Buffer-aided relaying for the multi-user uplink: outage analysis and power allocation

Bo Zhang\(^1\), Chen Dong\(^2\), Jing Lei\(^3\), Mohammed El-Hajjar\(^2\), Lie-Liang Yang\(^2\), Lajos Hanzo\(^2\)

\(^1\)State Key Laboratory of High Performance Computing, School of Computer, National University of Defense Technology, Changsha 410073, People’s Republic of China

\(^2\)School of Electronics and Computer Science, University of Southampton, SO17 1BJ, UK

\(^3\)School of Electronics and Electrical Engineering, National University of Defense Technology, Changsha 410073, People’s Republic of China

E-mail: lh@ecs.soton.ac.uk

Abstract: In this study, the authors consider a two-hop network, where multiple source nodes (SNs) transmit to a destination node (DN) via a relay node (RN). The RN is equipped with a buffer, which is capable of storing multiple frames received from the SNs. During each time slot, the proposed protocol activates either the SN–RN hop or the RN–DN hop, depending on the channel quality of each hop and the buffer state at the RN. To optimise the hop activation for the network, they design a hop quality metric and propose a multi-user buffer-aided-relaying uplink (MU-BR-UL) protocol, with the aid of the minimum signal-to-noise power ratio approximation. The benefits of the proposed protocol are analysed in terms of the end-to-end (e2e) outage probability and the e2e transmission delay. Then, the optimal power allocation is proposed for minimising the e2e outage probability under the total power constraint. The results indicate that the outage performance is significantly improved when the proposed power allocation is utilised in the MU-BR-UL protocol.

1 Introduction

In cooperative wireless networks, a source node (SN) may send its information to the destination node (DN) via a relay node (RN), which improves the link quality and extends the coverage area of the network [1]. In the conventional time-division multiple access (TDMA) based cooperative networks, the RN receives packets from the SN in its receive time slot (TS) and retransmits it to the DN in the transmit TS for the sake of accommodating the half-duplex nature of conventional transceivers, which indicates that the scheduling of the SN’s or the RN’s transmission is fixed [1].

The attainable performance may be improved if the SN–RN or the RN–DN links may be dynamically activated to transmit, based on their near-instantaneous channel quality (CQ). Recently, buffer-aided relaying has received significant attention [2, 3], demonstrating that it is capable of achieving a higher diversity gain than its traditional relaying based counterparts, where the RN is allowed to dynamically opt for receiving or transmitting in each TS, depending on the near-instantaneous CQ of the SN–RN and the RN–DN links. The research efforts on buffer-aided relaying may be divided into two categories: performance analysis and link-activation scheme design.

The performance of buffer-aided relaying has been analysed assuming a buffer of infinite size at the RNs, where Zlatanov et al. [4, 5] demonstrated that a simple two-hop scenario is capable of achieving a diversity order of $D = N$ for the network of $(N-1)$ intermediate RNs [6–9]. On the other hand, the bit error ratio (BER) and the outage performance of buffer-aided multi-hop networks relying on a realistic finite-buffer size were analysed in [6, 7] for transmissions over Rayleigh fading channels having identical average signal-to-noise power ratios (SNRs) per hop. By contrast, Nakagami-$m$ channels having non-identical average SNRs of the various hops were analysed by Dong et al. in [8, 9] and then the analysis was extended to adaptive modulation schemes in [10, 11].

On the other hand, the early contributions of the link-activation design in buffer-aided relaying mainly relied on a CQ of SNR, where the link having the maximum SNR was activated [2]. Recently, Dong et al. [3, 12, 13] proposed the concept of transmission activation probability space (TAPS), where they found that a linear TAPS is sub-optimal and hence proposed novel link-activation protocols relying on a carefully-designed non-linear TAPS. Furthermore, in order to support the practical employment of the link-activation protocols, the authors of [7, 11] proposed MAC layer protocols for buffer-aided multi-hop networks. Most contributions in the open literature considered a single SN, hence there is a paucity of contributions on multi-user buffer-aided relaying scenarios. To fill this gap, recently, buffer-aided relaying conceived for multi-source scenarios started to attract more substantial attention. Islam and Ikhlef [14] considered a multi-SN multi-RN network associated with a common DN, while Li et al. [15] considered a network consisting of multiple SN–DN pairs relying on an RN having a buffer, while attractive best-link selection schemes were proposed for maximising the attainable transmission rate of the resultant adaptive regime by Zlatanov et al. [16]. By contrast, in this paper, we are interested in a multi-source uplink scenario, where M SNs transmit to a common DN with the aid of an RN. We aim for optimising the outage probability of fixed-rate applications, where we propose a novel link-activation protocol and provide the associated outage performance analysis. The novel contributions of this paper can be summarised as follows:

(i) We analyse the outage performance of the multiple access channel (MAC) over the SN–RN hop and propose a minimum SNR (min-SNR) approximation. We formally show that this leads to the asymptotic outage probability expression at high SNRs and we verify this by simulations.

