A General Framework for Probabilistic Relay Selection in Asymmetric Buffer-Aided Cooperative Relaying Systems

Peng Xu, *Member, IEEE*, Chenghong Luo, Chong Huang, *Member, IEEE*, Gaojie Chen, *Senior Member, IEEE*, Yuanzhi He, Yong Li, *Member, IEEE* and Kai-Kit Wong, *Fellow, IEEE*

***Abstract*—This paper presents a general framework for probabilistic relay selection (RS) in asymmetric buffer-aided cooperative relaying systems, which caters to scenarios with both perfect and imperfect channel state information (CSI) during the RS process. The framework extends and generalizes many existing buffer-aided RS schemes. In particular, we introduce an auxiliary stochastic process which assigns varying selection probabilities to different links, considering the dynamic wireless channel and buffer states. Subsequently, we leverage the obtained outage probability and average packet delay (APD) to formulate outage optimization problems while adhering to APD. To address the intricate high-dimensional optimization problems, we employ a deep learning (DL) approach, which involves designing probability mass functions for the auxiliary stochastic process and developing an effective loss function to update the neural network. Simulation results unequivocally demonstrate the superior performance of the proposed DL-based probabilistic RS scheme compared to benchmark schemes, particularly in scenarios involving imperfect CSI.**

***Index Terms*—Buffer-aided relaying systems, probabilistic relay selection, outage probability, APD, imperfect CSI**

I. Introduction

With the rapid proliferation of new technologies for future networks, the escalating demand for high-speed mobile data traffic has intensified the need for the development of advanced communication technologies. Cooperative relaying, as a promising technology, can enhance the reliability of wireless communication systems [1], and buffer-aided cooperative relaying is widely considered as a solution that can achieve favourable tradeoffs between reliable connectivity and low delay [2], [3]. The buffer-aided relay selection (RS) technology has also been integrated with other emerging technologies, such as physical layer security [4]–[6], non-orthogonal multiple access technology [7]–[9], and free space optical (FSO) communication [10], [11].

Numerous research efforts have focused on designing deterministic buffer-aided RS schemes that rely solely on perfect channel state information (CSI) and buffer state to make selection decisions. One fundamental concept driving the development of deterministic RS schemes is mainly leveraging the instantaneous wireless channel conditions (e.g., [12]–[18]). The max-max RS scheme is an earlier buffer-aided RS scheme based on instantaneous perfect CSI [12], which achieved a diversity order equivalent to the number of relays, denoted as K. In [13], a max link RS scheme based on decode-and-forward (DF) was proposed to realize the full diversity order of $2K$. Then, the authors of [14] proved that the average packet delay (APD) of the max link RS scheme based on amplify-and-forward (AF) is $KL+1$, where L represents the buffer size. Additionally, various modified max-link RS schemes were introduced in [15]–[18], with the primary goal of improving outage or delay performance compared to the traditional max-link scheme. Conversely, in the pursuit of minimizing delays while ensuring high diversity, buffer-state-based (BSB) RS methods have garnered significant attention in buffer-aided cooperative relaying systems (e.g., [19]–[24]). The authors of [19] first demonstrated that a buffer size of 3 is sufficient to achieve the full diversity, and the APD of [19] does not increase rapidly as the buffer size increases. To further reduce the APD and overcome the drawback of increasing APD with an increasing number of relays, a threshold-based buffer-aided RS scheme was proposed in [21], which adjusts K threshold levels to balance the outage probability and APD.

P. Xu and C. Luo are with School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China and also with the Chongqing Key Laboratory of Mobile Communications Technology, Chongqing 400065, China (e-mail: xupeng@cqupt.edu.cn, S210131159@stu.cqupt.edu.cn).

C. Huang is with the Institute for Communication Systems (ICS), 5GIC & 6GIC, University of Surrey, Guildford, Surrey, GU2 7XH, U.K. (e-mail: chong.huang@surrey.ac.uk).

G. Chen is with the School of Flexible Electronics (SoFE) & State Key Laboratory of Optoelectronic Materials and Technologies (OEMT), Sun Yat-sen University, Shenzhen, Guangdong 518107, China and also with the 5GIC & 6GIC, University of Surrey, UK. (e-mail: gaojie.chen@ieee.org).

Y. He is with the Institute of Systems Engineering, Academy of Military Science, PLA, Beijing 100141, China (e-mail: he_yuanzhi@126.com).

Y. Li is with the College of Computer Science, Chongqing University, Chongqing 400044, China (e-mail: yongli@cqu.edu.cn).

K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (Email: kai-kit.wong@ucl.ac.uk).

For each of the aforementioned deterministic RS schemes, the selection outcome is a deterministic function of the buffer and channel states, which may result in unbalanced data flows for both receiving and transmitting at each relay. In contrast, several studies have focused on probabilistic relay or link selection in both single-relay and multi-relay buffer-aided systems [25]–[29]. These probabilistic schemes differ from deterministic approaches by incorporating additional auxiliary random variables that influence the selection outcomes. This allows for improved control over queue lengths at the buffers and enhances the balance of data flows for both source-to-relay (S2R) and relay-to-destination (R2D) links. For instance, in a single-relay scheme, if the S2R link is in poor condition while the R2D link is in good condition, a deterministic RS scheme is likely to frequently select the R2D link, resulting in the buffer tending to be empty. In contrast, using a PRS scheme allows the S2R link to have a non-zero probability of being selected across nearly all buffer states, enabling better control of the queue length by adjusting the selection probabilities. The authors of [27] and [28] proposed max-link-based (MLB) and BSB probabilistic relay selection (PRS) schemes for quasi-symmetric multi-relay systems[1], respectively. The basic idea is to initially identify a candidate from the S2R and R2D links separately. Subsequently, one of the two candidate links is chosen probabilistically. The PRS scheme, introduced in [29], strategically alternated between half-duplex and full-duplex modes using a probabilistic approach.

The RS schemes mentioned earlier operate under the assumption that centralized decision nodes have access to perfect CSI. Nevertheless, achieving perfect CSI in practical wireless systems poses challenges due to the dynamic nature of wireless channels and the presence of estimation errors. Motivated by this consideration, the authors of [30] proposed a single-relay buffer-aided adaptive link selection scheme for scenarios with imperfect CSI. Their numerical results demonstrated that this scheme can yield more significant coding gains when compared to traditional DF RS. In [31], the authors addressed the problem of maximizing throughput in a single-relay system with imperfect CSI. Their research revealed an optimal trade-off between the overhead required for CSI acquisition and the quality of the obtained CSI. Furthermore, in [32], the authors introduced the concepts of S2R broadcasting and distributed decision-making to overcome the challenge of outdated CSI in quasi-symmetric multi-relay systems.

In this paper, we investigate the problem of buffer-aided RS in cooperative relaying systems with K relays and asymmetric channel configurations, where all links are independent but non-identically distributed (i.n.i.d). The decision process takes into account both perfect and imperfect CSI. It is important to note that existing PRS schemes, such as those in [27], [28], were specifically designed for quasi-symmetric systems and focused on allocating selection probabilities for only two candidate links, which may not be suitable for the asymmetric relaying systems under consideration. The MLB and BSB candidate selection processes outlined in [27], [28] could result in significant performance degradation in asymmetric scenarios. To address the requirements of the asymmetric buffer-aided relaying systems, this paper presents a general framework for PRS by integrating an auxiliary stochastic process. This process assigns varying selection probabilities to all the $2K$ links, taking into account the dynamic wireless channel and buffer states. Moreover, unlike previous works on PRS schemes, which have not considered imperfect CSI scenarios (e.g., [27]–[29]), the framework proposed in this paper encompasses PRS schemes in both perfect and imperfect CSI scenarios. It is crucial to acknowledge that the uncertainty associated with the channels introduces complexity to the design of PRS schemes in imperfect CSI scenarios, thereby heightening the challenge of the task.

Furthermore, to evaluate the performance of the proposed framework for PRS, the outage probability and APD are derived based on Markov chains of buffer states. Additionally, outage optimization problems with delay constraints are formulated for both perfect and imperfect CSI. These optimization problems are highly challenging due to the involvement of a large number of time-varying selection probability parameters. Reinforcement learning has been shown to handle high-dimensional and time-varying problems in wireless communications [33], [34]. However, reinforcement learning has drawbacks such as inefficient sample collection and slow convergence rate, which limits its wide use in practical applications. To solve the formulated optimization problems, we use a deep learning (DL) approach, which employs a neural network to design conditional probability mass functions (PMFs) for the auxiliary stochastic process and designs an effective loss function to update the neural network using back propagation algorithm to improve the outage performance with APD constraints. The main contributions of this paper are summarized as follows:

- Proposing a general framework for PRS in the asymmetric buffer-aided cooperative relaying systems, which assigns varying selection probabilities for different links, according to the time-varying wireless channel and buffer states. This framework

[1]In a quasi-symmetric multi-relay system, either the S2R links or R2D links are independent and identically distributed (i.i.d.). However, it is important to note that the S2R links and R2D links have different channel gains.

encompasses PRS schemes in both perfect and imperfect CSI scenarios, and it generalizes many existing deterministic and probabilistic buffer-aided RS schemes.
- Formulating outage optimization problems with APD constraints for both perfect and imperfect CSI, based on the derivations of the outage probability and APD.
- Using a DL approach to tackle high-dimensional optimization problems, which involves designing PMFs for the auxiliary stochastic process conditioned on the channel and buffer states, and developing an effective loss function to update the neural network.
- Providing simulation results to demonstrate the performance improvement of the proposed PRS scheme compared to benchmark schemes. Notably, the advantage of the proposed scheme is more pronounced in scenarios with imperfect CSI.

The surge in emerging applications of mobile devices such as the Internet of Things (IoT), Vehicle-to-Everything (V2X) and Machine-to-Machine (M2M) communications has led to a significant increase in mobile traffic within Beyond Fifth Generation (B5G) and Sixth Generation (6G) networks. To meet the demands for Quality of Service (QoS) with enhanced network throughput, reduced network latency, and heightened network reliability, buffer-aided relaying has emerged as a promising technique. This approach offers increased flexibility in system scheduling of network resources. Designing an effective RS scheme for buffer-aided cooperative networks necessitates the consideration of various dynamic factors such as channel conditions, buffer states, delay performance, selection probabilities, and more. This paper introduces a general framework for PRS and presents a DL approach to address high-dimensional optimization challenges in buffer-aided relaying schemes. This approach shows promise in achieving the desired QoS levels in B5G and 6G networks. For instance, in IoT systems with extensive connections, buffer-aided relaying can pre-store shared information at the decider, facilitating data transmission between remote IoT nodes. By utilizing the proposed buffer-aided DL-PRS scheme, a more attractive high-reliability and low-latency method can be employed for delivering past or real-time information.

In this paper, $[K_1 : K_2]$ denotes the set $\{K_1, K_1+1, \ldots, K_2\}$, where $K_1 \leq K_2$ and both K_1 and K_2 are integers; $\mathbb{1}_\xi$ denotes the indicator function that takes the value 1 if the event ξ is true and 0, otherwise; $\mathbb{P}\{\xi\}$ denotes the probability of the event ξ; $\mathbb{E}\{X\}$ denotes the expectation of the random variable X. Moreover, $p_X(x) = \mathbb{P}\{X = x\}$, $p_{Y|X}(y|x) = \mathbb{P}\{Y = y|X = x\}$ and $\mathbb{E}_X\left[\mathbb{P}\{Y = y|X\}\right] = \sum_{x\in\mathcal{X}} p_X(x) p_{Y|X}(y|x)$.

