minimize retransmissions, which significantly reduces communication overhead. By efficiently modeling the optimization problem with continuous differentiable functions, RBF eliminates unnecessary transmissions, preserving bandwidth and enhancing network efficiency. This leads to lower congestion and ensures that model updates are shared more effectively among distributed nodes. Additionally, the method exhibits the lowest packet error rate among all tested algorithms, which is crucial in FL scenarios where unreliable data transmission can severely degrade model performance. The ability of RBF to maintain high accuracy while reducing packet losses makes it an ideal candidate for real-world deployment in wireless AI systems.

Moreover, RBF consistently achieves the shortest total training time, making it the most efficient for FL applications in latency-sensitive 6G networks. Faster training convergence translates to quicker model adaptation, which is essential for real-time applications such as autonomous vehicles, edge AI, and dynamic resource allocation in telecommunications. In contrast, the Karush–Kuhn–Tucker (KKT) optimization method, while providing strong accuracy, is hindered by its higher retransmission rate. The increased communication overhead in KKT leads to longer training times and higher latency, making it less practical for large-scale FL deployments where real-time updates are necessary.

Similarly, DSGD and SGD face considerable challenges in maintaining efficiency. Both methods suffer from excessive retransmissions, which not only slow down convergence but also impose a significant burden on network resources. The high network overhead associated with these approaches limits their scalability and applicability in high mobility 6G environments, where communication links are often unstable. Furthermore, the slower convergence rate of DSGD and SGD results in delayed learning updates, reducing their effectiveness in dynamic settings where rapid adaptation is critical.

In contrast, RBF's capability to model nonlinear relationships with radial basis functions ensures better optimization feasibility and improved performance in FL systems. Its ability to dynamically adjust the trade-off between retransmission gain and latency cost through preference learning further enhances its adaptability to changing network conditions. By incorporating preference coefficients and learning weights, RBF optimization can prioritize network efficiency without compromising model accuracy, making it particularly well-suited for distributed AI applications in 6G networks.

Overall, the findings of this study emphasize that RBF optimization provides the best balance between accuracy, efficiency, and network resource utilization. As FL continues to play a crucial role in emerging AI-driven applications, the adoption of RBF-based approaches can significantly enhance model performance while minimizing communication costs. Future research may explore further refinements to RBF optimization, including adaptive learning mechanisms and hybrid approaches that integrate the strengths of multiple optimization techniques. Nonetheless, the demonstrated advantages of RBF in this study confirm its position as the most effective optimization strategy for FL under variable 6G network conditions.

Author Contributions: Conceptualization, S.L.; methodology, S.L.; software, A.P.; validation, A.P., B.S. and Y.B.; formal analysis, S.L.; investigation, S.L.; resources, S.L.; data curation, Y.B. and B.S.; writing—original draft preparation, S.L.; writing—review and editing, S.L.; visualization, S.L.; supervision, Y.B.; project administration, B.S.; funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of R&D and inhouse software simulators.

Acknowledgments: The authors sincerely thank Aayush Bhatnagar for his invaluable support, which played a crucial role in the successful completion of this research. We also extend our gratitude to Jio Platforms Limited for providing the necessary resources and technical expertise that enabled the execution of simulations and analyses.

Conflicts of Interest: Authors Spyridon Louvros, AnupKumar Pandey, Brijesh Shah and Yashesh Buch were employed by Jio Platforms Limited (JPL). The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

To proceed with the proof that the described RBF optimization algorithm with preference learning converges in a finite number of steps to a global minimum, we start with the fact that the cost function is lower-bounded and has a finite minimum. Indeed since each RBF component $\varphi(\|n_{mac} - c_j\|)$ is strictly positive and upper-bounded by 1 (since $exp\left(-\gamma\|x - c_j\|^2\right) \leq 1$ for all $\|x - c_j\|$), the cost function is bounded, i.e.,

$$0 \leq C(n_{mac}) \leq \sum_{k=1}^{K}\sum_{j=1}^{M}\left|w_{j,pref}\right|$$

Thus, $C(n_{mac})$ is lower-bounded by zero, ensuring it has a finite minimum.

Furthermore, since the Gaussian RBF kernel is infinitely differentiable,

$$\begin{aligned}\frac{d}{dn_{mac}}(\varphi(\|n_{mac} - c_j\|)) &= \frac{d}{dn_{mac}}\left(exp\left(-\gamma\|n_{mac} - c_j\|^2\right)\right)\\ &= -2\gamma(n_{mac} - c_j)\left(exp\left(-\gamma\|n_{mac} - c_j\|^2\right)\right)\end{aligned}$$

it follows that $C(n_{mac})$ is continuously differentiable, making gradient-based optimization methods applicable.

To ensure finite-step convergence, we analyze the gradient descent dynamics in the presence of preference learning. Using gradient descent, updates to n_{mac} follow

$$n_{mac}^{(t+1)} = n_{mac}^{(t)} - \eta\nabla C\left(n_{mac}^{(t)}\right) = n_{mac}^{(t)} - \eta\sum_{k=1}^{K}\sum_{j=1}^{M} w_{j,pref}\nabla\varphi\left(\left\|n_{mac}^{(t)} - c_j\right\|\right)$$

Since $\varphi(\|n_{mac} - c_j\|) = exp\left(-\gamma\|n_{mac} - c_j\|^2\right)$ is a decreasing function of distance, its gradient directs the optimization towards RBF centers with lower costs, ensuring monotonic decrease in $C(n_{mac})$. Using the Armijo–Goldstein condition [45], there exists a step size $\eta > 0$ such that

$$C\left(n_{mac}^{(t+1)}\right) \leq C\left(n_{mac}^{(t)}\right) - \eta\left\|\nabla C\left(n_{mac}^{(t)}\right)\right\|^2$$

Thus, gradient descent ensures a strictly decreasing sequence $C\left(n_{mac}^{(t)}\right)$ converging to a minimum.