(ii) A hop quality metric (HQM) is designed for the MAC for the network of $(N-1)$ intermediate RNs [6–9]. On the other hand, the bit error ratio (BER) and the outage performance of buffer-aided multi-hop networks relying on a realistic finite-buffer size were analysed in [6, 7] for transmissions over Rayleigh fading channels having identical average signal-to-noise power ratios (SNRs) per hop. By contrast, Nakagami-$m$ channels having non-identical average SNRs of the various hops were analysed by Dong et al. in [8, 9] and then the analysis was extended to adaptive modulation schemes in [10, 11].
regime. We conceive a multi-user buffer-aided-relaying uplink (MU-BR-UL) protocol, which is capable of improving the outage performance of our two-hop system.

(iii) We analyse the end-to-end (e2e) outage performance of the proposed MU-BR-UL protocol. The lower-bound of the e2e outage probability is found by assuming an infinite buffer size at the RN. In addition, the closed-form outage performance expressions are derived for a finite buffer size, which are based on the min-SNR approximation and closely match the actual outage probability. It is shown that for an infinite buffer size, a diversity order of $D=2$ may be achieved. By contrast, for a finite buffer, only a diversity order of $D=1$ is achievable, but nevertheless, a significant outage performance improvement may be attained upon increasing the buffer size. In addition, we analyse the average delay of the proposed system, which is shown to increase linearly with the buffer size.

(iv) Based on the e2e outage probability derived, we formulate a power allocation problem for the sake of minimising the e2e outage probability under the constraint of a given total available transmit power. It is shown that compared to the equal-power philosophy, the proposed power allocation scheme is capable of significantly improving the outage performance. For instance, assuming a network of $M=4$ SNs and an RN equipped with a buffer size of $B=16$ frames as well as a maximum tolerable OP of $P_{out}=0.01$, the proposed power allocation scheme reduces the required transmit power by 5 dB compared to the equal-power allocation.

The paper is organised as follows. Our system model and its benchmark system are described in Section 2. Then, the proposed MU-BR-UL protocol is detailed in Section 3, while both the e2e outage probability and the delay analysis are discussed in Section 4. The optimal power allocation regime conceived for minimising the e2e outage probability is outlined in Section 5. Our simulation results are provided in Section 6, while our conclusions are offered in Section 7.

### 2 System model and benchmark system

We consider a network supporting ($M+2$) nodes, where $M$ SNs $\{S_m, 1 \leq m \leq M\}$ transmit their individual information to a common DN with the aid of a single RN. Each node is equipped with a single antenna and operates in a half-duplex mode. We assume a narrow-band Rayleigh block fading channel model, where the fading coefficients remain constant for the duration of a packet and then they are faded independently from one packet to another both in time and space. The direct links between the SNs and the DN are not considered, since we are interested in the scenario of using relaying for improving the uplink performance of cell-edge users. If the direct link is considered, a diversity order of $D=2$ may be achieved, but this does not affect the hop activation scheme to be proposed in Section 3. The additive noise at the receivers is modelled by independent zero-mean circularly symmetric complex Gaussian random variables.

The receivers are assumed to have perfect channel knowledge and rely on maximum likelihood (ML) detection. The buffer at the RN obeys the first-in-first-out regime, where the frames received first would be transmitted first. Due to the fact that ML detection results in joint detection success or in joint outage events, the packets received from the M SNs are either successfully recovered or corrupted concurrently. Therefore, if the ML detection is successful, all the $M$ detected codewords would be inserted into the RN’s buffer, otherwise all the codewords would be discarded. Hence, we consider a buffer storing $B$ $M$-codeword transmit frames at the RN.

In the conventional two-hop system associated with $B=1$, the schedule of hop activation obeys a fixed pattern. In the first TS, all the SNs concurrently transmit their messages at the rate of $R$, where the RN listens. The SN $S_m$ encodes a bit sequence $b_m$ into a codeword $c_m$ and transmits it to the RN, where the RN jointly decodes the codewords received from all the SNs. Therefore, the SN–RN hop may be modelled by a MAC and the criterion used for successful decoding is that satisfying

$$\sum_{m=0}^{\infty} R \leq \log \left(1 + \sum_{m=0}^{\infty} \gamma_{mB}\right), \quad \forall S \subseteq \{S_m, 1 \leq m \leq M\},$$

where $\gamma_{mB}$ represents the instantaneous received SNR of the $S_m$–RN link. The decoded bit sequences $\{b_0, b_1, \ldots, b_m\}$ would be concatenated and inserted into the buffer as a frame. In the next TS, the RN prepares the frame in its buffer for transmission, re-encodes the messages at the rate of MR and transmits them to the DN, where the criterion to be satisfied for successful decoding is

$$MR \leq \log (1 + \gamma_{dl}),$$

with $\gamma_{dl}$ representing the instantaneous received SNR at the DN.

We define the buffer state $V_b$, which represents that the buffer is storing $b \in [0, B]$ frames. For the buffer-aided system associated with $B=1$, the RN either opts for receiving from the SNs or for transmitting to the DN, according to the instantaneous CQ and to the buffer states $V_b$ in each TS. The proposed hop activation protocol will be described in the following section.

### 3 Buffer-aided two-hop protocol

By introducing a buffer having $B>1$ at the RN, the proposed MU-BR-UL protocol allows the RN to store a maximum of $B$ frames. In each TS, either the SN–RN hop or the RN–DN hop may be activated for transmission, depending both on their channel qualities as well as on the state of the buffer. In a certain TS, if the SN–RN hop is activated and all the $M$ codewords transmitted over the SN–RN hop are recovered at the RN, then they are entered into the buffer as a frame. If the RN–DN hop is activated for transmission, the specific frame at the top of the buffer is removed from the buffer and transmitted to the DN.