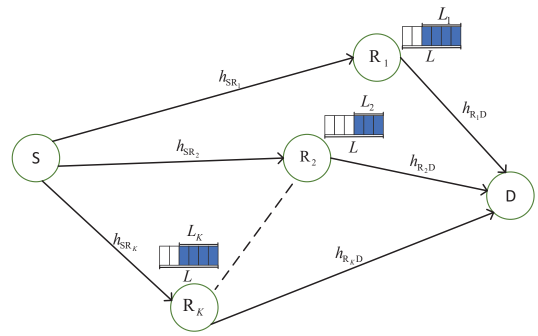


Fig. 1: Asymmetric buffer-aided relaying network.

II. System Model

Consider a buffer-aided RS system, as shown in Fig. 1, which consists of a source node S, K half-duplex[2] DF relays R_k ($k \in [1 : K]$), and a destination node D. Each relay has a buffer of size L to store information packets from S if an S2R link is selected and forward them to D if a R2D link is selected. Due to the path loss and shadowing effects, We assume that there is no direct link between S and D. Therefore, S can only communicate with D through K relays. All links are modelled as independent Rayleigh block fading channels, with channel coefficients remaining constant within a time slot but being independent for different time slots. During each time slot, the number of packets stored in the buffer of relay R_k is denoted by L_k, where $0 \leq L_k \leq L$. Relay R_k can receive a packet if its buffer is not full, i.e., $L_k < L$, and can send a packet if its buffer is not empty, i.e., $L_k > 0$. Consistent with most existing related works (e.g., [13]–[21]), the source node S is assumed to be saturated, ensuring it consistently maintains an infinite data backlog ready for transmission to the relay in each time slot. If all K relays are full or if no S2R links are selected during a particular time slot, S does not transmit any packets, and the queued packets at S will remain pending for transmission in subsequent time slots when the S2R links are selected.

Let $\mathcal{L}_{\mathrm{AB}}$ denote the link from node A to node B, where $\mathrm{A} \in \{\mathrm{S}\}$ and $\mathrm{B} \in \{\mathrm{R}_1, \ldots, \mathrm{R}_K\}$, or $\mathrm{A} \in \{\mathrm{R}_1, \ldots, \mathrm{R}_K\}$ and $\mathrm{B} \in \{\mathrm{D}\}$. If link $\mathcal{L}_{\mathrm{AB}}$ is selected to transmit a packet within a time slot t, the received signal at B is written

[2]This paper assumes that each relay operates under a half-duplex constraint, which provides benefits such as cost reduction and the prevention of self-interference in comparison to full-duplex relays. However, a drawback of employing a half-duplex configuration for the relays is the reduced transmission efficiency, given that each packet transmission consumes two time slots.

as[3]

$$r(t) = \sqrt{P} h_{\mathrm{AB}}(t) x_{\mathrm{A}}(t) + n_{\mathrm{B}}(t), \quad (1)$$

where P is the transmit power of each transmitter, which is a constant, $h_{\mathrm{AB}} = g_{\mathrm{AB}} d_{\mathrm{AB}}^{-\beta/2}$ denotes the channel coefficient of link $\mathcal{L}_{\mathrm{AB}}$, x_{A} is the data signal from A, n_{B} is the additive Gaussian white noise with power σ_n^2, g_{AB} is a complex Gaussian random variable with zero mean and unit variance, d_{AB} denotes the distance of the link $\mathcal{L}_{\mathrm{AB}}$, and β is the path loss exponent. Thus, the instantaneous channel capacity of the link $\mathcal{L}_{\mathrm{AB}}$ can be expressed as

$$C_{\mathrm{AB}} = \log_2 \left(1 + \gamma_{\mathrm{AB}}\right), \quad (2)$$

where $\gamma_{\mathrm{AB}} = \gamma_{\mathrm{tx}} d_{\mathrm{AB}}^{-\beta} |g_{\mathrm{AB}}|^2$, which denotes the received signal-to-noise ratio (SNR) and $\gamma_{\mathrm{tx}} = \frac{P}{\sigma_n^2}$ is the average transmit SNR. Moreover, the target rate is denoted as r_0 in bits per time slot. Since the transmission of each packet consumes two time slots, each packet includes $2r_0$ bits of data. If $C_{\mathrm{AB}} < 2r_0$, the link $\mathcal{L}_{\mathrm{AB}}$ is in outage, i.e., the link $\mathcal{L}_{\mathrm{AB}}$ is in outage if $\gamma_{\mathrm{AB}} < \gamma_{\mathrm{th}}$, where $\gamma_{\mathrm{th}} = 2^{2r_0} - 1$.

In addition, S is assumed to be the centralized decision node responsible for selecting either an S2R or R2D link at the beginning of each time slot. This decision is made based on the CSI and buffer length at each relay. While obtaining the buffer state is relatively straightforward by tracking the transmissions in each time slot, accessing the perfect CSI for each link at the decision node is challenging. This challenge arises from potential estimation errors that may occur during the CSI acquisition process. In this paper, we consider both scenarios: 1) perfect CSI, where the decision node has accurate knowledge of the CSI, and 2) imperfect CSI, where the decision node has to contend with estimation errors in the CSI measurements.

The estimated channel coefficient and the actual channel coefficient of link $\mathcal{L}_{\mathrm{AB}}$ are denoted as $\hat{h}_{\mathrm{AB}}$ and h_{AB}, respectively. In the perfect CSI scenario, the actual channel coefficient and the estimated channel coefficient are equal, i.e. $h_{\mathrm{AB}} = \hat{h}_{\mathrm{AB}}$. In the imperfect CSI scenario, h_{AB} conditioned on $\hat{h}_{\mathrm{AB}}$ follows a Gaussian distribution [35]

$$h_{\mathrm{AB}} \left| \hat{h}_{\mathrm{AB}} \right. \sim \mathcal{CN}\left(\rho_{\mathrm{AB}} \hat{h}_{\mathrm{AB}}, \left(1 - \rho_{\mathrm{AB}}^2\right) \sigma_{\mathrm{AB}}^2\right), \quad (3)$$

where $\rho_{\mathrm{AB}} \in [0, 1)$ denotes the correlation coefficient between the envelopes of h_{AB} and $\hat{h}_{\mathrm{AB}}$, and $\sigma_{\mathrm{AB}}^2 = d_{\mathrm{AB}}^{-\beta}$ denotes the variance of the channel coefficient of link $\mathcal{L}_{\mathrm{AB}}$. The parameter ρ_{AB} is assumed to be the same for all links for the sake of brevity (i.e. $\rho_{\mathrm{AB}} = \rho, \forall \mathrm{AB}$).

[3]For the sake of brevity, the time index t will be removed from any stochastic process for the remainder of this paper, as long as it does not cause any confusion.

The estimated received SNR and the actual received SNR coefficient of link $\mathcal{L}_{\mathrm{AB}}$ are denoted as $\hat{\gamma}_{\mathrm{AB}}$ and γ_{AB}, respectively. In the perfect CSI scenario, the actual received SNR and the estimated received SNR are equal, i.e. $\gamma_{\mathrm{AB}} = \hat{\gamma}_{\mathrm{AB}}$. In the imperfect CSI scenario, γ_{AB} obeys a non-central chi-square distribution with 2 degrees of freedom conditional on the estimated SNR $\hat{\gamma}_{\mathrm{AB}}$, and the expression of its conditional probability density function (PDF) is as follows [36]

$$\begin{aligned} & f_{\gamma_{\mathrm{AB}} | \hat{\gamma}_{\mathrm{AB}}}\left(\gamma_{\mathrm{AB}} \left| \hat{\gamma}_{\mathrm{AB}} \right.\right) = \frac{1}{\bar{\gamma}_{\mathrm{AB}} \left(1 - \rho^2\right)} \times \\ & \exp\left(-\frac{\gamma_{\mathrm{AB}} + \rho^2 \hat{\gamma}_{\mathrm{AB}}}{\bar{\gamma}_{\mathrm{AB}} \left(1 - \rho^2\right)}\right) \times I_0\left(\frac{2\sqrt{\gamma_{\mathrm{AB}} \rho^2 \hat{\gamma}_{\mathrm{AB}}}}{\bar{\gamma}_{\mathrm{AB}} \left(1 - \rho^2\right)}\right), \end{aligned} \quad (4)$$

where $\bar{\gamma}_{\mathrm{AB}} = \gamma_{\mathrm{tx}} d_{\mathrm{AB}}^{-\beta}$ is the average received SNR of link $\mathcal{L}_{\mathrm{AB}}$, and $I_0(\cdot)$ represents the zero-order modified Bessel function of the first kind.

III. General Framework for Probabilistic Relay Selection

A general framework for PRS is summarized in three steps.

1) Determining the available links: An available link refers to an S2R link when its associated buffer is not full, or a R2D link when its associated buffer is not empty. Clearly, only the available links have non-zero probabilities of being selected. Therefore, the set of available links can be determined by the buffer state, which can be expressed as

$$\mathcal{A} = \left\{ \mathcal{L}_{\mathrm{SR}_k} \left| L_k < L, \forall k \right. \right\} \cup \left\{ \mathcal{L}_{\mathrm{R}_k \mathrm{D}} \left| L_k > 0, \forall k \right. \right\}, \quad (5)$$

where L_k denotes the number of stored packets in the buffer of relay R_k in a certain time slot.

2) Determining the qualified links: A qualified link refers to a link for which the estimated received SNR exceeds a certain threshold γ_{th}. Hence, the set of qualified links can be identified based on either perfect or imperfect estimated CSI, which can be expressed as

$$\mathcal{Q} = \left\{ \mathcal{L}_{\mathrm{AB}} \left| \tilde{\gamma}_{\mathrm{AB}} \geq \gamma_{\mathrm{th}}, \forall \mathrm{AB} \right. \right\}, \quad (6)$$

where $\tilde{\gamma}_{\mathrm{AB}} = \gamma_{\mathrm{AB}}$ in the perfect CSI scenario and $\tilde{\gamma}_{\mathrm{AB}} = \hat{\gamma}_{\mathrm{AB}}$ in the imperfect CSI scenario. It is important to note that the set $\mathcal{Q}$ defined in (6) is a stochastic set that is dependent on the time-varying channel state, introduced for ease of explanation. In the scenario with perfect CSI, we can initially ascertain which links are capable of successfully transmitting a packet of $2r_0$ bits. Subsequently, we can formulate the qualified set of links $\mathcal{Q}$. Consequently, no outages are observed for the links within the set $\mathcal{Q}$, while outages are encountered for all links outside of $\mathcal{Q}$. In contrast, in the imperfect CSI scenario, for any link $\mathcal{L}_{\mathrm{AB}} \in \mathcal{Q}$, there still exists a non-zero probability that it may result in an outage.

3) Probabilistically selecting a link: We introduce an auxiliary discrete stochastic process, denoted as Y, with possible integer outcomes from 0 to $2K$, i.e., $Y \in [0:2K]$. In each time slot, the decision maker S utilizes a random number generator to stochastically generate a realization of Y based on the predetermined PMF conditioned on the buffer and channel states. The conditional PMF of Y for each pair of the buffer and channel states will be designed later in Section V. Subsequently, a link is selected according to the following PRS scheme:

$$\mathcal{L}^* = \begin{cases} \mathcal{L}_{\mathrm{SR}_k}, & \text{if } \mathcal{L}_{\mathrm{SR}_k} \in \mathcal{A} \cap \mathcal{Q} \text{ and } Y = k, \\ \mathcal{L}_{\mathrm{R}_k\mathrm{D}}, & \text{if } \mathcal{L}_{\mathrm{R}_k\mathrm{D}} \in \mathcal{A} \cap \mathcal{Q} \text{ and } Y = K + k, \\ \emptyset, & \text{otherwise}, \end{cases} \tag{7}$$

where $k \in [1:K]$ and $\emptyset$ denotes the empty set, i.e., no link is selected.