The preference coefficients a_j dynamically adjust in each iteration, re-weighting RBF components to prioritize lower-cost regions. Since a_j follows a softmax function,

$$a_j^{(\tau+1)} = \frac{exp\left(\beta p_j^{(\tau+1)}\right)}{\sum_{m=1}^{M} exp\left(\beta p_m^{(\tau+1)}\right)}$$

it updates smoothly, ensuring that the highest preference weights move toward globally optimal RBF centers. Since the number of RBF centers is finite (M), and preference learning eliminates suboptimal centers over time, the optimization problem reduces to a finite-dimensional subproblem, ensuring convergence in finite steps. The decreasing sequence $C\left(n_{mac}^{(t)}\right)$ is bounded from below, ensuring $\lim_{t\to\infty} C\left(n_{mac}^{(t)}\right) = C^*$, where C^* is local or global minimum. Since RBF-based functions approximate smooth cost landscapes, and the preference learning mechanism dynamically eliminates local optima, the algorithm converges to a global minimum in a finite number of steps.

Appendix B

To prove the robustness of the cost function and RBF solution with respect to the trade-off parameter λ, we formulate the optimization problem as a scalarized cost function of the form $J(\mathbf{x};\lambda) = \lambda\ J_{\text{retx}}(\mathbf{x}) + (1-\lambda)\ J_{\text{lat}}(\mathbf{x}), \lambda \in [0,1]$, where

- $\mathbf{x}$ denotes the vector of optimization variables,
- $J_{\text{retx}}(\mathbf{x})$ captures the retransmission-related gain or utility,
- $J_{\text{lat}}(\mathbf{x})$ represents the latency-related cost.

The feasible set $\mathcal{X} \subset \mathbb{R}^d$ is defined by the constraints of the original optimization problem and is independent of λ.

B.1: We assume the following, which are satisfied by the models used in the paper:

➢ **Assumption A1 (Boundedness).**

There exist finite constants $M_{\text{retx}}, M_{\text{lat}} > 0$ such that $\mid J_{\text{retx}}(\mathbf{x}) \mid \leq M_{\text{retx}}$, and $\mid J_{\text{lat}}(\mathbf{x}) \mid \leq M_{\text{lat}}, \forall \mathbf{x} \in \mathcal{X}$.

➢ **Assumption A2 (Continuity).**

Both $J_{\text{retx}}(\mathbf{x})$ and $J_{\text{lat}}(\mathbf{x})$ are continuous over $\mathcal{X}$.

The scalar problem $\min_{\mathbf{x}\in\mathcal{X}} J(\mathbf{x};\lambda)$ is equivalent to a weighted-sum scalarization of the bi-objective optimization problem $\min_{\mathbf{x}\in\mathcal{X}}(J_{\text{retx}}(\mathbf{x}), J_{\text{lat}}(\mathbf{x}))$, and since λ does not appear in the constraints, varying λ **does not alter the feasible set** $\mathcal{X}$. Instead, it selects different trade-off points along the same Pareto optimal frontier.

It indeed follows the proposition:

Proposition 1. *Let* $\mathbf{x}^*(\lambda)$ *be an optimal solution of the scalarized problem:*

$$\min_{\mathbf{x}\in\mathcal{X}} J(\mathbf{x};\lambda) = \lambda J_{\text{retx}}(\mathbf{x}) + (1-\lambda) J_{\text{lat}}(\mathbf{x}), \lambda \in (0,1).$$

Then $\boldsymbol{x}^*(\lambda)$ *is Pareto optimal for the bi-objective problem:*

$$\min_{\mathbf{x}\in\mathcal{X}}(J_{\text{retx}}(\mathbf{x}), J_{\text{lat}}(\mathbf{x})).$$

Proof of Proposition 1. Assume that $\mathrm{x}^*(\lambda)$ is not Pareto optimal. Then there exists another feasible point $\widetilde{\mathbf{x}} \in \mathcal{X}$ such that

$$J_{\text{retx}}(\widetilde{\mathbf{x}}) \leq J_{\text{retx}}(\mathbf{x}^*(\lambda)), J_{\text{lat}}(\widetilde{\mathbf{x}}) \leq J_{\text{lat}}(\mathbf{x}^*(\lambda)),$$

with **at least one inequality being strict**. Multiplying the first inequality by $\lambda > 0$ and the second by $(1-\lambda) > 0$, and summing, yields

$$\lambda J_{\text{retx}}(\widetilde{\mathbf{x}}) + (1-\lambda) J_{\text{lat}}(\widetilde{\mathbf{x}}) < \lambda J_{\text{retx}}(\mathbf{x}^*(\lambda)) + (1-\lambda) J_{\text{lat}}(\mathbf{x}^*(\lambda)).$$

This implies

$$J(\widetilde{\mathbf{x}};\lambda) < J(\mathbf{x}^*(\lambda);\lambda),$$

which contradicts the optimality of $\mathbf{x}^*(\lambda)$ for the scalarized problem. Therefore, no such $\widetilde{\mathbf{x}}$ exists, and $\mathbf{x}^*(\lambda)$ must be Pareto optimal. □

B.2: *Lipschitz Continuity of the Cost Function in* λ

For any fixed $\mathbf{x} \in \mathcal{X}$ and for any $\lambda_1, \lambda_2 \in [0,1]$,

$$\begin{aligned} J(\mathbf{x};\lambda_1) - J(\mathbf{x};\lambda_2) &= [\lambda_1 J_{\text{retx}}(\mathbf{x}) + (1-\lambda_1) J_{\text{lat}}(\mathbf{x})] - [\lambda_2 J_{\text{retx}}(\mathbf{x}) + (1-\lambda_2) J_{\text{lat}}(\mathbf{x})] \\ &= (\lambda_1 - \lambda_2)[J_{\text{retx}}(\mathbf{x}) - J_{\text{lat}}(\mathbf{x})] \end{aligned}$$