The criterion used for activating a hop in each TS should satisfy the following rules:

(i) If the RN’s buffer is empty, the SN–RN hop is activated. This is natural, because the RN has no frames to send over the RN–DN hop.

(ii) If the RN’s buffer is full, the RN–DN hop is activated, otherwise the RN’s buffer overflows.

(iii) The SN–RN and RN–DN hops should be activated with equal probability of 50%.

The first two rules are plausible; hence we would focus our attention on justifying the third one. For example, when the number of users is $M=1$, we always activate the specific hop having the highest instantaneous channel $\text{SNR}$ in order to fully exploit the selective diversity potential of the SN–RN and the RN–DN hops. However, assuming that the average SNR of the SN–RN hop is higher than that of the RN–DN hop, the SN–RN hop would be activated more frequently and therefore the buffer at the RN would overflow. Hence, in order to avoid buffer-overflow, we should activate the SN–RN and RN–DN hops with an equal probability of 50%, relying on an appropriate HQM, where a hop with a higher HQM would be activated. Therefore, the HQM design is crucial for our system.

For the special case of $M=1$ and for an unequal CQ of the different hops, the hop activation problem has been addressed in [9], where the cumulative distribution function (CDF) value of the fading is used as the HQM. However, for the scenario of $M>1$, the hop activation strategy of [9] cannot be used in the multi-user uplink model, because the SN–RN hop is a MAC channel associated with a vector of $M$ instantaneous channel SNR values, while the RN–DN hop is a point-to-point (P2P) channel having a single instantaneous channel SNR value. Hence we are unable to directly estimate and compare the CDF value. Therefore, we design a specific HQM for our...
multiple-user uplink system, where the SN–RN and RN–DN hops are guaranteed to be activated with an equal probability of 50%.

First, let us investigate the OP of the SN–RN hop. If any of the inequalities in (1) is not satisfied, the transmissions over the SN–RN hop are bound to be corrupted. Consider the instantaneous SNRs $\gamma_{mn}$, $m = 1, \ldots, M$, of the channels between the SN $S_m$ and the RN, where these are sorted so that $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_M$ [Note that $\gamma_1$ is the smallest of $\gamma_{mn}$ and it is not necessarily $\gamma_1$]. Then the criterion used for successful decoding in (1) may be expressed as

$$\sum_{i=1}^{M} \gamma_i \geq g_{SR} - 1, \quad 1 \leq m \leq M.$$  \hspace{1cm} (3)

The SNR sorting technique reduces the number of formulas in the decoding criteria from $2^M$ to $M$, as we may find out by comparing (1) and (3). In this way, the outage probability derivation is more tractable, especially for a large $M$. The OP of the SN–RN hop may be derived as follows (see (4))

where $f_{\gamma_i}$ represents the probability density function (PDF) of the specific channel SNR, while the factor $M$ represents the total number of different sorting outcomes for $M$ channel SNRs. In addition, the second line in (4) uses the relationship between the conditional probabilities and the joint probability. Furthermore, note that the SNRs are exponentially distributed, since the channel considered is a Rayleigh fading link.

In order to illustrate (4), an example of the $M = 3$ scenario is presented, where the exact OP of the SN–RN hop may be derived using (4) as follows (see (5))

where each term may be derived separately. Similarly, there are 14 terms for the $M = 4$ scenario. However, as the number of SNs $M$ increases, the closed-form expression of the exact OP expression in (4) becomes excessively complicated due to the large number of integrations.

If any of the inequalities in (1) is not satisfied, the transmissions over the SN–RN hop are bound to be corrupted. On the other hand, we may observe that when the minimum SNR of the $M$ channels spanning from the SNs to the RN, which is defined as $\gamma_{\text{min}} = \min_{m=1}^{M} \gamma_{mn}$, becomes lower than the threshold of $g_{SR} = 2^{M} - 1$ that has to be exceeded for successful decoding, the SN–RN hop becomes erroneous. In Fig. 1, we compare the outage performance of the $M$-user MAC channel employing the theoretical OP characterised in (4) to that of a simulated $M$-user

![Comparison of the outage probability of the first hop in the proposed two-hop system, when using $M = 3$ and 4 users. Theory curve is based on (4), while the min-SNR approximation curve corresponds to the outage of the hop having the min-SNR among the $M$ users. Min-SNR approximation curve and the simulation curve are both based on simulation](image-url)
MAC channel employing ML detection and also to that of a single link having the minimum SNR $\gamma_{\text{min}}$ of the $M$-user system based on simulation. The results of Fig. 1 are shown for the $M=3$ and $M=4$ scenarios. As shown in Fig. 1, the theory and simulation of the $M$-user MAC channel are perfectly matching, while the OP of the ‘minimum SNR approximation’ matches with the theory and simulation for $\text{OP} \leq 10\%$. Therefore, we model the OP performance of the $M$-user MAC in the context of the SN–RN hop based on the performance of the specific SN–RN link having the minimum SNR $\gamma_{\text{min}}$, which facilitates a close approximation of the outage probability.