Remark 1: According to the general framework for PRS, the link selection outcome in each time slot is determined by the instantaneous channel state, the buffer state, and the realization of the introduced auxiliary stochastic process Y. Note that the probabilities of the events $Y = k$ and $Y = K+k$ may be assigned different values, influencing the selection probabilities of the S2R link $\mathcal{L}_{\mathrm{SR}_k}$ and the R2D link $\mathcal{L}_{\mathrm{R}_k\mathrm{D}}$, respectively.

Remark 2: The design of the conditional PMF of Y in each time slot is crucial for the system performance of the proposed framework for PRS. For example, by increasing $\mathbb{P}\{Y \geq K+1\}$, the probability of selecting a R2D link can be increased, and the APD can be decreased. However, The conditional PMF of Y in each time slot needs to adapt to the instantaneous wireless channel conditions and buffer state.

Remark 3: The proposed framework for PRS offers a generalization of several related existing buffer-aided PRS schemes. This can be demonstrated through simplification of the considered asymmetric system into a quasi-symmetric system, along with the assignment of specific values to the conditional PMFs of Y. Such simplification leads to a reduction of the proposed scheme to the existing PRS schemes presented in works such as [27], [28].

Remark 4: For each given pair of the buffer and channel states, if we assign a probability of 1 to a specific realization of Y while assigning a probability of 0 to all other realizations, the proposed framework for PRS simplifies to a series of existing deterministic buffer-aided RS schemes (e.g., [12]–[16], [19]–[21]).

A. Retransmission Mechanism

In the perfect CSI scenario, the selected link $\mathcal{L}^*$ guarantees a reliable transmission without any outage events. Consequently, there is no need for a retransmission mechanism, as the receiver can successfully decode the transmitted information. However, in the imperfect CSI scenario, the transmission over the selected link $\mathcal{L}^*$ does not assure a successful decoding at the receiver. Recognizing this limitation, it becomes necessary to develop a retransmission mechanism.

The retransmission process operates on the principle of the Acknowledgement (ACK)/Negative-Acknowledgement (NACK) mechanism. This mechanism allows the receiver associated with $\mathcal{L}_{\mathrm{AB}}$ to provide feedback to the transmitter over a separate and error-free channel[4] (e.g. [13], [37]). For instance, in the case where the receiver is a relay R_k, if it successfully decodes a packet, the packet is stored in the buffer, and an ACK is sent back to S. On the other hand, if the decoding is unsuccessful, the buffer state remains unchanged, and a NACK is transmitted back to S.

The retransmission mechanism operates instantaneously for each packet transmission within each time slot. Additionally, because a short-length packet is used for each retransmission, only a brief phase is required, which is negligible in comparison to the transmission phase of each data packet. Consequently, the retransmission mechanism has an insignificant impact on the packet delay.

B. The Conditional PMFs of Y

The PMFs of the auxiliary stochastic process Y are conditioned on the buffer and channel states. The number of packets stored in the buffer is defined as the buffer state $\boldsymbol{S}_{\mathrm{B}}$, which can be determined by following equation

$$\boldsymbol{S}_{\mathrm{B}} = [L_1, \ldots, L_K]. \tag{8}$$

In the perfect CSI scenario, our main concern is assessing whether the links are qualified or not. This information is crucial in determining whether these links are experiencing an outage or not. Thus, the channel state vector in the perfect CSI scenario is defined as

$$\boldsymbol{S}_{\mathrm{PCSI}} = \left[\mathbb{1}_{\mathcal{L}_{\mathrm{SR}_1} \in \mathcal{Q}}, \ldots, \mathbb{1}_{\mathcal{L}_{\mathrm{SR}_K} \in \mathcal{Q}}, \mathbb{1}_{\mathcal{L}_{\mathrm{R}_1\mathrm{D}} \in \mathcal{Q}}, \ldots, \mathbb{1}_{\mathcal{L}_{\mathrm{R}_K\mathrm{D}} \in \mathcal{Q}}\right]. \tag{9}$$

In the imperfect CSI scenario, there is a non-zero outage (resp. not outage) probability for link $\mathcal{L}_{\mathrm{AB}}$, even if $\mathbb{1}_{\mathcal{L}_{\mathrm{AB}} \in \mathcal{Q}} = 1$ (resp. $\mathbb{1}_{\mathcal{L}_{\mathrm{AB}} \in \mathcal{Q}} = 0$). In other words, we cannot determine whether a link is in outage. However, the probability that link $\mathcal{L}_{\mathrm{AB}}$ is in outage generally decreases with the value of $\hat{\gamma}_{\mathrm{AB}}$. Thus, we need to consider the continuous value of $\hat{\gamma}_{\mathrm{AB}}$ to design the conditional PMFs of auxiliary stochastic process Y. Thus, the channel state vector in the imperfect CSI scenario is defined as

$$\boldsymbol{S}_{\mathrm{ICSI}} = [\hat{\gamma}_{\mathrm{SR}_1}, \ldots, \hat{\gamma}_{\mathrm{SR}_K}, \hat{\gamma}_{\mathrm{R}_1\mathrm{D}}, \ldots, \hat{\gamma}_{\mathrm{R}_K\mathrm{D}}]. \tag{10}$$

[4]This paper focuses on investigating the performance of the PRS in buffer-aided cooperative relaying systems under perfect/imperfect CSI assumptions. Therefore, implementation issues (CSI acquisition, feedback implementation, etc.) are beyond the scope of this paper.

Let $p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(y\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})$ denote the probability of the event $Y=y$ given that the buffer state is $\boldsymbol{s}_{\mathrm{B}}$ and the channel state is $\boldsymbol{s}_{\Delta}$, where $\Delta\in\{\mathrm{PCSI},\mathrm{ICSI}\}$. Note that the terms "PCSI" and "ICSI" denote perfect CSI and imperfect CSI, respectively. To simplify the constraints in equation (7), suitable constraints are added to the PMFs of Y such that when the event $Y=k$ (resp. $Y=K+k$) occurs, the constraint $\mathcal{L}_{\mathrm{SR}_k}\in\mathcal{A}\cap\mathcal{Q}$ (resp. $\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\in\mathcal{A}\cap\mathcal{Q}$) always holds. Therefore, for a given state pair $(\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})$, $p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(k\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})$ and $p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(K+k\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})$ need to satisfy the following constraints

$$\begin{cases} p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(k\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})\,\mathbb{1}_{\mathcal{L}_{\mathrm{SR}_k}\notin\mathcal{A}\cap\mathcal{Q}}=0,\\ p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(K+k\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})\,\mathbb{1}_{\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\notin\mathcal{A}\cap\mathcal{Q}}=0,\end{cases}\tag{11}$$

$\forall k\in[1:K]$, where $\boldsymbol{s}_{\mathrm{B}}$ and $\boldsymbol{s}_{\Delta}$ are two arbitrary realizations of $\boldsymbol{S}_{\mathrm{B}}$ and $\boldsymbol{S}_{\Delta}$, respectively. Based on (11), it can be inferred that when $\mathcal{A}\cap\mathcal{Q}=\emptyset$, $p_{Y|\boldsymbol{S}_{\mathrm{B}},\boldsymbol{S}_{\Delta}}(0\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta})=1$, and in this case, the value of Y must be 0.

IV. Outage Optimization Problem with APD Constraints Formulation

In this section, we first analyze the performance of the proposed general framework for PRS in terms of outage probability and APD. Then, we formulate the optimization problem of minimizing the outage probability with APD.

A. Outage and Delay Performance

Markov chain is used to analyze the steady-state distribution, outage probability and APD of the proposed scheme. The j-th buffer state of the Markov chain is denoted by

$$\boldsymbol{s}_{\mathrm{B}}^{(j)}=\left[l_1^{(j)},\ldots,l_K^{(j)}\right],\tag{12}$$

where $l_k^{(j)}$ denotes the number of packets stored in the buffer of relay R_k at buffer state $\boldsymbol{s}_{\mathrm{B}}^{(j)}$. Given that $l_k^{(j)}\in[0:L]$, the relationship between the index j and $l_k^{(j)}$ is defined as $j=1+\sum_{k=1}^{K}l_k^{(j)}(L+1)^{k-1}$. Therefore, there are a total of $N_{\mathrm{B}}=(L+1)^K$ states involved in the Markov chain, i.e. $1\le j\le N_{\mathrm{B}}$.

The transition matrix of the Markov chain with $(L+1)^K$ states is represented by $\boldsymbol{F}$. The (i,j)-th element of $\boldsymbol{F}$, denoted by $F_{i,j}$, corresponds to the probability of transitioning from buffer state $\boldsymbol{s}_{\mathrm{B}}^{(j)}$ at time t to $\boldsymbol{s}_{\mathrm{B}}^{(i)}$ at time $t+1$. In other words, $F_{i,j}$ denotes $\mathbb{P}\left\{\boldsymbol{S}_{\mathrm{B}}(t+1)=\boldsymbol{s}_{\mathrm{B}}^{(i)}\,\middle|\,\boldsymbol{S}_{\mathrm{B}}(t)=\boldsymbol{s}_{\mathrm{B}}^{(j)}\right\}$, where $\boldsymbol{S}_{\mathrm{B}}(t)$ denotes the buffer state at time slot t.

The formula for each element of the transition matrix can be expressed by the following equation

$$F_{i,j}=\begin{cases} p_{\mathrm{out}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)}\right), & \text{if } \boldsymbol{s}_{\mathrm{B}}^{(i)}=\boldsymbol{s}_{\mathrm{B}}^{(j)},\\ p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{SR}_k}\right), & \text{if } \boldsymbol{s}_{\mathrm{B}}^{(i)}=\boldsymbol{s}_{\mathrm{B}}^{(j)}+\boldsymbol{I}_k,\\ p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\right), & \text{if } \boldsymbol{s}_{\mathrm{B}}^{(i)}=\boldsymbol{s}_{\mathrm{B}}^{(j)}-\boldsymbol{I}_k,\\ 0, & \text{otherwise,}\end{cases}\tag{13}$$

where $\boldsymbol{I}_k$ denotes the k-th row of the identity matrix of size K; $p_{\mathrm{out}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)}\right)$ denotes the outage probability at buffer state $\boldsymbol{s}_{\mathrm{B}}^{(j)}$; $p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{SR}_k}\right)$ and $p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\right)$ denote the probabilities of selecting links $\mathcal{L}_{\mathrm{SR}_k}$ and $\mathcal{L}_{\mathrm{R}_k\mathrm{D}}$ at the buffer state $\boldsymbol{s}_{\mathrm{B}}^{(j)}$, respectively. Therefore, the transition probabilities of $p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{SR}_k}\right)$ and $p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\right)$ are given by

$$\begin{aligned} &p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{SR}_k}\right)\\ &=\mathbb{P}\left\{\mathcal{L}^*=\mathcal{L}_{\mathrm{SR}_k},\gamma_{\mathrm{SR}_k}\ge\gamma_{\mathrm{th}}\,\middle|\,\boldsymbol{S}_{\mathrm{B}}=\boldsymbol{s}_{\mathrm{B}}^{(j)}\right\}\\ &=\mathbb{P}\left\{Y=k,\gamma_{\mathrm{SR}_k}\ge\gamma_{\mathrm{th}}\,\middle|\,\boldsymbol{S}_{\mathrm{B}}=\boldsymbol{s}_{\mathrm{B}}^{(j)}\right\},\end{aligned}\tag{14}$$

$$\begin{aligned} &p_{\mathrm{sel}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)},\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\right)\\ &=\mathbb{P}\left\{\mathcal{L}^*=\mathcal{L}_{\mathrm{R}_k\mathrm{D}},\gamma_{\mathrm{R}_k\mathrm{D}}\ge\gamma_{\mathrm{th}}\,\middle|\,\boldsymbol{S}_{\mathrm{B}}=\boldsymbol{s}_{\mathrm{B}}^{(j)}\right\}\\ &=\mathbb{P}\left\{Y=K+k,\gamma_{\mathrm{R}_k\mathrm{D}}\ge\gamma_{\mathrm{th}}\,\middle|\,\boldsymbol{S}_{\mathrm{B}}=\boldsymbol{s}_{\mathrm{B}}^{(j)}\right\}.\end{aligned}\tag{15}$$

Let π_j denote the steady-state probability that the buffer state is $\boldsymbol{s}_{\mathrm{B}}^{(j)}$. It can be proven that the Markov chain is column stochastic, irreducible, and aperiodic. Hence, the stationary probability vector can be computed by [13]

$$\boldsymbol{\pi}=(\boldsymbol{F}-\boldsymbol{I}+\boldsymbol{B})^{-1}\boldsymbol{b},\tag{16}$$

where $\boldsymbol{\pi}=[\pi_1,\ldots,\pi_{N_B}]^T$; $\boldsymbol{I}$ denotes the identity matrix of size N_{B}; $\boldsymbol{B}$ denotes the all-ones matrix of size N_{B}; $\boldsymbol{b}$ denotes the all-ones vector of length N_{B}.