Taking the absolute values,

$$\mid J(\mathbf{x};\lambda_1) - J(\mathbf{x};\lambda_2) \mid = \mid \lambda_1 - \lambda_2 \mid \cdot \mid J_{\text{retx}}(\mathbf{x}) - J_{\text{lat}}(\mathbf{x}) \mid$$

In our problem statement the number of HARQ retransmissions is **bounded** since $0 \leq n_{\text{mac}} \leq \nu$ while latency is **finite** because the frame duration is finite, HARQ processes are bounded and scheduling constraints exist.

Therefore, $\mathcal{X}$ is compact (closed and bounded)

If: $J_{\text{retx}}(\mathbf{x})$ and $J_{\text{lat}}(\mathbf{x})$ are continuous (which they are, by construction), then **by the Extreme Value Theorem**, they attain finite maxima, i.e., $M_{\text{retx}} < \infty, M_{\text{lat}} < \infty$. Consequently there are $M_{\text{retx}} \mid \triangleq \mid \sup_{\mathbf{x}\in\mathcal{X}} J_{\text{retx}}(\mathbf{x}), M_{\text{lat}} \mid \triangleq \mid \sup_{\mathbf{x}\in\mathcal{X}} J_{\text{lat}}(\mathbf{x})$ defined as the **maximum possible (worst-case) value** of the retransmission-related cost and the latency-related cost over the feasible set. This means that

$$0 \leq J_{\text{retx}}(\mathbf{x}) \leq M_{\text{retx}}, 0 \leq J_{\text{lat}}(\mathbf{x}) \leq M_{\text{lat}}, \forall \mathbf{x} \in \mathcal{X}.$$

Then, $\mid J_{\text{retx}}(\mathbf{x}) - J_{\text{lat}}(\mathbf{x}) \mid \leq \mid J_{\text{retx}}(\mathbf{x}) \mid + \mid J_{\text{lat}}(\mathbf{x}) \mid \leq M_{\text{retx}} + M_{\text{lat}}$, being the **triangle inequality,** implies that

$$\mid J(\mathbf{x};\lambda_1) - J(\mathbf{x};\lambda_2) \mid = \mid \lambda_1 - \lambda_2 \mid \cdot \mid J_{\text{retx}}(\mathbf{x}) - J_{\text{lat}}(\mathbf{x}) \mid \leq \mid \lambda_1 - \lambda_2 \mid (M_{\text{retx}} + M_{\text{lat}}).$$

Hence, since $\mid J(\mathbf{x};\lambda_1) - J(\mathbf{x};\lambda_2) \mid \leq \mid \lambda_1 - \lambda_2 \mid (M_{\text{retx}} + M_{\text{lat}})$, $J(\mathbf{x};\lambda)$ is **Lipschitz continuous with respect to** λ, uniformly over $\mathcal{X}$, with Lipschitz constant $(M_{\text{retx}} + M_{\text{lat}})$.

In the proposed RBF-based solution, the cost function is approximated as

$$\hat{J}(\mathbf{x};\lambda) = \sum_{k=1}^{K} \alpha_k(\lambda)\, \phi_k(\mathbf{x})$$

where

- $\phi_k(\mathbf{x})$ are fixed radial basis functions,
- $\alpha_k(\lambda)$ are coefficients determined by the scalarized cost in $\hat{J}(\mathbf{x};\lambda)$.

Since $\phi_k(\mathbf{x})$ are independent of λ and the original cost function is linear in λ, the coefficients $\alpha_k(\lambda)$ vary **affinely** with λ. Consequently, $\hat{J}(\mathbf{x};\lambda)$ inherits the Lipschitz continuity of $J(\mathbf{x};\lambda)$.

B.3: Let $\mathbf{x}^*(\lambda) = \arg\min_{\mathbf{x}\in\mathcal{X}} \hat{J}(\mathbf{x};\lambda)$. Since $\mathcal{X}$ is compact and $\hat{J}(\mathbf{x};\lambda)$ is continuous in $\mathbf{x}$ and Lipschitz in λ, standard results from parametric optimization imply that the solution mapping $\mathbf{x}^*(\lambda)$ is **upper hemicontinuous** and varies smoothly with respect to λ. For sufficiently small changes in λ, the deviation in the optimal solution satisfies $\| \mathbf{x}^*(\lambda_1) - \mathbf{x}^*(\lambda_2) \| \leq c \cdot \mid \lambda_1 - \lambda_2 \mid$; for some finite constant $c > 0$, demonstrating that variations in the trade-off parameter λ leads to **bounded and smooth changes** in the optimal solution obtained via the RBF-based method.

Therefore

- Any solution $\mathbf{x}^*(\lambda)$ obtained for $\lambda \in (0,1)$ is Pareto optimal with respect to the original bi-objective formulation;
- Changes in λ correspond to preference shifts rather than structural changes in the optimization problem;

- The proposed framework is robust against moderate variations of λ, and the qualitative performance trends reported in the simulations are preserved across a wide range of trade-off values.

Appendix C

To calculate computational complexity and the overhead costs we will use the $\mathcal{O}(\cdot)$ notation (i.e., the Big-O notation), which characterizes the asymptotic growth rate of computational complexity, the upper bound on computational or communication cost as a function of problem size, while ignoring constant and lower-order terms. In particular, $\mathcal{O}(md)$ represents linear scaling with respect to the number of local training samples m and the model dimensionality d, while constant factors and lower-order terms are omitted. This notation will be used throughout to compare the relative scalability of the proposed RBF-based preference learning approach against conventional SGD-based FEEL methods.