In the following, we show that the approximate outage probability expression derived in (7) is the asymptotic outage probability

$$P_{e,\text{SK}} = \text{Pr}\left\{ \sum_{m=1}^{M} \gamma_i < 2^{nR} - 1 \right\} \leq \sum_{m=1}^{M} \text{Pr}\left\{ \sum_{i=1}^{m} \gamma_i < 2^{nR} - 1 \right\},$$

where the first term for $m=1$ is the min-SNR approximation and has a diversity order of 1, and we prove that the remaining terms have a higher diversity order. Without loss of generality, we show that the term for $m=2$ has a diversity order higher than 1 in the following (see (8))

$$\text{Pr}\left\{ \gamma_1 + \gamma_2 < 2^{2R} - 1 | \gamma < 2^{R} - 1 \right\},$$

and

$$\text{Pr}\left\{ \gamma_1 + \gamma_2 < 2^{2R} - 1 | \gamma \geq 2^{R} - 1 \right\} \leq \text{Pr}\left\{ \gamma < 2^{R} - 1 \right\}.$$ (9)

Furthermore, we define the HQM of the RN–DN link as the CDF ordinate value of

$$F_{\text{RN}}(\gamma_{\text{DN}}) = 1 - \exp\left( -\frac{\gamma_{\text{DN}}}{\gamma_{\text{RN}}} \right).$$ (12)

In order to satisfy the third rule by activating both the SN–RN and the RN–DN hops with an equal probability [9], when the buffer is neither empty nor full, the SN–RN hop is activated for transmission, provided that the CDF ordinate value of the SN–RN hop is higher than that of the RN–DN hop, which is equivalent to the following condition: $\gamma_{\text{min}}^\text{RN} / \gamma_{\text{min}}^\text{DN} > \gamma_{\text{min}}^\text{RN}$. By contrast, when $\gamma_{\text{min}}^\text{RN} / \gamma_{\text{min}}^\text{DN}$ is satisfied and the buffer is neither full nor empty, then the RN–DN hop is activated for transmission.

Based on the above hop activation strategy, we design our MU-BR-UL protocol in two stages: the hop activation stage and the data transmission stage. In the hop activation stage, the SNs and DN broadcast pilots for channel estimation in orthogonal TSs, while the RN estimates the instantaneous CQ over both the SN–RN and the RN–DN channels [Note that the duration of transmitting pilots can be considered negligible compared to the data transmission duration in the data transmission stage.]. Then, the RN calculates the HQM for both the SN–RN and the RN–DN hops according to (11) and (12), where the hop having a higher HQM would be activated for transmission. Finally, the RN would either transmit a clear-to-send (CTS) signal, if the RN–DN hop is selected or a request-to-send (RTS) signal if the SN–RN hop is selected and then the data transmission will be initiated. It is assumed that the duration of the CTS and RTS signals is negligible, when compared to the length of a TS in the data transmission.

The design of the medium access control layer protocol is discussed here. In our previous work [7, 11], the medium access control protocol for the buffer-aided multi-hop networks is proposed and in this paper, we consider the two-hop case, where the RN can serve as a central control unit. Based on the HQM and the hop activation strategy proposed in this paper, the best hop may be activated as follows. In the first $M$ symbol durations, each SN broadcasts the channel-quality estimation pilot over time-orthogonal channels, whereas in the $(M+1)$th symbol duration, the DN broadcasts the same pilot. The RN may estimate the channel SNR of each the SN–RN and the RN–DN channels and may choose the hop having a higher HQM. Then, during the $(M+2)$th symbol duration, the RN transmits an RTS signal if the first hop is chosen, whereas it sends a CTS signal if it selects the second hop. Data transmission starts from the $(M+3)$th symbol duration. Therefore, the selection process requires $(M + 2)$ symbol durations. Throughout this paper and in the following analysis, it is assumed that the CQ estimation is perfect.

### 4 Performance analysis

The outage event is defined as the event, when any of the codewords received from the SNs is not correctly recovered at the DN. The e2e outage probability of the two-hop system may be expressed as

$$P_{\text{out}} = 1 - (1 - P_{e,\text{SK}})(1 - P_{e,\text{RD}}).$$ (13)

where $P_{\text{out}}$ represents the e2e outage probability of the outage system, while $P_{e,\text{SK}}$ and $P_{e,\text{RD}}$ stand for the error probabilities of the SN–RN and the RN–DN hops, respectively.
4.1 Outage probability lower bound

We first assume that the buffer at the RN has an infinite size and that there are always packets in the buffer ready to be transmitted. In other words, the buffer is assumed to be neither full nor empty. Therefore, the system is always at liberty to choose between the SN–RN and the RN–DN hops to be activated and hence it benefits from a selective diversity gain. In practice, the buffer size is finite and the buffer may be empty, in which case no selective diversity gain is achieved. Therefore, by relying on the assumption mentioned before, we may derive the lower-bound \( P_{\text{out},L} \) of the e2e outage probability, which may be formulated as follows

\[
P_{\text{out},L} = 1 - (1 - P_{\text{select}}) (1 - P_{\text{RD}}),
\]

where \( P_{\text{select}} \) and \( P_{\text{RD}} \) stand for the error probability lower-bounds, when benefiting from a hop selection diversity by appropriately activating the SN–RN and the RN–DN hops [The error performance of the SN–RN and the RN–DN hops achieves the hop activation diversity, when the buffer is neither full nor empty.].