The system is considered to be in outage when the relays are unable to receive or forward a packet successfully. By considering all buffer states, the outage probability of the proposed scheme can be expressed as

$$P_{\mathrm{out}}=\sum_{j=1}^{(L+1)^K}\pi_j p_{\mathrm{out}}\left(\boldsymbol{s}_{\mathrm{B}}^{(j)}\right).\tag{17}$$

On the other hand, the packet delay consists of two parts: the delay at the source and the delay at the relay. According to [19], the delay at the source node is determined by the outage probability, while the delay at the relay is jointly determined by the outage probability and the average queue length of data packets. The APD can be calculated as follows [19]

$$\bar{D}=\bar{D}_{\mathrm{S}}+\bar{D}_{\mathrm{R}},\tag{18}$$

where $\bar{D}_{\text{S}} = \frac{1+P_{\text{out}}}{1-P_{\text{out}}}$ and $\bar{D}_{\text{R}} = \frac{2K\bar{L}}{1-P_{\text{out}}}$ denote the delay at the source node and the delay at the relay node, respectively; $\bar{L} = \frac{1}{K}\sum_{j=1}^{N_{\text{B}}}\pi_j\sum_{k=1}^{K} l_k^{(j)}$ represents the average queue length of the data packets for each buffer.

Remark 5: The expression of $\bar{D}$ in (18) is derived primarily based on Little's law [38] and the calculations of the average queue lengths and average throughputs at S and the relays. It is evident that $\bar{D}_{\text{S}}$ and $\bar{D}_{\text{R}}$ have distinct expressions due to the different packet storage and delay measurement mechanisms at S and the relays. Specifically, S is assumed to always have packets ready for transmission, and thus, the queue length at S is influenced by how frequently it is selected for transmission, which, in turn, hinges on the probability of selecting an S2R link. In contrast, the delay of a packet at a relay is defined as the duration between its arrival at the relay and its departure from the relay.

Remark 6: Analyzing explicit expressions regarding the trade-off between the APD and outage probability presents significant challenges due to the complex implicit nature of the stationary probability vector, as shown in (16). To the best of the authors' knowledge, only the authors of [21] and [24] have derived asymptotic explicit expressions for the trade-offs between the APD and diversity order, based on simple selection rules. However, the proposed PRS scheme involves a series of conditional PMFs for Y, which are associated with the buffer and channel states. As a result, deriving the trade-off between the APD constraint and outage probability becomes intractable, presenting an interesting issue for future research.

B. Analysis of the Transition Probabilities

1) Analysis for perfect CSI: In the perfect CSI scenario, Ω_{PCSI} is used to represent the set of possible values for the channel state vector $\boldsymbol{S}_{\text{PCSI}}$ defined in (9). The m-th element in Ω_{PCSI} is denoted by $\boldsymbol{s}_{\text{PCSI}}^{(m)} = \left[\delta_1^{(m)}, \delta_2^{(m)}, \ldots, \delta_{2K}^{(m)}\right]$, where $\delta_\mu^{(m)}$ signifies a realization of the random variable $\mathbb{1}_{\mathcal{L}_{\text{SR}_\mu}\in\mathcal{Q}}$ (for $\mu \in [1:K]$) or $\mathbb{1}_{\mathcal{L}_{\text{R}_{\mu-K}\text{D}}\in\mathcal{Q}}$ (for $\mu \in [K+1:2K]$) at state $\boldsymbol{s}_{\text{PCSI}}^{(m)}$. Given that $\delta_\mu^{(m)} \in \{0,1\}$, the relationship between the index m and $\delta_\mu^{(m)}$ is defined as $m = 1 + \sum_{\mu=1}^{2K} 2^{\mu-1}\delta_\mu^{(m)}$.

Now, let $p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$ denote the value of the transition probability $p_{\text{sel}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$ in the perfect CSI scenario. In the perfect CSI scenario, according to (6) and (11), it can be inferred that $\gamma_{\text{SR}_k} \geq \gamma_{\text{th}}$ if $Y = k$, where $k \in [1:K]$. At the buffer state $\boldsymbol{s}_{\text{B}}^{(j)}$, the event of selecting link $\mathcal{L}_{\text{SR}_k}$ and this link not being in outage is equivalent to the event $Y = k$. Applying the law of total probability at (14), it yields

$$\begin{aligned}
&p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right) \\
&= \mathbb{P}\left\{Y = k \,\middle|\, \boldsymbol{S}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}\right\} \\
&= \sum_{m=1}^{N_{\text{PCSI}}} \mathbb{P}\left\{Y = k \,\middle|\, \boldsymbol{S}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\} \\
&\qquad\quad \times \mathbb{P}\left\{\boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\} \\
&= \sum_{m=1}^{N_{\text{PCSI}}} p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\text{PCSI}}}\left(k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{PCSI}}^{(m)}\right) \mathbb{P}\left\{\boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\},
\end{aligned} \tag{19}$$

where $N_{\text{PCSI}} = 2^{2K}$ represents the total number of elements in $\mathcal{S}_{\text{PCSI}}$; $\boldsymbol{s}_{\text{PCSI}}^{(m)}$ denotes the m-th element in $\mathcal{S}_{\text{PCSI}}$; $p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{s}_{\text{PCSI}}}\left(k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{PCSI}}^{(m)}\right)$ is a conditional PMF of Y; $\mathbb{P}\left\{\boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\}$ can be determined by the following equation

$$\begin{aligned}
\mathbb{P}\left\{\boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\} &= \mathbb{P}\left\{\mathcal{Q} = Q\left(\boldsymbol{s}_{\text{PCSI}}^{(m)}\right)\right\} \\
&= \prod_{\mathcal{L}_{\text{AB}}\in Q\left(\boldsymbol{s}_{\text{PCSI}}^{(m)}\right)} \left(1 - q_{\mathcal{L}_{\text{AB}}}\right) \prod_{\mathcal{L}_{\text{AB}}\notin Q\left(\boldsymbol{s}_{\text{PCSI}}^{(m)}\right)} q_{\mathcal{L}_{\text{AB}}},
\end{aligned} \tag{20}$$

where $Q\left(\boldsymbol{s}_{\text{PCSI}}^{(m)}\right)$ represents the set of qualified links at channel state $\boldsymbol{s}_{\text{PCSI}}^{(m)}$; the random set $\mathcal{Q}$ is defined in (6); $q_{\mathcal{L}_{\text{AB}}}$ represents the outage probability of link $\mathcal{L}_{\text{AB}}$, which can be expressed as

$$\begin{aligned}
q_{\mathcal{L}_{\text{AB}}} &= \mathbb{P}\{\gamma_{\text{AB}} < \gamma_{\text{th}}\} \\
&= \mathbb{P}\left\{\gamma_{\text{tx}} d_{\text{AB}}^{-\beta} |g_{\text{AB}}|^2 < \gamma_{\text{th}}\right\} \\
&= 1 - \exp\left(-\frac{\gamma_{\text{th}}}{\gamma_{\text{tx}} d_{\text{AB}}^{-\beta}}\right).
\end{aligned} \tag{21}$$

Furthermore, let $p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ denote the value of the transition probability $p_{\text{sel}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ in the perfect CSI scenario. By following similar derivation steps of $p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$, $p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ can be obtained as

$$\begin{aligned}
&p_{\text{sel}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right) \\
&= \sum_{m=1}^{N_{\text{PCSI}}} p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\text{PCSI}}}\left(K+k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{PCSI}}^{(m)}\right) \mathbb{P}\left\{\boldsymbol{S}_{\text{PCSI}} = \boldsymbol{s}_{\text{PCSI}}^{(m)}\right\}.
\end{aligned} \tag{22}$$

In addition, let $p_{\text{out}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ denote the value of the transition probability $p_{\text{out}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ in the perfect CSI scenario. In the perfect CSI scenario, an outage event occurs only if $\mathcal{L}^* = \emptyset$. In other words, an outage event occurs only if all available links are in outage. Therefore, $p_{\text{out}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ can be expressed as

$$p_{\text{out}}^{\text{PCSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right) = \prod_{\mathcal{L}_{\text{AB}} \in A\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)} q_{\mathcal{L}_{\text{AB}}}, \tag{23}$$

where $A\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ represents the set of available links at buffer state $\boldsymbol{s}_{\text{B}}^{(j)}$.

2) Analysis for imperfect CSI: Let $p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$ denote the value of the transition probability $p_{\text{sel}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$ in the imperfect CSI scenario. In the imperfect CSI scenario, it exists a non-zero probability that the link $\mathcal{L}_{\text{SR}_k}$ is in outage even if $Y = k$. Conditioned on the imperfect channel state $\boldsymbol{S}_{\text{ICSI}}$, applying the law of total probability at (14) yields

$$\begin{aligned}
&p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right) \\
&= \mathbb{E}_{\boldsymbol{S}_{\text{ICSI}}}\left[\mathbb{P}\left\{Y = k, \gamma_{\text{SR}_k} \geq \gamma_{\text{th}} \,\middle|\, \boldsymbol{S}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{S}_{\text{ICSI}}\right\}\right] \\
&= \mathbb{E}_{\boldsymbol{S}_{\text{ICSI}}}\left[\mathbb{P}\left\{Y = k \,\middle|\, \boldsymbol{S}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{S}_{\text{ICSI}}\right\}\right. \\
&\qquad\qquad \left. \times \mathbb{P}\{\gamma_{\text{SR}_k} \geq \gamma_{\text{th}} \,|\hat{\gamma}_{\text{SR}_k}\}\right] \\
&= \mathbb{E}_{\boldsymbol{S}_{\text{ICSI}}}\left[\mathbb{P}\left\{Y = k \,\middle|\, \boldsymbol{S}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{S}_{\text{ICSI}}\right\}\left(1 - q_{\text{SR}_k|\hat{\gamma}_{\text{SR}_k}}\right)\right],
\end{aligned} \tag{24}$$

where $q_{\text{SR}_k|\hat{\gamma}_{\text{SR}_k}}$ represents the outage probability of link $\mathcal{L}_{\text{SR}_k}$ conditioned on that the estimated SNR is $\hat{\gamma}_{\text{SR}_k}$, which can be calculated as

$$\begin{aligned}
&q_{\text{SR}_k|\hat{\gamma}_{\text{SR}_k}} \\
&= \mathbb{P}\{\gamma_{\text{SR}_k} < \gamma_{\text{th}} \,|\hat{\gamma}_{\text{SR}_k}\} \\
&= \int_0^{\gamma_{\text{th}}} f_{\gamma_{\text{SR}_k}|\hat{\gamma}_{\text{SR}_k}}\left(\gamma_{\text{SR}_k} \,|\hat{\gamma}_{\text{SR}_k}\right) \mathrm{d}\gamma_{\text{SR}_k} \\
&= 1 - Q_1\left(\sqrt{\frac{2\rho^2 \hat{\gamma}_{\text{SR}_k}}{\bar{\gamma}_{\text{SR}_k}(1-\rho^2)}}, \sqrt{\frac{2\gamma_{\text{th}}}{\bar{\gamma}_{\text{SR}_k}(1-\rho^2)}}\right),
\end{aligned} \tag{25}$$

where $Q_1(\cdot)$ stands for the first-order Marcum Q-function.