In the subsequent analysis we shall use the following notation:

Table A1. Big-O notation for performance analysis.

Notation	Meaning
O(1)	Constant time (independent of problem size)
O(m)	Linear in number of samples
O(d)	Linear in model dimension
O(md)	Linear in both samples and model size
O(mdT)	Repeated over TTT rounds

Computational Complexity

Baseline: SGD-Based FEEL Optimization

From Section 3 Equations (17) and (18), SGD updates the model parameters as

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \frac{1}{m} \sum_{i=1}^{m} \nabla Q_i(\mathbf{w})$$

The gradient computation per iteration is $\mathcal{O}(md)$ and the total SGD complexity is estimated to be $\mathcal{O}(Tmd)$. In FEEL, this cost is **multiplied by communication rounds**, since gradients must be retransmitted under HARQ.

Proposed RBF-Based Global Optimization

Step 1: Continuous relaxation (negligible cost)

Mapping $n_{\text{mac}} \in \{0, \ldots, \nu\} \rightarrow [0, \nu]$ is constant in time, i.e., $\mathcal{O}(1)$

Step 2: RBF surrogate construction

From (21)–(23), the cost function is approximated using M RBF centers

$$f(n_{\text{mac}}) = \sum_{j=1}^{M} w_j \phi(\| \, n_{\text{mac}} - c_j \, \|)$$

Evaluating one RBF kernel is $\mathcal{O}(1)$(scalar distance). The cost evaluation per device is estimated to be $\mathcal{O}(M)$ and for K devices $\mathcal{O}(KM)$

Step 3: Preference-weighted cost evaluation

From Section 3, Equation (26)

$$C = \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} w_{j,pref} \varphi(\|n_{mac} - c_j\|)$$

Which implies that cost evaluation complexity = $\mathcal{O}(KM)$

Step 4: Softmax preference update

From Section 3, Equation (25)

$$a_j = \frac{e^{\beta p_j}}{\sum_{m=1}^{M} e^{\beta p_m}}$$

Softmax over M actions: $\mathcal{O}(M)$, the preference score update (policy-gradient style) is $\mathcal{O}(1)$ and the total preference update per iteration is $\mathcal{O}(M)$

Step 5: Overall optimization iterations

Let I be the number of preference learning iterations (typically $I \ll T$), then the total complexity is $\mathcal{O}(I(KM + M)) = \mathcal{O}(IKM)$. Typically $M \ll md$ and $I \ll T$, therefore the SGD-based FEEL is worse in computational complexity than RBF approach with preference learning since $\mathcal{O}(IKM) \ll \mathcal{O}(Tmd)$

This proves **strict computational efficiency improvement**

Communication and Overhead Cost Analysis

SGD-Based FEEL Overhead:

In FEEL with HARQ, each device transmits:

- Gradient vector of size d,
- With average retransmissions $\overline{n}_{\text{mac}}$.

Then per round communication cost is $\mathcal{O}(Kd\overline{n}_{\text{mac}})$ and the Total training overhead is $\mathcal{O}(TKd\overline{n}_{\text{mac}})$ which dominates system latency.

Proposed Approach Overhead:

I. Transmitted information

In this case instead of gradients, devices transmit

- Scalar feedback $r_j(t)$,
- ACK/NACK indicators,
- Optional latency metric $\tau_j(t)$.

Per device $\mathcal{O}(1)$, while per iteration $\mathcal{O}(K)$ and the total overhead is $\mathcal{O}(IK)$

II. HARQ Retransmission Reduction

Since n_{mac} is **explicitly optimized**, the expected retransmissions satisfy

$$\mathbb{E}\left[n_{\text{mac}}^{\text{proposed}}\right] \leq \mathbb{E}\left[n_{\text{mac}}^{\text{SGD}}\right]$$

Thus, latency overhead is **provably upper-bounded**.

Concluding, the SGD-based FEEL is worse in overhead cost analysis than RBF approach with preference learning since $\mathcal{O}(IK) \ll \mathcal{O}(TKd\overline{n}_{\text{mac}})$

This proves **strict overhead cost analysis improvement**

Boundary Condition Analysis

Theorem A1. *(SGD–RBF Dominance Boundary in FEEL with HARQ)*

Consider a FEEL system employing HARQ with maximum retransmission limit ν, operating under SINR-dependent BLER, i.e., $BLER(SINR)$. Let SGD-based FEEL minimize the empirical risk via gradient aggregation and let the proposed RBF-based preference learning framework minimize the surrogate cost function as per (26):

$$C = \min_{n_{mac}} \sum_{k=1}^{K} \sum_{j=1}^{M} w_{j,pref} \varphi(\|n_{mac} - c_j\|)$$

Then SGD asymptotically outperforms the proposed RBF-based approach if and only if the following condition holds:

$$\frac{\sigma^2}{\mu T} \cdot \frac{1}{1 - \mathrm{BLER}(\mathrm{SINR})} < \epsilon_{\mathrm{RBF}}^{\mathrm{approx}} + \epsilon_{\mathrm{pref}} + \epsilon_{\mathrm{proj}}$$

where

- σ^2 *is the gradient variance,*
- μ *is the effective curvature (strong convexity constant),*
- T *is the training horizon,*
- ϵ_{RBF}^{approx} *is the RBF approximation error,*
- ϵ_{pref} *is the preference learning error,*
- ϵ_{proj} *is the integer projection error.*

Otherwise, the proposed RBF-based preference learning framework yields lower expected training loss and latency.