According to our analysis provided in Section 3, the MAC channel of the SN–RN hop may be modelled by a P2P channel having the minimum received SNR \( \gamma_{\text{SR}} \) amongst all the channels spanning from the SNs to the RN. According to (11), we may arrive at the PDF of \( \gamma_{\text{SR}} \) in the form of

\[
f_{\gamma_{\text{SR}}} (\gamma) = \frac{dF_{\gamma_{\text{SR}}} (\gamma)}{d\gamma} = \frac{1}{\gamma_{\text{min}}^{\text{SR}}} \exp \left( -\frac{\gamma}{\gamma_{\text{min}}^{\text{SR}}} \right),
\]

Therefore, we may formulate \( P_{\text{select}} \) using the min-SNR approximation as

\[
P_{\text{select}} \approx \Pr \left( \min \gamma_{\text{SR}} < \gamma_{\text{SR}}^{\text{min}} \right) = \Pr \left( \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} > 1 \right).
\]

\[
P_{\text{select}} \approx 2 \Pr \left( \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} > 1 \right).
\]

\[
P_{\text{select}} \approx 2 \left( \int_0^1 f_{\gamma_{\text{SR}}} (\gamma) dx \right) f_{\gamma_{\text{SR}}} (\gamma) \text{d}\gamma.
\]

\[
P_{\text{select}} \approx 1 + \exp \left( -2 \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} \right) - 2 \exp \left( -\frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} \right).
\]

where we exploit the relation of \( \Pr \left( \min \gamma_{\text{SR}} < \gamma_{\text{SR}}^{\text{min}} \right) \approx \Pr \left( \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} \leq 1 \right) \).

Similarly, we may arrive at

\[
P_{\text{RD}} = \Pr \left( \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} < 1 \right) = \exp \left( -2 \frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} \right) - 2 \exp \left( -\frac{\gamma_{\text{SR}}^{\text{min}}}{\gamma_{\text{SR}}} \right).
\]

4.2 Exact outage probability analysis

Let us now remove the assumption of having a buffer, which is neither full nor empty as stipulated in the previous section and take into account the scenarios, when the buffer is either empty or full. In these cases, according to the hop activation rules, we have to activate the SN–RN or the RN–DN hops without the benefit of choice and therefore no selective diversity gain is achieved. We define the buffer size as \( B \) frames, while \( b \) is the number of frames stored in the buffer.

According to our hop activation strategy, we can derive the approximate error probability of \( P_{e,RD} \) using the min-SNR approximation and the actual \( P_{e,RD} \) as follows

\[
P_{e,RD} \approx \Pr (b = 0) P_{e,RD|b=0} + \Pr (0 < b < B) P_{e,RD|b=B}.
\]

\[
P_{e,RD} \approx \Pr (b = 0) P_{e,RD|b=0} + \Pr (0 < b < B) P_{e,RD|b=B}.
\]

where \( P_{e,RD} \) stands for the error probability of the SN–RN hop, when the buffer is empty, i.e. \( b = 0 \) and this can be expressed as

\[
P_{e,RD|b=0} \approx 1 - \exp \left( -\frac{s_{\text{SR}}}{\gamma_{\text{SR}}} \right).
\]

On the other hand, \( P_{e,RD|b=B} \) stands for the error probability of the RN–DN hop, when the buffer is full, i.e. \( b = B \) and we can express \( P_{e,RD|b=B} \) as

\[
P_{e,RD|b=B} \approx 1 - \exp \left( -\frac{s_{\text{SR}}}{\gamma_{\text{SR}}} \right).
\]

In order to derive \( P_{e,RD|b} \), we have to derive the probability of \( \Pr (b = 0) \) as well as of \( \Pr (b = B) \) and hence we adopt the state-transition analytical tools proposed in [7]. We first define the state \( V_j \) as the specific state, in which the number of frames stored in the buffer is \( b \), \( 0 \leq b \leq B \). Then we arrive at the state transition matrix \( T \), where the specific element \( T_{ij} \) in the \( i \)th row and the \( j \)th column stands for the transition probability from the state \( V_i \) to \( V_j \). For example, the transition matrix \( T \) of the \( B = 3 \) scenario is

\[
T = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0.5 & 0.5 & 0 & 0.5 \\
0 & 0.5 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}.
\]

When the buffer size \( B \) increases, the dimension of \( T \) is as large as \( (B + 1) \times (B + 1) \), but it is easy to construct:

(i) For the first row, the current state \( V_0 \) stands for the empty buffer. It evolves to state \( V_1 \) with a probability of 1. 
(ii) For the \( (B + 1) \)th row, the current state \( V_B \) stands for the full buffer. It has to evolve to \( V_{B+1} \) with a probability of 1, because the RN–DN hop is activated and the first frame of the buffer would be removed and sent to the DN.
(iii) For the rest of the rows, the current state \( V_j \) represents neither empty nor full buffer. When the RN–DN hop is selected, \( V_j \) traverses to \( V_{j+1} \) with a probability of 0.5. When the SN–RN hop is activated, \( V_j \) traverses to \( V_{j-1} \) with a probability of 0.5.