The closed-form expression for the expectation in (24) is intractable, since it contains the Marcum Q-function. Alternatively, we employ the Monte Carlo estimation method for approximating this complicated computational problem through random sampling. Therefore, applying the Monte Carlo estimation at (24) yields

$$\begin{aligned}
&p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right) \\
&\approx \frac{1}{N_{\text{ICSI}}} \sum_{n=1}^{N_{\text{ICSI}}} \mathbb{P}\left\{Y = k \,\middle|\, \boldsymbol{s}_{\text{B}} = \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{S}_{\text{ICSI}} = \boldsymbol{s}_{\text{ICSI}}^{(n)}\right\} \\
&\qquad\qquad\qquad\qquad \times \left(1 - q_{\text{SR}_k|\hat{\gamma}_{\text{SR}_k}^{(n)}}\right) \\
&= \frac{1}{N_{\text{ICSI}}} \sum_{n=1}^{N_{\text{ICSI}}} p_{Y|\boldsymbol{S}_{\text{B}}, \boldsymbol{s}_{\text{ICSI}}}\left(k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{ICSI}}^{(n)}\right)\left(1 - q_{\text{SR}_k|\hat{\gamma}_{\text{SR}_k}^{(n)}}\right),
\end{aligned} \tag{26}$$

where N_{ICSI} is the total number of samples; $\boldsymbol{s}_{\text{ICSI}}^{(n)}$ denotes the n-th random sample; $\hat{\gamma}_{\text{SR}_k}^{(n)}$ denotes the estimated SNR of link $\mathcal{L}_{\text{SR}_k}$ for the n-th random sample; $p_{Y|\boldsymbol{S}_{\text{B}}, \boldsymbol{s}_{\text{ICSI}}}\left(k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{ICSI}}^{(n)}\right)$ is a conditional PMF of Y.

Furthermore, let $p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ denote the value of the transition probability $p_{\text{sel}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ in the imperfect CSI scenario. By following similar derivation steps of $p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right)$, $p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)$ can be obtained as

$$\begin{aligned}
&p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right) \\
&\approx \frac{1}{N_{\text{ICSI}}} \sum_{n=1}^{N_{\text{ICSI}}} p_{Y|\boldsymbol{S}_{\text{B}}, \boldsymbol{s}_{\text{ICSI}}}\left(K + k \,\middle|\, \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\text{ICSI}}^{(n)}\right) \\
&\qquad\qquad\qquad\qquad \times \left(1 - q_{\text{R}_k\text{D}|\hat{\gamma}_{\text{R}_k\text{D}}^{(n)}}\right).
\end{aligned} \tag{27}$$

In addition, let $p_{\text{out}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ denote the value of the transition probability $p_{\text{out}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ in the imperfect CSI scenario. Since the sum of the elements of the columns of the transition matrix is equal to 1, $p_{\text{out}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ can be given by

$$\begin{aligned}
&p_{\text{out}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right) \\
&= F_{j,j} = 1 - \sum_{i=1, i \neq j}^{N_{\text{B}}} F_{i,j} \\
&= 1 - \sum_{k=1}^{K}\left(p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{SR}_k}\right) + p_{\text{sel}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}, \mathcal{L}_{\text{R}_k\text{D}}\right)\right).
\end{aligned} \tag{28}$$

Remark 7: In the imperfect CSI scenario, $p_{\text{out}}^{\text{ICSI}}\left(\boldsymbol{s}_{\text{B}}^{(j)}\right)$ depends on a finite sample of random channels, and thus the outage probability and APD are only estimated values. Furthermore, it can be concluded that the outage probability is low if the random channel sample has a high average gain, and conversely, it is high if the average gain is low.

C. Problem Formulation

In this paper, we consider a constraint on the APD, and aim to minimize the outage probability in buffer-aided relay systems. The optimization problem can be formulated as

$$\min_{\boldsymbol{G}} \quad P_{\text{out}}\left(\boldsymbol{G}\right) \tag{29}$$

$$\text{s.t.} \quad \bar{D}\left(\boldsymbol{G}\right) \leq \bar{D}^c, \tag{29a}$$

$$\mathbb{1}_{\mathcal{L}_{\text{SR}_k} \notin A\left(\boldsymbol{s}_{\text{B}}^{(j)}\right) \cap Q\left(\boldsymbol{s}_{\Delta}^{(i)}\right)} p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(k \left| \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\Delta}^{(i)}\right.\right) = 0, \tag{29b}$$

$$\mathbb{1}_{\mathcal{L}_{\text{R}_k\text{D}} \notin A\left(\boldsymbol{s}_{\text{B}}^{(j)}\right) \cap Q\left(\boldsymbol{s}_{\Delta}^{(i)}\right)} p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(K+k \left| \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\Delta}^{(i)}\right.\right) = 0, \tag{29c}$$

$$0 \leq p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(y \left| \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\Delta}^{(i)}\right.\right) \leq 1, \forall y \in [0:2K], \tag{29d}$$

$$\sum_{y=0}^{2K} p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(y \left| \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\Delta}^{(i)}\right.\right) = 1, \tag{29e}$$

where $\boldsymbol{G}$ is a three-dimensional matrix of size $N_{\text{B}} \times N_{\Delta} \times (2K+1)$, the (j,i,y)-th element of matrix $\boldsymbol{G}$ is denoted by $G_{j,i,y}$, $G_{j,i,y} = p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(y \left| \boldsymbol{s}_{\text{B}}^{(j)}, \boldsymbol{s}_{\Delta}^{(i)}\right.\right)$ represents the variable to be optimized, referred to as the decision matrix $\boldsymbol{G}$; $P_{\text{out}}\left(\boldsymbol{G}\right)$ and $\bar{D}\left(\boldsymbol{G}\right)$ represent the outage probability and average packet delay, respectively, as functions of the conditional PMFs of Y; $\bar{D}^c$ denotes the APD upper bound. (29a) ensures that the APD does not exceed the set upper bound, (29b) and (29c) ensure that each PMF of Y satisfies equation (11). This paper mainly focuses on solving the complex high-dimensional optimization problem in (29), where the conditional PMFs for Y are associated with the buffer and channel states, which do not have explicit expressions. Consequently, deriving the scaling behaviors for the performance of the proposed RRS scheme is currently a complex and unresolved task, presenting an interesting area for future research.

This problem involves numerous conditional PMFs of the auxiliary stochastic process Y varying with the buffer and channel states, making it a complex high-dimensional optimization problem. The key to solving the optimization problem (29) lies in finding a function that maps the buffer and channel states to the conditional PMFs of Y. A simple approach is to use a table with N_{B} rows and N_{Δ} columns to record the mapping relationship, and then use gradient descent to update the elements in the table. However, when the number of buffer and channel states is large, the table becomes very large, which makes convergence difficult. Furthermore, although the number of channel states in the optimization problem (29) is limited, the number of imperfect channel states in practice is infinite. Therefore, it is necessary to use interpolation to determine the conditional PMFs of Y corresponding to the imperfect channel states that are not in the table, which can lead to issues related to the curse of dimensionality and overfitting. DL can learn the distribution of the channel, use neural networks to accomplish complex nonlinear mapping, and has strong generalization capabilities. Therefore, to overcome the challenges posed by the high-dimensional of this optimization problem and the infinite number of system states in the imperfect CSI scenario, we are motivated to explore the application of DL-based solutions for solving the optimization problem described in (29).

V. Solving the Optimization Problem in (29)

A DL approach is adopted in this section to solve the problem in (29), since it has emerged as a highly promising tool for tackling high-dimensional problems with a bulk of time-varying selection probability parameters. With their parallel computing capabilities and efficient training algorithms, these networks offer faster convergence rates compared to traditional heuristic algorithms. One of the key advantages of neural networks lies in their ability to learn intricate relationships between input and output data, enabling them to approximate functions even with limited data points. By leveraging neural networks, it becomes feasible to overcome the constraints imposed by numerous variables and imperfect CSI. These networks provide a more efficient and accurate approach to solving the optimization problem in (29).

To solve the optimization problem in (29), a DL-based algorithm is designed, involving utilizing a neural network to formulate the conditional PMFs of the auxiliary stochastic process Y and designing an effective loss function for updating the neural network to enhance the outage performance with APD constraints.

A. Design of the Conditional PMFs for Y

Based on whether the intersection between sets $\mathcal{A}$ and $\mathcal{Q}$ is empty, we will discuss the conditional PMFs of auxiliary stochastic process Y separately for the following two cases.

1) The case of $\mathcal{A} \cap \mathcal{Q} = \emptyset$: If $\mathcal{A} \cap \mathcal{Q} = \emptyset$ in a time slot, according to (11), it is known that the value of Y must be 0. Therefore, the conditional PMFs of Y in this case can be determined by the following equation

$$p_{Y|\boldsymbol{S}_{\text{B}},\boldsymbol{S}_{\Delta}}\left(y \left| \boldsymbol{s}_{\text{B}}, \boldsymbol{s}_{\Delta}\right.\right) = \begin{cases} 1, & y = 0, \\ 0, & y \in [1:2K]. \end{cases} \tag{30}$$

2) The case of $\mathcal{A} \cap \mathcal{Q} \neq \emptyset$: If $\mathcal{A} \cap \mathcal{Q} \neq \emptyset$ in a time slot, a neural network and an activation function are employed to determine the conditional PMFs of Y, referred to as the decision network. The system state consisting of buffer and channel states is defined as follows

$$\boldsymbol{s} = \left[\boldsymbol{s}_{\text{B}}, \boldsymbol{s}_{\Delta}\right]. \tag{31}$$

Algorithm 1: DL-based algorithm for solving the optimization problem in (29)

Generate N_{B} buffer states;
Generate $N_{\mathrm{T},\Delta}$ channel states;
Randomly select N_{Δ} channel states from the generated channel states without replacement;
First, construct $N_{\mathrm{B}} \times N_{\Delta}$ system states using all buffer states and the selected N_{Δ} channel states, then initialize the decision matrix using (30) and (32);
repeat
    Randomly select M_{C} channel states from the generated channel states without replacement;
    First, construct $N_{\mathrm{B}} \times M_{\mathrm{C}}$ system states using all buffer states and the selected M_{C} channel states, then update the decision matrix using (30) and (32);
    if *CSI is perfect* **then**
        The transition matrix is computed using (13), (19), (22) and (23);
    else
        The transition matrix is computed using (13), (26), (27) and (28);
    end
    The stationary probability is computed using (16);
    The outage probability and APD are calculated using (17) and (18), respectively;
    The loss function is determined using (34);
    Update the decision network based on (34);
until *the end of learning*;

The decision network used in the DL-based algorithm is a neural network with an input feature size of $3K$ and an output feature size of $2K$. The output of the decision network is transformed into the conditional PMFs of Y using the Softmax activation function. Considering the constraint of the conditional PMFs of Y as specified in (11), not all neurons in the decision network output layer need to be activated.