Theorem Proof.

1. Fundamental performance models

1.1. SGD-based FEEL convergence under HARQ

For SGD, the expected optimality gap after T rounds, using stochastic optimization theory, satisfy (standard result):

$$\mathbb{E}[Q(w_T) - Q(w^*)] \leq \mathcal{O}\left(\frac{\sigma^2}{\mu T} + \frac{1}{T}\right)$$

where

- σ^2: variance of stochastic gradients
- μ: strong convexity constant (or effective curvature)
- T: number of successful gradient aggregation rounds

Considering HARQ, each effective gradient update requires an average transmission number of

$$\mathbb{E}[n_{\mathrm{mac}}] = \frac{1}{1 - \mathrm{BLER}(\mathrm{SINR})}$$

The effective number of useful rounds becomes $T_{\mathrm{eff}} = \frac{T}{\mathbb{E}[n_{\mathrm{mac}}(\mathrm{SINR})]}$, hence the effective convergence rate degrades to $\mathbb{E}[Q_{\mathrm{SGD}}] \sim \mathcal{O}\left(\frac{\sigma^2 \mathbb{E}[n_{\mathrm{mac}}]}{\mu T}\right)$

1.2. RBF-based FEEL optimization error

Our proposed framework minimizes a **surrogate cost function**:

$$C(n_{\mathrm{mac}}) = \sum_{j=1}^{M} w_{j,\mathrm{pref}} \phi(\parallel n_{\mathrm{mac}} - c_j \parallel)$$

The proposed method does not rely on stochastic gradients; hence the total error is bounded by

$$\epsilon_{\mathrm{RBF}} = \underbrace{\epsilon_{\mathrm{approx}}}_{\text{RBF modeling}} + \underbrace{\epsilon_{\mathrm{pref}}}_{\text{preference learning}} + \underbrace{\epsilon_{\mathrm{proj}}}_{\text{integer projection}}$$

Now let $C(n_{\mathrm{mac}}) \in \mathcal{C}^1([a, b])$ be the true scalar cost function (smooth, bounded first derivative) over a compact 1-D domain of MAC attempts. The RBF approximation is

$$\hat{C}_M(n_{\mathrm{mac}}) = \sum_{j=1}^{M} w_j \, \phi(\parallel n_{\mathrm{mac}} - c_j \parallel)$$

with following assumptions:

- $\phi(\cdot)$ Lipschitz continuous (e.g., Gaussian, multiquadric),
- centers $\{c_j\}$ uniformly spaced over $[a, b]$,
- optimal weights w_j (least-squares or projection).

For functions in $\mathcal{C}^1$ on a compact interval, the best RBF approximation satisfies [46]:

$$\| C - \hat{C}_M \|_\infty | \leq | K h \sup_{x \in [a,b]} | C'(x) |$$

where

- $h = \max_j | c_{j+1} - c_j |$ is the **fill distance**,
- K is a constant depending only on ϕ.

For uniformly spaced centers: $h = \frac{b-a}{M-1} = \mathcal{O}(M^{-1})$, therefore,

$$\varepsilon_{\text{approx}} = \| C - \hat{C}_M \|_\infty = \mathcal{O}(M^{-1})$$

Now the projection error arises because the optimizer selects the nearest discrete action $n_{\text{mac}}^{(j)}$ instead of the true continuous minimizer $n_{\text{mac}}^\star$: $\varepsilon_{\text{proj}} = | C(n_{\text{mac}}^{(j)}) - C(n_{\text{mac}}^\star) |$.

Since $C \in \mathcal{C}^1$, there exists ξ between $n_{\text{mac}}^\star$ and $n_{\text{mac}}^{(j)}$ such that

$$C(n_{\text{mac}}^{(j)}) - C(n_{\text{mac}}^\star) = C'(\xi)\,(n_{\text{mac}}^{(j)} - n_{\text{mac}}^\star)$$

And taking absolute values: $\varepsilon_{\text{proj}} \leq \sup_x | C'(x) | \cdot | n_{\text{mac}}^{(j)} - n_{\text{mac}}^\star |$. Let the action grid spacing be Δn, then by nearest-neighbor selection:

$$| n_{\text{mac}}^{(j)} - n_{\text{mac}}^\star | \leq \frac{\Delta n}{2}$$

Assuming normalized MAC index ($\Delta n = 1$): $| n_{\text{mac}}^{(j)} - n_{\text{mac}}^\star | \leq \frac{1}{2}$, and the final bound becomes

$$\varepsilon_{\text{proj}} \leq \frac{1}{2} \sup_{n_{\text{mac}}} | C'(n_{\text{mac}}) |$$

Finally, $\epsilon_{\text{pref}} \to 0$ as $t \to \infty$ (softmax policy gradient)

1.3. Exact mathematical dominance condition

SGD becomes preferable when $\mathbb{E}[Q_{\text{SGD}}] < \epsilon_{\text{RBF}}$. Substituting, this is the exact crossover condition:

$$\frac{\sigma^2 \mathbb{E}[n_{\text{mac}}(\text{SINR})]}{\mu T} < \epsilon_{\text{approx}} + \epsilon_{\text{pref}} + \epsilon_{\text{proj}}$$

And the theorem is proved.

2. Translating the condition into system-level parameters

2.1. SINR regime

HARQ retransmissions satisfy approximately $\mathbb{E}[n_{\text{mac}}] \approx \frac{1}{1-\text{BLER}(\text{SINR})}$

Thus, in high-SINR regime SINR $\gg$ MCS threshold $\Rightarrow$ BLER $\to 0 \Rightarrow \mathbb{E}[n_{\text{mac}}] \approx 1$

Hence SGD may outperform RBF, because in high-SINR regime:

- Gradients are reliable;
- Retransmissions are rare;
- Communication noise is negligible.