When the buffer is in steady state, the effects of the system represented by a multiplication operation of the system transfer matrix \( T \) no longer changes the system state \( \pi \). Therefore, the state probabilities \( \pi \) may be computed using [7]

\[
\pi = T^\pi \pi_0.
\]
approximation as well as the $P_{e,\text{RD}}$ formulated in (18) as follows

\[
P_{e,\text{SR}} \simeq 1 - \frac{2B - 1}{B} \exp \left( -\frac{\gamma_0^u}{\gamma_{\text{min}}^u} \right) + \frac{B - 1}{B} \exp \left( -\frac{2\gamma_0^u}{\gamma_{\text{rd}}^u} \right),
\]

\[
P_{e,\text{RD}} = 1 - \frac{2B - 1}{B} \exp \left( -\frac{\gamma_0^u}{\gamma_{\text{rd}}^u} \right) + \frac{B - 1}{B} \exp \left( -\frac{2\gamma_0^u}{\gamma_{\text{rd}}^u} \right),
\]

which can be substituted into (13) in order to arrive at the exact OP for an arbitrary buffer size $B$ as follows:

\[
P_{\text{out}} \simeq 1 - \left[ \frac{2B - 1}{B} \exp \left( -\frac{\gamma_0^u}{\gamma_{\text{min}}^u} \right) - \frac{B - 1}{B} \exp \left( -\frac{2\gamma_0^u}{\gamma_{\text{rd}}^u} \right) \right] \times \left[ \frac{2B - 1}{B} \exp \left( -\frac{\gamma_0^u}{\gamma_{\text{rd}}^u} \right) - \frac{B - 1}{B} \exp \left( -\frac{2\gamma_0^u}{\gamma_{\text{rd}}^u} \right) \right].
\]

### 4.3 Diversity analysis

The expression of the exact OP in (24) is somewhat complex, while further insights may be offered by its high-SNR approximations. First, using the Taylor series expansion, we may formulate the per-hop exact error probability of $P_{e,\text{SR}}$ and $P_{e,\text{RD}}$, as follows

\[
P_{e,\text{SR}} \simeq 1 - \frac{1}{B} \frac{\gamma_0^u}{\gamma_{\text{min}}^u} + \frac{2B - 3}{2B} \left( \frac{\gamma_0^u}{\gamma_{\text{min}}^u} \right)^2 + O\left( \gamma^{-2} \right)
\]

\[
= 1 - \frac{1}{B} \frac{\gamma_0^u}{\gamma_{\text{min}}^u} + O\left( \gamma^{-1} \right)
\]

\[
P_{e,\text{RD}} = 1 - \frac{1}{B} \frac{\gamma_0^u}{\gamma_{\text{rd}}^u} + \frac{2B - 3}{2B} \left( \frac{\gamma_0^u}{\gamma_{\text{rd}}^u} \right)^2 + O\left( \gamma^{-2} \right)
\]

\[
= 1 - \frac{1}{B} \frac{\gamma_0^u}{\gamma_{\text{rd}}^u} + O\left( \gamma^{-1} \right),
\]

where $O\left( \gamma^{-1} \right)$ denotes the components associated with a diversity order higher than $k$. For a finite buffer size of $B$, the error probability is dominated by the first terms in (25) and (26), since they are associated with a diversity order of $D = 1$, showing that the buffer-aided system fails to achieve a diversity gain at high SNRs. However, compared to the traditional scheme associated with $B = 1$, the OP of the proposed protocol at high SNRs may be reduced by a factor as $B$ as seen below

\[
P_{\text{out}} = P_{e,\text{SR}} + P_{e,\text{RD}} = P_{e,\text{SR}}P_{e,\text{RD}}
\]

\[
= \frac{1}{B} \left( \frac{\gamma_0^u}{\gamma_{\text{min}}^u} + \frac{\gamma_0^u}{\gamma_{\text{rd}}^u} \right) + O\left( \gamma^{-1} \right).
\]

Therefore, as the buffer size $B$ increases, the outage performance improves by a factor of $B$. In the extreme case of $B = \infty$, the OP would no longer be dominated by the components associated with the diversity order of $D = 1$ at high SNRs, where a diversity order of $D = 2$ may be achieved as the lower bound, which is given by

\[
P_{\text{out,L}} = \left( \frac{\gamma_0^u}{\gamma_{\text{min}}^u} \right)^2 + \left( \frac{\gamma_0^u}{\gamma_{\text{rd}}^u} \right)^2 + O\left( \gamma^{-2} \right).
\]

### 4.4 e2e transmission delay

In conventional two-hop systems, the frame received by the RN may be readily retransmitted immediately in the following TS of the same frame, which results in a constant delay of two TSs. The outage performance improvement of the proposed buffer-aided relaying system arises from the selective diversity gain achieved by activating the better of the SN–RN hop or the RN–DN hop, relying on the RN’s ability to store the frames for a certain number of TSs, before retransmitting them to the DN.

For the buffer-aided two-hop system, the minimum delay for a TSs and the maximum delay for a frame is $2B$ TSs.

The minimum delay corresponds to the scenario, when the frame is received by the RN in a TS and then immediately transmitted to the DN in the following TS. On the other hand, the maximum delay of $2B$ TSs is related to the specific case, when a frame is received by the RN in TS 1 and hence it is inserted at the bottom of the buffer. Then it waits for $(B - 1)$ TSs for all the previous received frames stored already in the buffer to be transmitted, followed by another $(B - 1)$ TSs for $(B - 1)$ newly received frames to be inserted in the buffer, until the frame reaches the top of the full buffer. Then, it is transmitted to the DN using another single TS.