Therefore, the conditional PMFs of Y in this case can be determined by the following equation

$$p_{Y|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta}}\left(y\,|\boldsymbol{s}_{\mathrm{B}},\boldsymbol{s}_{\Delta}\right)= \begin{cases} \dfrac{\exp\left\{f_{\mathrm{DN}}\left(\boldsymbol{s},k;\boldsymbol{\theta}\right)\right\}}{\xi}, & y=k, \mathcal{L}_{\mathrm{SR}_k}\in\mathcal{A}\cap\mathcal{Q},\\ \dfrac{\exp\left\{f_{\mathrm{DN}}\left(\boldsymbol{s},K+k;\boldsymbol{\theta}\right)\right\}}{\xi}, & y=K+k, \mathcal{L}_{\mathrm{R}_k\mathrm{D}}\in\mathcal{A}\cap\mathcal{Q},\\ 0, & \text{otherwise}, \end{cases} \tag{32}$$

where $f_{\mathrm{DN}}\left(\boldsymbol{s},k;\boldsymbol{\theta}\right)$ and $f_{\mathrm{DN}}\left(\boldsymbol{s},K+k;\boldsymbol{\theta}\right)$ represent the values of the k-th and $(K+k)$-th neurons in the output layer of the decision network, respectively, when the input is $\boldsymbol{s}$; $\boldsymbol{\theta}$ denotes the vector of parameters of the decision network; and ξ is used to ensure that the sum of probabilities is always equal to 1, which can be determined as

$$\xi=\sum_{\mathcal{L}_{\mathrm{SR}_k}\in\mathcal{A}\cap\mathcal{Q}}\exp\left(f_{\mathrm{DN}}\left(\boldsymbol{s},k;\boldsymbol{\theta}\right)\right)+ \sum_{\mathcal{L}_{\mathrm{R}_k\mathrm{D}}\in\mathcal{A}\cap\mathcal{Q}}\exp\left(f_{\mathrm{DN}}\left(\boldsymbol{s},K+k;\boldsymbol{\theta}\right)\right). \tag{33}$$

B. Training of the Decision Network

(30) and (32) transform the variables of the optimization problem in (29) from the decision matrix $\boldsymbol{G}$ to the parameters of the decision network $\boldsymbol{\theta}$. Therefore, the parameters of the decision network can be adjusted to improve the outage performance with APD constraints of the buffer-aided relaying systems. The DL-based algorithm is described by Algorithm 1.

1) Generation of the training data: First, generate N_{B} buffer states according to (8). Then, in the perfect CSI scenario, generate $N_{\mathrm{T,\,PCSI}}$ channel states according to (9), where $N_{\mathrm{T,\,PCSI}}=N_{\mathrm{PCSI}}$; in the imperfect CSI scenario, generate $N_{\mathrm{T,\,ICSI}}$ channel states based on the channel model and (10), where $N_{\mathrm{T,\,ICSI}}>N_{\mathrm{ICSI}}$, ensuring sufficient data for computing transition probabilities in the imperfect CSI scenario.

2) Initialization of the decision matrix: Before updating the decision network, it is necessary to initialize the decision matrix because it is used to compute the transition matrix, which determines the loss function. First, for any buffer state, N_{Δ} channel states are extracted from the generated channel states to compose the system state used for initializing the decision matrix. Then, the decision matrix $\boldsymbol{G}$ is initialized according to formulas (30) and (32).

3) Update of the decision network: We divide the process of updating the decision network into the following sub-steps.

a) Only some of the elements of the decision matrix need to be updated in each iteration, which helps to reduce the consumption of computational resources, and the whole decision matrix can be optimized in several iterations. First, for each buffer state, M_{C} ($M_{\mathrm{C}}\leq N_{\mathrm{T},\Delta}$) channel states are extracted from the training data. Then, according to (30) and (32), the extracted states are used to update the elements of the corresponding decision matrix.
b) In the perfect CSI scenario, the transition matrix is computed using (13), (19), (22) and (23). In the imperfect CSI scenario, the transition matrix is computed using (13), (26), (27) and (28).
c) The stationary probability is computed using (16).

d) The outage probability and APD are calculated using (17) and (18), respectively.

e) APD is optimized if APD does not satisfy the constraints, and the probability of interruption is optimized otherwise. Considering that the outage probability is often a very small value (e.g., 10^{-4}), a logarithm is taken when optimizing the outage probability in this paper. Hence, the loss function is determined as follows

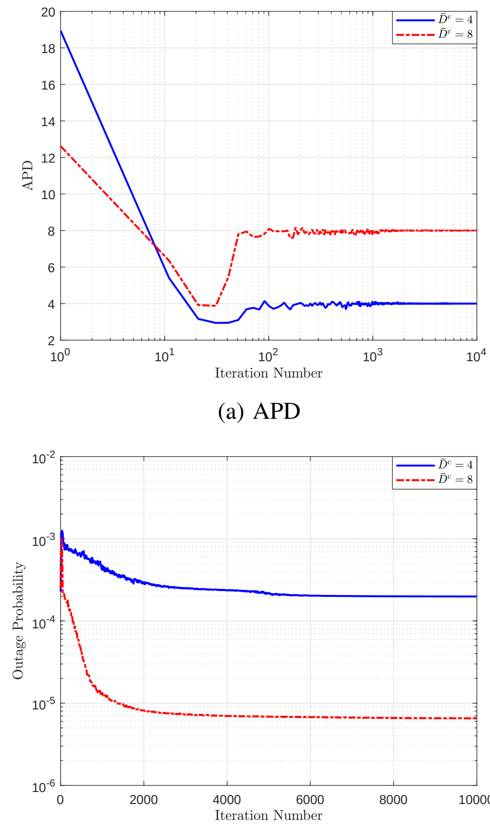
$$\mathcal{O}(\boldsymbol{\theta}) = \begin{cases} \log\left(P_{\text{out}}(\boldsymbol{\theta})\right), & \text{if } \bar{D}(\boldsymbol{\theta}) \leq \bar{D}^c, \\ \bar{D}(\boldsymbol{\theta}), & \text{otherwise.} \end{cases} \quad (34)$$

f) Update the neural network using back propagation based on the loss function (34). In particular, we use the Adam algorithm to update the parameters for better convergence [39], [40].

Remark 8: In the proposed algorithm, we address both perfect and imperfect CSI scenarios by utilizing an auxiliary stochastic process to allocate selection probabilities among various links based on wireless channel conditions and buffer states. The effective loss function in this algorithm leverages back-propagation to establish a mapping between the system environment and the solution. Through iterative training, the deep neural network captures the distribution of environmental variations within the system and outputs favourable choices. This algorithm presents a promising solution for PRS with high reliability and low latency in future cooperative networks.

Remark 9: The challenge of optimal link selection has been addressed in buffer-aided single-relay systems without delay constraints [25], [41]. However, finding optimal solutions for buffer-aided link/relay selection with APD constraints remains a difficult open problem, even for buffer-aided single-relay systems [25], [26]. The optimization problem in (29) involves APD constraints in a buffer-aided multi-relay system, making its optimal solution difficult to obtain. Consequently, it is challenging to theoretically prove that the obtained solution surpasses all deterministic RS schemes. Nevertheless, empirical evidence from simulation results regarding existing probabilistic link/relay selection schemes consistently demonstrates their superior performance over deterministic schemes [25]–[28]. In the next section, we will present simulation results to validate the effectiveness of the proposed PRS scheme.

Remark 10: Based on the experimental results and accumulated experience, we chose a five-layer structure that achieves a good balance between performance and efficiency. Additionally, we generated the training data using a widely used Rayleigh distribution. By taking advantage of the robust fitting capabilities of neural networks, our model can be easily adapted to various practical multipath fading models that utilize different training data distributions. Furthermore, the simulation results presented in the following section demonstrate that the proposed DL-PRS scheme for scenarios with imperfect CSI remains advantageous under rapidly varying channel conditions.



(a) APD

(b) Outage Probability

Fig. 2: APD and outage probability of the proposed DL-PRS scheme versus iteration numbers for the perfect CSI scenario, where $\gamma_{\text{tx}} = 25$ dB.

VI. Simulation Results

In this section, we present simulation results to compare the proposed DL-PRS scheme and three benchmark schemes: the threshold-based scheme in [21], the max-link scheme in [13] and the max-link-based probabilistic (MLBP) scheme in [27]. The threshold-based scheme utilizes K threshold levels to adjust the outage probability and APD[5]. The max-link scheme selects the strongest available link in each time slot, which does not consider

[5]It should be noted that finding optimal threshold levels for the threshold-based scheme suffers a high complexity. In particular, each threshold level is taken from the set $[0 : L-1]$, and hence there are a total of L^K different threshold level combinations. In order to obtain the minimum outage probability with APD constraints, it is necessary to exhaustively enumerate all possible threshold level combinations.

delay constraints. In this section, we simply adjust the buffer size to control the APD of this scheme. The MLBP scheme combines the concepts of the max-link and probabilistic schemes, which control the outage and delay performance by adjusting the selecting probability of the S2R or R2D links. Since the APD is a type of QoS requirement addressed in this paper, as illustrated in (29), this section primarily concentrates on comparing the outage performance between the proposed scheme and existing approaches under the APD constraints. We do not separately compare the APD performance between the proposed scheme and existing approaches.

Several parameters for the considered buffer-aided relay system are set as follows: Unless otherwise stated, the number of relays is $K = 5$, the buffer size is $L = 3$, the path loss exponent is $\beta = 3$, and the target rate is $r_0 = 1$ bits per time slot. Additionally, a two-dimensional asymmetric network topology is considered, where all nodes are located on a two-dimensional plane measured in meters. Specifically, S is located at the point $(0, 0)$, and D is located at $(5, 0)$. The positions of the K relays are as follows: $(3.51, -1.45)$, $(2.51, 1.88)$, $(2.90, -2.12)$, $(1.84, 2.30)$, $(3.08, -1.80)$, $(2.36, -2.30)$ and $(2.02, 1.73)$.[6] Moreover, we set the learning rate to 10^{-3}, the number of sampling updates $M_C = 64$. In the imperfect CSI scenario, the total number of the channel states $N_{T,ICSI} = 10^4$, and the total number of samples N_{ICSI} in (26) and (27) is 1024. In addition, the decision network in this paper is a 5-layer perceptron, where the hidden layer size is 64. Finally, the APD upper bound $\bar{D}_c$ is set to either 4 or 8, unless specified otherwise. The motivation behind this setting is to consider a relatively small or large APD bound. Specifically, as can be observed from (18), when $K = 5$, the APD $\bar{D} \to 1 + 2K\bar{L}$ as $P_{out} \to 0$, indicating that the average queue length $\bar{L}$ for each buffer approaches 0.3 for $\bar{D}_c = 4$ and 0.7 for $\bar{D}_c = 8$. This implies that, on average, each relay needs to be empty during at least 70% of the time slots to meet the APD requirement of $\bar{D}_c = 4$, while each buffer is permitted to store one packet on average during 70% of the time slots for $\bar{D}_c = 8$.

A. Perfect CSI

Fig. 2 shows the APD and outage probability of the proposed scheme versus the iteration numbers under different APD upper bounds for the perfect CSI scenario, where the transmit SNR $\gamma_{tx} = 25$ dB. In Fig. 2(a), it can be observed that when the APD upper bound $\bar{D}^c$ is 4 and 8, the APD converges to 4 and 8, respectively, indicating that the proposed DL-PRS scheme can solve the outage probability optimization problem with APD constraints

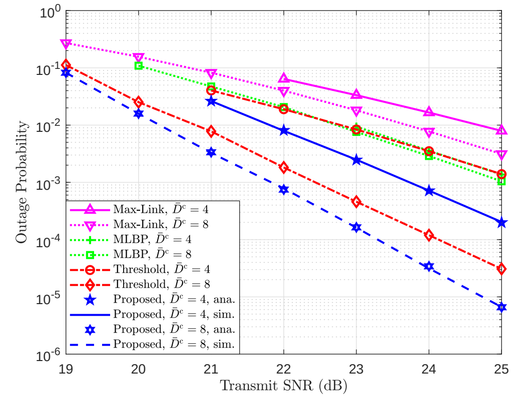


Fig. 3: Outage probability versus transmit SNR for the perfect CSI scenario.