And $\frac{\sigma^2}{\mu T} < \epsilon_{\text{RBF}}$

In Low-to-medium SINR regime BLER increases $\Rightarrow E[n_{mac}] \gg 1$ resulting into

$$\frac{\sigma^2 \mathbb{E}[n_{\text{mac}}]}{\mu T} \gg \epsilon_{\text{RBF}}$$

Thus RBF-based FEEL dominates, because

- SGD variance explodes;
- Gradient distortion accumulates;
- RBF smooths noisy feedback.

2.2. Numerical SINR Thresholds (Concrete Values)

We now quantify the dominance boundary using typical NR UL assumptions (QPSK–64QAM, target BLER = 10%).

The HARQ behavior approximation is

$$\mathbb{E}[n_{\text{mac}}] \approx \frac{1}{1 - \text{BLER}}$$

Table A2. Boundary condition analysis—SINR threshold estimation.

SINR (dB)	BLER	E[nmac]
≥15 dB	≤1%	≈1.01
10 dB	5%	≈1.05
5 dB	15%	≈1.18
0 dB	30%	≈1.43
−5 dB	≥50%	≥2

Let $\epsilon_{\text{RBF}} \approx 10^{-2}$, $\sigma^2/\mu \approx 1$ and $T = 100$ and substitute into

$$\frac{\sigma^2 \mathbb{E}[n_{\text{mac}}(\text{SINR})]}{\mu T} < \epsilon_{\text{approx}} + \epsilon_{\text{pref}} + \epsilon_{\text{proj}}$$

Then SGD dominates when

$$\frac{1}{100} \cdot \mathbb{E}[n_{\text{mac}}] < 10^{-2} \Rightarrow \mathbb{E}[n_{\text{mac}}] < 1$$

This only occurs when BLER $\approx$ **0**, i.e., typical values SINR $\gtrsim$ 15 dB [47], which implies the resulting conclusion that when SINR $\gtrsim$ 15 dB the dominant algorithm is SGD while for any condition SINR $<$ 15 dB the preferred and dominant algorithm is FEEL/RBF

2.3. Trade-off factor λ

Recall that $C(n_{\text{mac}}) = -\lambda G_r + (1 - \lambda) G_T$

SGD becomes favorable when $\lambda \to 1$, meaning

- Reliability dominates,
- Latency cost negligible,
- Retransmissions encouraged.

In this case, SGD can exploit reliable gradients without penalty.

RBF becomes favorable when $\lambda \leq 0.5$, meaning

- Latency is critical,
- Over-retransmission is harmful,
- Explicit optimization of n_{mac} is needed.

2.4 HARQ constraints

SGD becomes better when $\nu \leq 1$, meaning

- Almost no HARQ flexibility,

- Essentially fixed-rate transmission,
- Discrete optimization advantage disappears,

RBF dominates when $\nu \geq 3$, because,

- The optimization space becomes non-trivial,
- SGD cannot reason over-retransmission structure,
- RBF captures global cost trends.

Follows a comparison table for the Boundary Condition Analysis:

Table A3. Boundary condition analysis—comparison Table.

Condition	SGD Better	RBF/Preference Better
SINR	High ($\gtrsim$15 dB)	Low/medium
BLER	$\approx$0	$\geq$10%
Avg. HARQ rounds	$\approx$1	$\geq$2
λ	$\rightarrow$1	$\leq$0.5
HARQ max ν	$\leq$1	$\geq$3
Training Horizon T	Very large	Limited
Latency constraint	Loose	Strict

In the next Figure A1, a dominance region for SGD vs. RBF-based preference learning in FEEL is illustrated, based on our previous boundary conditions analysis:

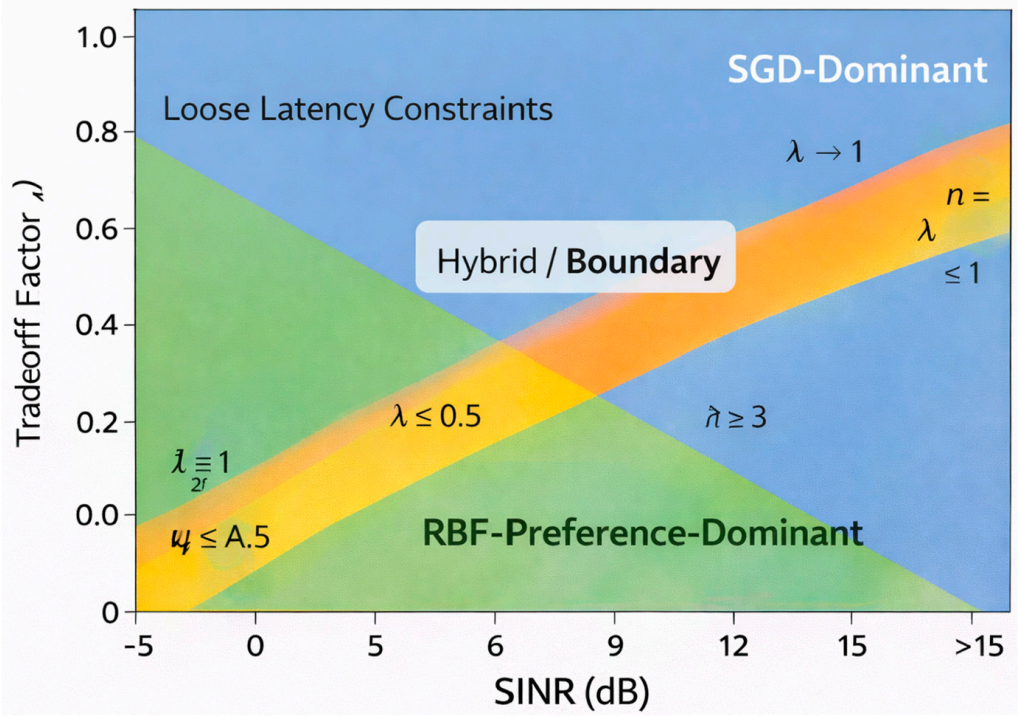


Figure A1. SGD vs. RBF-based preference learning in FEEL dominance regions.