Even though a certain frame may experience a longer delay than that in the conventional two-hop system, the overall delay for a block of $M$ frames is $2M$, which is identical to that in the conventional system. This is because in a certain TS, a frame is moving forward by one hop, either on the SN–RN or the RN–DN hops. Let us now consider delay of the frame. According to Little’s law [17], the average time duration that a frame is stored in the buffer is given by

\[
d = \frac{b}{A}, \tag{29}
\]

where $A = 0.5$ is the average arrival rate experienced in terms of frames per TS, while $b$ is the average queue length, which can be expressed as

\[
b = \sum_{i=0}^{B} i P(r = i), \tag{30}
\]

Upon solving (22), we arrive at $P(r = B) = P(r = 0) = 1/2B$ and $P(r = i) = 1/B, \forall i \in [1, B - 1]$. By substituting these probabilities into (30), we arrive at $b = B/2$ and the average time duration that a frame is stored in the buffer therefore becomes $B$ TSs. While one TS is required for transmitting the frame from the RN to the DN, the average frame delay becomes $(B + 1)$ TSs, indicating that the average frame delay increases linearly with the buffer size $B$.

As derived in Section 4.2, when the buffer size $B$ increases, the e2e outage probability decreases by a factor of $B$. Therefore, the outage-versus-delay trade-off should be considered in a practical design. In the following section, we aim for improving the e2e outage probability by finding the optimal power allocation among the SNs and the RN. By fixing the buffer size, the optimal power allocation improves the e2e outage performance without increasing the delay. On the other hand, the optimal power allocation may be capable of reducing the delay, while maintaining the same e2e outage performance, as an equal-power allocation scheme.

### 5 Optimal power allocation

In this section, we design a power allocation algorithm for minimising the e2e outage probability of the proposed buffer-aided two-hop system subject to the constraint of a total power budget given by $P_m = (M + 1)P_m$, where $P_m$ is the average power assigned to a transmit node among the SNs and the RN. The maximum power per transmit node is given by $P_{\text{max}} = KP_m$, where $K$ satisfies $1 \leq K \leq (M + 1)$. The problem is formulated as the following maximisation problem

\[
\max \ 1 - P_{\text{out}}
\]

\[\text{s.t. } \sum_{m=1}^{M} P_m + P_r = P_A \tag{31}\]

\[P_m \leq P_{\text{max}}, \quad m = 1, \ldots, M, \quad P_r \leq P_{\text{max}}, \]
where \( P_1, P_2, \ldots, P_M \) denotes the transmitted power of the SNs, while \( P_d \) represents the transmitted power of the RN. The average SNR of the link between nodes \( i \) and \( j \) is \( \gamma_{ij} = G_{ij} P_d \), where \( G_{ij} = (N_0 \times dB)^{-1} \) captures the effect of both the pathloss and of the noise power.

Here, we adopt the Lagrange multiplier-based maximisation method by ignoring the constraints of the maximum power per transmit node [18], where the Lagrange function can be expressed as

\[
L = (1 - P_{\text{out}}) + \lambda \left( P_d - \sum_{m=1}^{M} P_m - P_r \right),
\]

By substituting the exact formulation of \( P_m \) from (24) into (32) and setting the partial derivative of \( L \) to 0 with respect to \( P_m \) and \( P_r \), the optimal power allocation can be obtained as

\[
\begin{align*}
P^*_m &= P_d - \frac{K(P^*_d, P^*_r) G_d^{-1/2}}{\sum_{m=1}^{M} G_m^{-1/2} + K(P^*_d, P^*_r) G_d^{-1/2}}, \\
P^*_r &= (P_d - P^*_m) G_d^{-1/2} \sum_{m=1}^{M} G_m^{-1/2}, \\
\end{align*}
\]

where \( K(P^*_d, P^*_r) \) is defined as (see (34)).

The solution in (34) is not in closed form. However, it may be solved by the successive approximation method using the iteration of \( [\hat{P}^{k+1}_d, \ldots, \hat{P}^{k+1}_d, \hat{P}^{k+1}_r] = f([\hat{P}^{k+1}_d, \ldots, \hat{P}^{k+1}_d, \hat{P}^{k+1}_d]) \), where \( f(\cdot) \) is defined by the right-hand side of (34) and the initial solution in iteration \( k = 1 \) is assumed to be a uniform allocation of \( P_d = P_1 = \ldots = P_M = P_r \). If the solution of (34) violates the constraints of the maximum affordable power per node, clipping is applied by assigning the maximum power \( P_{\text{max}} \) to each node in the set \( N_v \) comprising the violating nodes. Then we re-optimise the modified problem by changing the first constraint in (31) as \( \sum_{v \in N_v} P_m = P_d - |N_v| P_{\text{max}} \).