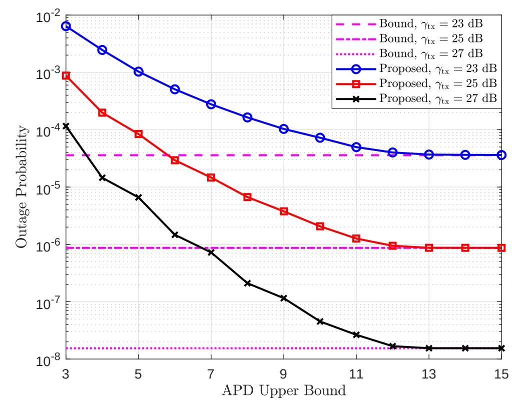


Fig. 4: Outage probability of the proposed DL-PRS scheme versus the APD upper bound $\bar{D}_c$ for the perfect CSI scenario.

in the perfect CSI scenario. Moreover, the APD constraint is satisfied with very few iteration numbers, which is due to the fact that the APD can be reduced by simply selecting R2D links as much as possible. It is worth that the APD continues to descend even if it has satisfied the constraints because Adam is a gradient descent algorithm using momentum, which will consider the historical gradient when updating the neural network. In Fig. 2(b), it can be observed that convergence is achieved around 6000 iterations, with the outage probability converging to 1.98×10^{-4} and 6.54×10^{-6} when the APD upper bound is 4 and 8, respectively. Early in the iteration, even if the APD constraint has been met, there is still huge room for optimization in the outage probability, which is due to the buffer not being fully utilized. Later in the iteration, both the outage probability and APD have small fluctuations, which is due to the outage probability and APD being optimized in turn.

Fig. 3 shows the outage probability of the proposed

[6]If the number of relays K is less than 7, only the first K relay points will be used for their respective positions.

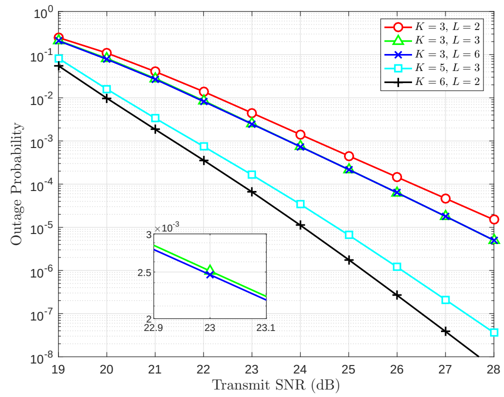
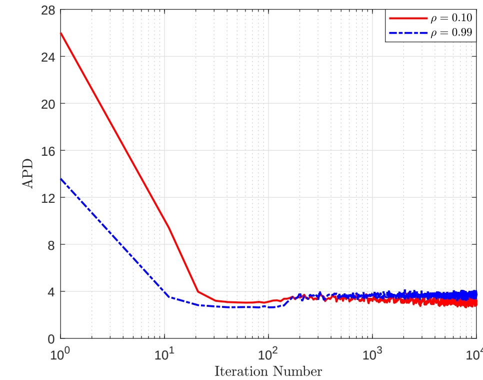


Fig. 5: Outage probability of the proposed DL-PRS scheme versus transmit SNR for the perfect CSI scenario, with $\bar{D}^c = 8$ and different values of K and L.

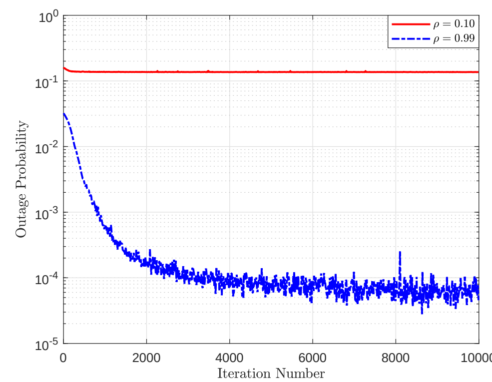
scheme and the benchmark schemes versus the transmit SNR under different APD constraints. Please note that the starting points of the curves in Fig. 3 are not the same. This is because each scheme may require a different minimum SNR to meet the APD constraint. It can be observed that the analytical results (ana.) and Monte Carlo simulation results (sim.) of the proposed scheme match well. It is worth noting that among the three benchmark schemes, the threshold-based scheme exhibits the best performance. However, the outage probability of the proposed DL-PRS scheme is significantly lower than that of the threshold-based scheme, under different delay constraints. For instance, when the transmit SNR is set at 25 dB and the APD constraint is defined as $D^c = 4$, the outage probabilities of the proposed DL-PRS and the threshold-based schemes are 2.03×10^{-4} and 1.39×10^{-3}, respectively.

Fig. 4 illustrates the relationship between the outage probability and the delay constraint D^c for the perfect CSI case, with transmit SNR values of 23, 25, and 27 dB. As the delay constraint increases, the outage probability of the proposed DL-PRS scheme decreases. Additionally, it is worth noting that once the delay constraint surpasses approximately 12, the outage probability of the proposed DL-PRS scheme tends to converge to the selection bound. This implies that for larger delay constraints, the proposed scheme performs very close to the optimal selection bound, which assumes ideal conditions where all buffer queues are always neither full nor empty.

Fig. 5 illustrates the outage performance of the proposed DL-PRS scheme relative to the transmit SNR, with $\bar{D}^c = 8$ and different values of K and L. Observing this figure, it is evident that the outage probability decreases with both K and L. However, it should be noted that the outage performance with $K = 3$ and $L = 6$ only slightly outperforms that with $K = 3$ and $L = 3$. On the other hand, increasing K can notably reduce the outage probability. For instance, substantial performance gains can be observed between the curve with $K = 5$ and $L = 3$ and the one with $K = 6$ and $L = 2$.



(a) APD



(b) Outage Probability

Fig. 6: APD and outage probability of the proposed DL-PRS scheme versus iteration numbers for the imperfect CSI scenario, where $\bar{D}^c = 4$ and $\gamma_{\text{tx}} = 28$ dB.

B. Imperfect CSI

Fig. 6 shows the APD and outage probability of the proposed DL-PRS scheme versus the training iterations under different correlation coefficients for the imperfect CSI scenario, where the APD upper bound $\bar{D}^c = 4$ and the transmit SNR $\gamma_{\text{tx}} = 28$ dB. In Fig. 6(a), a conclusion similar to that in Fig. 2(a) can be observed. The difference is that the APD is significantly smaller than the APD upper bound, which is due to the fact that a stricter APD constraint to ensure that the APD satisfies the constraint even if there is an error in the transmission probability due to the Monte Carlo estimation used in (30) and (32). In addition, we observed fluctuations in the convergence plot of imperfect CSI. This is because different random channel samples are used each time

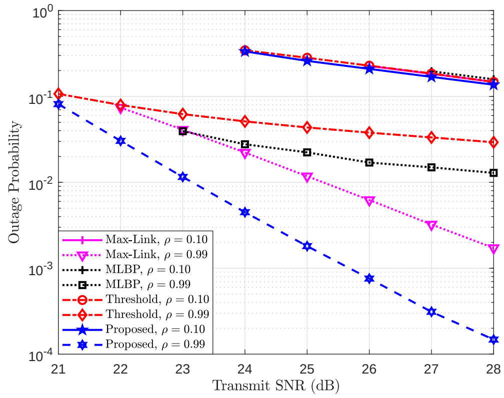


Fig. 7: Outage probability versus transmit SNR for the imperfect CSI scenario, where $\bar{D}^c = 4$.

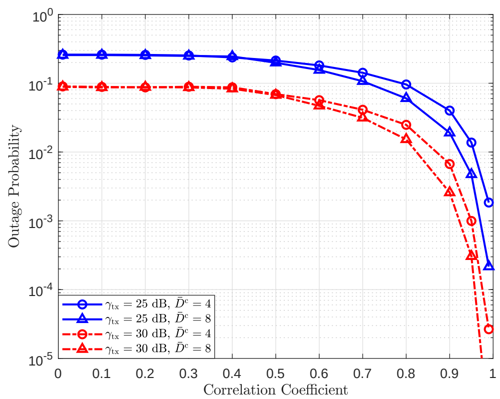
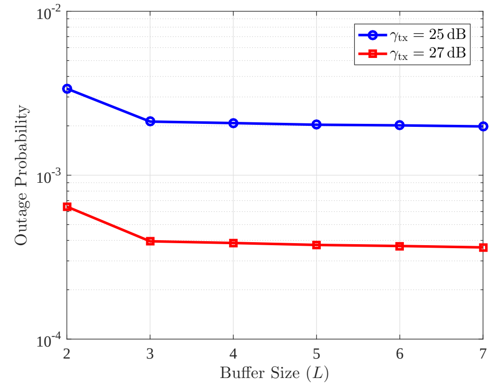
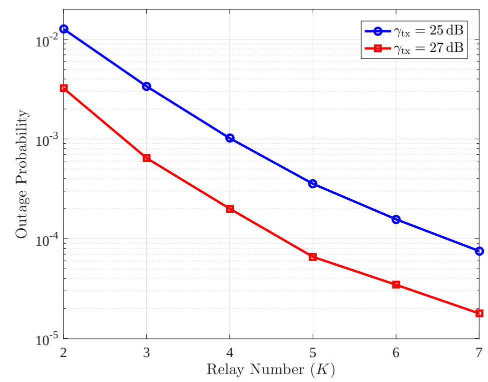


Fig. 8: Outage probability of the proposed DL-PRS scheme versus correlation coefficient ρ.



(a) Outage probability vs L, where $K = 3$



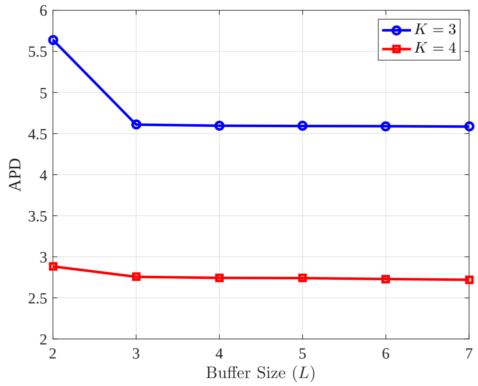
(b) Outage probability vs K, where $L = 2$

Fig. 9: Outage probability versus K and L for the imperfect CSI scenario, where $\rho = 0.99$ and $\bar{D}^c = 8$.

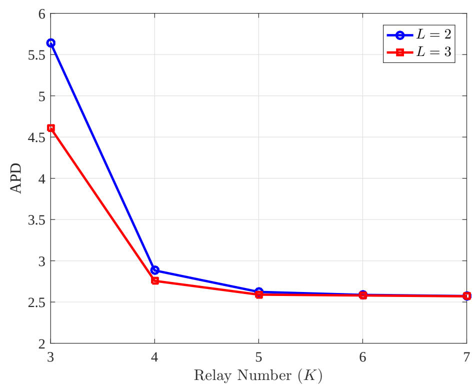
to calculate the outage probability. If the quality of the random channel samples is higher, the outage probability will be lower. Conversely, if the quality of the random channel samples is lower, the outage probability will be higher. Therefore, to ensure the accuracy of numerical results for imperfect CSI, we will use Monte Carlo simulation in the upcoming experiments.