Overall Conclusion: Unlike gradient-based FEEL optimization, which incurs linear dependence on model dimensionality and retransmission count, the proposed framework replaces gradient exchange with scalar preference feedback and global RBF modeling. This results in provably lower computational complexity and communication overhead while preserving optimization accuracy. □

References

1. 3GPP TR 37.817 v17.0.0. Study on Enhancement for Data Collection for NR and EN-DC, Release 17 April 2022. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3817 (accessed on 15 November 2024).
2. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [CrossRef]
3. Aledhari, M.; Razzak, R.; Parizi, R.M.; Saeed, F. Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Access* **2020**, *8*, 140699–140725. [CrossRef]
4. Peter, K.; Brendan, M.H.; Brendan, A.; Aurélien, B.; Mehdi, B.; Nitin, B.A.; Kallista, B.; Zachary, C.; Graham, C.; Rachel, C.; et al. Advances and Open Problems in Federated Learning. *Found. Trends Mach. Learn.* **2021**, *14*, 1–210.
5. 3GPP TR 23.700-84 v0.2.0. Study on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML), Release 19 March 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4252 (accessed on 20 November 2024).
6. 3GPP TR 23.288 v19.1.0. Architecture Enhancements for 5G System (5GS) to Support Network Data Analytics Services, Release 19 December 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3579 (accessed on 6 January 2025).
7. 3GPP TR 23.501 v19.2.1. System Architecture for the 5G System (5GS), Release 19 January 2025. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144 (accessed on 10 February 2025).
8. 3GPP TR 23.502 v19.2.0. Procedures for the 5G System (5GS), Release 19 December 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145 (accessed on 10 February 2025).
9. 3GPP TR 23.503 v19.2.0. Policy and Charging Control Framework for the 5G System (5GS), Release 19 December 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3334 (accessed on 10 February 2025).
10. Qu, M.; Tang, X.; Zhang, Y.; Chen, Z.; Zhang, T.; Zhao, Y.; Li, Y. An Overview of Enabling Artificial Intelligence in 3GPP 5G-Advanced Networks. In Proceedings of the IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Danzhou, China, 17–21 December 2023; pp. 65–70.
11. Lin, X. Artificial Intelligence in 3GPP 5G-Advanced: A Survey. *IEEE ComSoc Technology News*, 6 September 2023.
12. 3GPP TR 22.874 v18.2.0. Study on Traffic Characteristics and Performance Requirements for AI/ML Model Transfer. December 2021. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3721 (accessed on 12 February 2025).
13. 3GPP TS 22.261 v18.10.0. Service Requirements for the 5G System. June 2023. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107 (accessed on 10 February 2025).
14. 3GPP TR 38.843 v18.0.0. Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR Air Interface, Release 18 December 2023. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3983 (accessed on 18 January 2025).
15. El Mokadem, R.; Ben Maissa, Y.; El Akkaoui, Z. eXtreme Federated Learning (XFL): A layer-wise approach. *Clust. Comput.* **2024**, *27*, 5741–5754. [CrossRef]
16. Tao, M.; Zhou, Y.; Shi, Y.; Lu, J.; Cui, S.; Lu, J.; Letaief, K.B. Federated Edge Learning for 6G: Foundations, Methodologies, and Applications. *Proc. IEEE* **2024**, *113*, 1075–1113. [CrossRef]
17. Liu, Y.; Yuan, X.; Xiong, Z.; Kang, J.; Wang, X.; Niyato, D. Federated learning for 6G communications: Challenges, methods, and future directions. *China Commun.* **2020**, *17*, 105–118. [CrossRef]
18. Parra-Ullauri, J.M.; Zhang, X.; Bravalheri, A.; Moazzeni, S.; Wu, Y.; Nejabati, R.; Simeonidou, D. Federated Analytics for 6G Networks: Applications, Challenges, and Opportunities. *IEEE Netw.* **2024**, *38*, 9–17. [CrossRef]
19. Jiao, L.; Shao, Y.; Sun, L.; Liu, F.; Yang, S.; Ma, W.; Li, L.; Liu, X.; Hou, B.; Zhang, X.; et al. Advanced Deep Learning Models for 6G: Overview, Opportunities, and Challenges. *IEEE Access* **2024**, *12*, 133245–133314. [CrossRef]
20. Driss, M.B.; Sabir, E.; Elbiaze, H.; Saad, W. Federated Learning for 6G: Paradigms, Taxonomy, Recent Advances and Insights. *arXiv* **2023**, arXiv:2312.04688. [CrossRef]
21. Yu, C.; Shen, S.; Wang, S.; Zhang, K.; Zhao, H. Efficient Multi-Layer Stochastic Gradient Descent Algorithm for Federated Learning in E-health. In Proceedings of the ICC 2022—IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022; pp. 1263–1268.
22. Gao, Y. Federated learning: Impact of different algorithms and models on prediction results based on fashion-MNIST data set. *Appl. Comput. Eng.* **2024**, *86*, 204–212. [CrossRef]

23. Acar, D.A.E.; Zhao, Y.; Navarro, R.M.; Mattina, M.; Whatmough, P.N.; Saligrama, V. Federated Learning Based on Dynamic Regularization. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual, 3–7 May 2021; pp. 1–36.
24. Yeganeh, Y.; Farshad, A.; Navab, N.; Albarqouni, S. Inverse Distance Aggregation for Federated Learning with Non-IID Data. In Proceedings of the Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning: Second MICCAI Workshop, DART 2020, and First MICCAI Workshop, DCL 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, 4–8 October 2020.
25. Salari, A.; Johnson, S.J.; Vucetic, B.; Shirvanimoghaddan, M. Rate-convergence tradeoff of federated learning over wireless channel. *J. IEEE Internet Things* **2022**, *10*, 22703–22716. [CrossRef]
26. Ye, H.; Liang, L.; Li, G.Y. Decentralized Federated Learning with Unreliable Communications. *IEEE J. Sel. Top. Signal Process.* **2022**, *16*, 487–500. [CrossRef]
27. Jeong, E.; Zecchin, M.; Kountouris, M. Asynchronous decentralized learning over unreliable wireless networks. In Proceedings of the IEEE International Conference on Communications (ICC), Seoul, Republic of Korea, 16–20 May 2022; pp. 607–612.
28. Shah, S.W.H.; Mahboob-Ur-Rahman, M.; Mian, A.N.; Dobre, O.A.; Crowcroft, J. Effective capacity analysis of HARQ-enabled D2D communication in multi-tier cellular networks. *IEEE Trans. Veh. Technol.* **2021**, *70*, 9144–9159. [CrossRef]
29. Campolo, C.; Iera, A.; Molinaro, A. Network for Distributed Intelligence: A Survey and Future Perspectives. *IEEE Access* **2023**, *11*, 52840–52861. [CrossRef]
30. Hu, J.; Wang, F.; Xu, W.; Gao, H.; Zhang, P. SemHARQ: Semantic-Aware HARQ for Multi-task Semantic Communications. *arXiv* **2024**, arXiv:2404.08490.
31. Xu, X.; Liu, S.; Yu, G.D. Adaptive retransmission design for wireless federated edge learning. *ZTE Commun.* **2023**, *21*, 3–14.
32. Amiri, M.M.; Gündüz, D. Federated learning over wireless fading channels. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 3546–3557. [CrossRef]
33. Yu, T.; Huang, P.; Zhang, S.; Chen, X.; Sun, Y.; Wang, X. IREE Oriented Green 6G Networks: A Radial Basis Function-Based Approach. *IEEE J. Sel. Areas Commun.* **2024**, *42*, 3246–3261. [CrossRef]
34. Loh, C.-H.; Chen, Y.-C.; Su, C.-T. Using Transfer Learning and Radial Basis Function Deep Neural Network Feature Extraction to Upgrade Existing Product Fault Detection Systems for Industry 4.0: A Case Study of a Spring Factory. *Appl. Sci.* **2024**, *14*, 2913. [CrossRef]
35. Banholzer, D.; Fliege, J.; Werner, R. A radial basis function method for noisy global optimisation. *Math. Program.* **2025**, *211*, 49–92. [CrossRef]
36. Xi, Y.; Burr, A.; Wei, J.; Grace, D. A General Upper Bound to Evaluate Packet Error Rate over Quasi-Static Fading Channels. *IEEE Trans. Wirel. Commun.* **2011**, *10*, 1373–1377. [CrossRef]
37. Spiros, L.; Paraskevas, M. Analytical average throughput and delay estimations for LTE uplink cell edge users. *Comput. Electr. Eng.* **2014**, *40*, 1552–1563. [CrossRef]
38. 3GPP TS 38.321. Access Network NR; Medium Access Control (MAC) Protocol Specification, Release 18. 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3194 (accessed on 11 January 2026).
39. 3GPP TS 38.211. Access Network NR; Physical Channels and Modulation, Release 18 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213 (accessed on 11 January 2026).
40. 3GPP TS 38.300. Access Network NR and NG-RAN Overall Description. Stage 2, Release 18. 2024. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191 (accessed on 11 January 2026).
41. Louvros, S.; Iossifides, A.C.; Aggelis, K.; Baltagiannis, A.; Economou, G. A Semi-Analytical Macroscopic MAC Layer Model for LTE Uplink. In Proceedings of the 5th International Conference on New Technologies, Mobility and Security (NTMS), Istanbul, Turkey, 7–10 May 2012; pp. 1–5.
42. Göktepe, B.; Hellge, C.; Schierl, T.; Stanczak, S. Distributed Machine-Learning for Early HARQ Feedback Prediction in Cloud RANs. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 31–44. [CrossRef]
43. Stanojev, I.; Simeone, O.; Bar-Ness, Y. Performance Analysis of Collaborative Hybrid-ARQ Incremental Redundancy Protocols Over Fading Channels. In Proceedings of the IEEE 7th Workshop on Signal Processing Advances in Wireless Communications, Cannes, France, 2–5 July 2006; pp. 1–5.
44. McDonald, D.; Grantham, W.; Tabor, W.; Murphy, M. Global and local optimization using radial basis function response surface models. *Appl. Math. Model.* **2007**, *31*, 2095–2110. [CrossRef]
45. Armijo, L. Minimization of functions having Lipschitz continuous first partial derivatives. *Pac. J. Math.* **1966**, *16*, 1–3. [CrossRef]

46. Buhmann, M.D. *Radial Basis Functions: Theory and Implementations*; Cambridge University Press: Cambridge, UK, 2003.
47. Méndez-Monsanto, L.; MacQuarrie, A.; Ghourtani, M.R.; Morales, M.J.L.; Armada, A.G.; Burr, A. BLER-SNR Curves for 5G NR MCS under AWGN Channel with Optimum Quantization. In Proceedings of the 2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall), Washington, DC, USA, 7–10 October 2024; pp. 1–6.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.