Fig. 7 illustrates the outage performance of the proposed DL-PRS scheme and three benchmark schemes in the imperfect CSI case, where $\bar{D}^c = 4$ and different correlation coefficients are considered. It is worth noting that the threshold-based scheme performs poorly; the max-link scheme outperforms the threshold-based and MBLP schemes. In the scenario of a low correlation coefficient of $\rho = 0.1$, the outage probability of the proposed DL-PRS scheme is similar to that of the other schemes. This can be attributed to the large variance of the channel estimation error, which hampers effective decision-making. However, in the scenario of a high correlation coefficient of $\rho = 0.99$, the proposed DL-PRS scheme obviously exhibits significant performance improvement compared to the existing schemes. For instance, when the transmit SNR is 25 dB, the outage probabilities of the proposed DL-PRS scheme and max-link scheme are 1.49×10^{-4} and 1.72×10^{-3}, respectively.

Fig. 8 illustrates the relationship between the outage probability of the proposed DL-PRS scheme and the correlation coefficient ρ for the imperfect CSI case, considering different transmit SNRs. From this figure, it can be observed that the outage probability generally decreases as the correlation coefficient ρ increases. When ρ is lower than 0.5, all curves exhibit a slow change as ρ increases. However, when ρ approaches 1, the outage probability decreases rapidly with increasing ρ. Therefore, the relationship between the outage probability and the correlation coefficient ρ is complex. When ρ is small, large changes in ρ generally only result in slight changes in the outage probability. However, when ρ is relatively large, even slight changes in ρ can lead to significant changes in the outage probability. This highlights the sensitivity of the system's performance to the correlation



(a) APD vs L



(b) APD vs K

Fig. 10: APD versus K or L under the outage constraint, where $\gamma_{\text{tx}} = 27$ dB, $\rho = 0.99$ and the outage constraint $P_{\text{out}}^{c} = 10^{-3}$.

coefficient in such cases.

Fig. 9 illustrates the outage performance of the proposed DL-PRS scheme in the presence of imperfect CSI, with $\gamma_{\text{tx}} = 25$ or 27 dB, $\bar{D}^{c} = 8$, and $\rho = 0.99$. Fig. 9(a) specifically depicts the outage probability as a function of the buffer size L for $K = 3$. From this figure, it is clear that the outage probability experiences a slight decrease when L exceeds 3. This indicates that simply increasing the buffer size beyond 3 does not lead to significant reductions in outage probability, an observation that has been highlighted in several prior works [19], [21], [28]. Conversely, Fig. 9(b) demonstrates a significant decrease in outage probability as the number of relays K increases from 2 to 7, for $L = 2$. Note that simulating scenarios with larger values of K presents challenges, as the total number of buffer states grows exponentially with K. From Fig. 9, the optimal value for L is determined to be 3, while the appropriate value for K depends on the specific outage requirement.

Fig. 10 illustrates the performance of the proposed DL-PRS scheme in terms of APD as a function of the parameters K and L under an outage constraint, where $\gamma_{\text{tx}} = 27$ dB, $\rho = 0.99$ and the constraint on the outage probability is $P_{\text{out}}^{c} = 10^{-3}$. To generate this figure, the target function in (29) is substituted with $\bar{D}(\boldsymbol{G})$ and the corresponding constraint on the APD outlined in (29a) is modified to $P_{\text{out}}(\boldsymbol{G}) \leq P_{\text{out}}^{c}$. Subsequently, the DL-based algorithm detailed in Section V was slightly adjusted to effectively address this APD minimization problem. In Fig. 10(a), the APD is illustrated versus L, for $K = 3$ or 4. Similar to the trends observed in outage performance shown in Fig. 9, this figure indicates that the APD remains nearly constant as L increases beyond 3. Meanwhile, Fig. 10(b) shows the APD as a function of K, for $L = 2$ or 3. It is notable that the APD similarly stabilizes as K surpasses 5. Thus, based on the parameter configurations in Fig. 10, suitable values for L and K are determined to be 3 and 5, respectively.

VII. Conclusions

This paper presents a general framework for DL-PRS in asymmetric buffer-aided cooperative relaying systems. The framework extends and generalizes numerous existing buffer-aided RS schemes. Both perfect and imperfect CSI cases are considered to make selection decisions. An auxiliary stochastic process is employed to assign selection probabilities to different links based on the wireless channel and buffer states. We formulated outage optimization problems with APD constraints, including a large number of time-varying selection probability parameters. Then, we used a DL approach which involves a neural network to design the conditional PMFs for the auxiliary stochastic process and to develop an effective loss function to update the neural network, aiming to improve the outage performance under APD constraints. Simulation results demonstrated that the proposed DL-PRS scheme significantly outperforms the benchmark schemes, particularly in the imperfect CSI scenario.

References

[1] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, “A simple cooperative diversity method based on network path selection,” *IEEE Journal on selected areas in communications*, vol. 24, no. 3, pp. 659–672, 2006.

[2] N. Zlatanov, A. Ikhlef, T. Islam, and R. Schober, “Buffer-aided cooperative communications: Opportunities and challenges,” *IEEE Communications Magazine*, vol. 52, no. 4, pp. 146–153, 2014.

[3] M. El-Rajab, C. Abou-Rjeily, and R. Kfouri, “Buffer-aided relaying: a survey on relay selection policies,” *IET Communications*, vol. 14, no. 21, pp. 3715–3734, 2020.

[4] R. Nakai and S. Sugiura, “Physical layer security in buffer-state-based max-ratio relay selection exploiting broadcasting with cooperative beamforming and jamming,” *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 431–444, 2018.

[5] J. He, J. Liu, W. Su, Y. Shen, X. Jiang, and N. Shiratori, “Jamming and link selection for joint secrecy/delay guarantees in buffer-aided relay system,” *IEEE Transactions on Communications*, vol. 70, no. 8, pp. 5451–5468, 2022.

[6] G. Srirutchataboon and S. Sugiura, "Physical layer security of buffer-aided hybrid virtual full-duplex and half-duplex relay selection," in *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE, 2022, pp. 1–5.

[7] M. Alkhawatrah, Y. Gong, G. Chen, S. Lambotharan, and J. A. Chambers, "Buffer-aided relay selection for cooperative NOMA in the internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5722–5731, 2019.

[8] J. Ren, X. Lei, and P. T. Mathiopoulos, "Jointly adaptive distributed beamforming and resource allocation for buffer-aided multiple-relay NOMA networks," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7603–7617, 2021.

[9] N. Nomikos, T. Charalambous, D. Vouyioukas, G. K. Karagiannidis, and R. Wichman, "Hybrid NOMA/OMA with buffer-aided relay selection in cooperative networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 524–537, 2019.

[10] A. K. Shukla and M. R. Bhatnagar, "Mitigating jamming in FSO cooperative networks: A buffer-aided relaying approach," *IEEE Communications Letters*, vol. 27, no. 8, pp. 2078–2082, 2023.

[11] C. Abou-Rjeily, "A distance-aware buffer-aided relaying protocol for cooperative FSO communications," *IEEE Wireless Communications Letters*, 2024.

[12] A. Ikhlef, D. S. Michalopoulos, and R. Schober, "Max-max relay selection for relays with buffers," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1124–1135, 2012.

[13] I. Krikidis, T. Charalambous, and J. S. Thompson, "Buffer-aided relay selection for cooperative diversity systems without delay constraints," *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 1957–1967, 2012.

[14] Z. Tian, G. Chen, Y. Gong, Z. Chen, and J. A. Chambers, "Buffer-aided max-link relay selection in amplify-and-forward cooperative networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 2, pp. 553–565, 2015.

[15] M. Oiwa, C. Tosa, and S. Sugiura, "Theoretical analysis of hybrid buffer-aided cooperative protocol based on max-max and max-link relay selections," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 9236–9246, 2016.

[16] B. R. Manoj, R. K. Mallik, and M. R. Bhatnagar, "Performance analysis of buffer-aided priority-based max-link relay selection in DF cooperative networks," *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 2826–2839, 2018.

[17] A. K. Shukla and M. R. Bhatnagar, "Buffer-aided max-link relaying for free space optical communication with delay constraints," *IEEE Photonics Journal*, vol. 14, no. 6, pp. 1–13, 2022.

[18] ——, "Differential modulation-based buffer-aided cooperative relaying network," *IEEE Systems Journal*, vol. 17, no. 2, pp. 2590–2601, 2023.

[19] S. Luo and K. C. Teh, "Buffer state based relay selection for buffer-aided cooperative relaying systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5430–5439, 2015.

[20] A. A. M. Siddig and M. F. M. Salleh, "Balancing buffer-aided relay selection for cooperative relaying systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8276–8290, 2017.

[21] S. El-Zahr and C. Abou-Rjeily, "Threshold based relay selection for buffer-aided cooperative relaying systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 6210–6223, 2021.

[22] ——, "Relaying strategies and asymptotic analysis for half-duplex buffer-aided serial relaying systems," *IEEE Communications Letters*, vol. 26, no. 10, pp. 2327–2331, 2022.

[23] ——, "Buffer state based relay selection for half-duplex buffer-aided serial relaying systems," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 3668–3681, 2022.

[24] P. Xu, J. Quan, G. Chen, Z. Yang, Y. Li, and I. Krikidis, "A novel link selection in coordinated direct and buffer-aided relay transmission," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3296–3309, 2023.

[25] N. Zlatanov and R. Schober, "Buffer-aided relaying with adaptive link selection–Fixed and mixed rate transmission," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2816–2840, 2013.

[26] V. Jamali, N. Zlatanov, and R. Schober, "Bidirectional buffer-aided relay networks with fixed rate transmission–Part II: Delay-constrained case," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1339–1355, 2015.

[27] P. Xu, Z. Yang, Z. Ding, I. Krikidis, and Q. Chen, "A novel probabilistic buffer-aided relay selection scheme in cooperative networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4548–4552, 2020.

[28] P. Xu, G. Chen, Z. Yang, and H. Lei, "Buffer-state-based probabilistic relay selection for cooperative networks with delay constraints," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1855–1859, 2020.

[29] A. A. Siddig, A. S. Ibrahim, and M. H. Ismail, "A low-delay hybrid half/full-duplex link selection scheme for cooperative relaying networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 10, pp. 11 174–11 188, 2022.

[30] T. Islam, D. S. Michalopoulos, R. Schober, and V. K. Bhargava, "Buffer-aided relaying with outdated CSI," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1979–1997, 2015.

[31] V. Jamali, N. Waly, N. Zlatanov, and R. Schober, "Optimal buffer-aided relaying with imperfect CSI," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1309–1312, 2016.

[32] N. Nomikos, T. Charalambous, D. Vouyioukas, and G. K. Karagiannidis, "Low-complexity buffer-aided link selection with outdated CSI and feedback errors," *IEEE Transactions on Communications*, vol. 66, no. 8, pp. 3694–3706, 2018.

[33] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, 2020.

[34] C. Huang, G. Chen, and Y. Gong, "Delay-constrained buffer-aided relay selection in the internet of things with decision-assisted reinforcement learning," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 10 198–10 208, 2021.

[35] J. L. Vicario, A. Bel, J. A. Lopez-Salcedo, and G. Seco, "Opportunistic relay selection with outdated CSI: Outage probability and diversity analysis," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2872–2876, 2009.

[36] J. L. Vicario and C. Anton-Haro, "Analytical assessment of multi-user vs. spatial diversity trade-offs with delayed channel state information," *IEEE Communications Letters*, vol. 10, no. 8, pp. 588–590, 2006.

[37] D. Peng, C. Hu, C. Pan, Y. Dong, P. Fan, and K. B. Letaief, "Timely communications with and without relaying and buffering," *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 24 903–24 918, 2022.

[38] J. D. Little and S. C. Graves, "Little's law," in *International Series in Operations Research & Management Science*. Springer-Verlag, 2008, pp. 81–100.

[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, San Diega, CA, USA, 2015.

[40] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[41] V. Jamali, N. Zlatanov, and R. Schober, "Bidirectional buffer-aided relay networks with fixed rate transmission–Part I: Delay-unconstrained case," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1323–1338, 2015.