The above power allocation method is based on the exact OP. However, in practice the asymptotic OP-based power allocation is simpler and may provide more insights. Specifically, by substituting the asymptotic OP of (27) that neglects the \( O(\gamma^{-1}) \) terms in high SNRs, the Lagrange function in (32) becomes (see (35))

By taking the partial derivatives of \( L \), we obtain

\[
\frac{\partial L}{\partial P_m} = \frac{\gamma_{th}}{BG_{\text{aw}}} P_m^{-2} - \lambda = 0
\]

\[
\frac{\partial L}{\partial P_r} = \frac{\gamma_{th}}{BG_{\text{aw}}} P_r^{-2} - \lambda = 0.
\]

Then, since we have \( \sum_{m=1}^{M} P_m + P_r = P_d \), the optimal \( P_m \) and \( P_r \)

\[
K(P^*_d, P^*_r) = \frac{(2B - 1) \exp(-\gamma_{th}/\gamma_{th}^{\text{out}})}{(2B - 1) \exp(-\gamma_{th}/\gamma_{th}^{\text{out}}) - 2(\gamma_{th}/\gamma_{th}^{\text{out}}) \gamma_{th}^{\text{out}}},
\]

\[
L = 1 - \frac{1}{B} \left( \frac{\gamma_{th}^{\text{out}} + \gamma_{th}^{\text{out}}}{\gamma_{th}^{\text{out}}} \right) + \lambda \left( P_d - \sum_{m=1}^{M} P_m - P_r \right),
\]

\[
= 1 - \frac{1}{B} \left( \gamma_{th}^{\text{out}} \sum_{m=1}^{M} P_m G_m^{-1} + \gamma_{th}^{\text{out}} P_r G_d^{-1} \right) + \lambda \left( P_d - \sum_{m=1}^{M} P_m - P_r \right).
\]
The impact of the number of SNs

6.2 Impact of the number of users

The impact of the number of SNs $M$ on the e2e outage performance is illustrated in Fig. 4. In the examples shown in Fig. 4, we assume $d_{a}=80$ m and $d_{d}=120$ m, while the transmission rate of each SN is $R=1$ bps/Hz.

As the number of SNs $M$ increases, it is observed in Fig. 4 that the outage performance degrades. The reasons for the performance degradation are two-fold. First, the average minimum SNR $\eta_{a}=\frac{\sum_{i=1}^{M}R_{i}}{M}$ on the SN–RN hop decreases, which results in a higher error probability over the SN–RN hop, as derived in (23). Second, as the number of SNs increases, a higher rate of MR is achieved over the RN–DN hop, which results in a higher decoding threshold of $\Delta_{dn}=2^{M}−1$ and a higher error probability over the RN–DN hop.

It is also observed in Fig. 4 that the proposed MU-BR-UL protocol using $B>1$ improves the outage performance, regardless of the number of SNs involved. In the example illustrated in Fig. 4, an SNR gain of approximately 8 dB may be achieved by the proposed protocol with a buffer size $B=8$, when compared to a conventional two-hop system having a buffer size $B=1$ for a system supporting $M=4$ users.

6.3 Impact of the relay position

Let us now investigate the impact of the RN position on the outage performance. In the example illustrated in Fig. 5, we assume that the SNs, the RN and the DN are placed in a straight line and the distance between a SN and the DN is $d_{rd}=200$ m. In order to quantify the position of the RN, we introduce the normalised position of $X_{sr}=(d_{sr}/d_{rd})\in[0,1]$. The transmission rate of each SN is $R=1$ bps/Hz and the transmission power of each node is $P_{t}=10$ dBm, while the buffer size at the RN is $B=8$.

Observe in Fig. 5 that when the number of SNs increases, the e2e outage performance degrades, as also seen in Fig. 4. The optimal RN position quantified in terms of the minimum e2e outage performance moves from the SNs closer to the DN, because a higher rate of MR is transmitted over the RN–DN hop. Therefore, from a network design...
We opted for activating both hops with an equal probability of to balance the data traffic between the two-hop system and the optimal RN position for \( M = 1 \) is exactly in the middle of the SN–DN. Furthermore, in the example of Fig. 5, the buffer size is set to \( B = 8 \), albeit an increased buffer size may further improve the achievable outage performance as shown in the results of Fig. 2.

The impact of power allocation is illustrated in Fig. 6 for different RN positions, where it is shown that for arbitrary RN positions, the optimal power allocation scheme outperforms its equal-power allocation counterpart. Specifically, it is observed in Fig. 6 that the outage performance is less sensitive to the RN positions with the assistance of the optimal power allocation, while the outage improvement of optimal power allocation over uniform power allocation is more significant, when the RN is closer to the SNs or when the number of SNs increases. It is also observed that for specific RN positions, the outage performance results of the optimal power allocation and of the equal-power allocation are similar, because in these cases, the solutions of the optimal power allocation regime are close to the equal-power allocation.

7 Conclusions

In this paper, we have proposed and investigated a multi-user buffer-aided-relaying uplink protocol conceived for two-hop systems. With the aid of the min-SNR approximation technique, we analysed the \( e_2e \) outage probability and the transmission delay of the proposed protocol. Our analysis and simulation results showed that by exploiting the selection diversity of the SN–RN and the RN–DN hops, significant outage improvements may be achieved compared to the conventional systems, albeit this is attained at the expense of an increased \( e_2e \) delay. The optimal power allocation scheme was also conceived, which achieved a better \( e_2e \) outage performance than the equal-power allocation scheme.

In this paper, we considered an identical SN rate scenario and an identical TS allocation scheme, where the SNs and the RN are allocated an identical TS-duration. Hence, the RN should re-encode the messages with a rate of MR. Furthermore, in order to balance the data traffic on the SN–RN and the RN–DN hops, we opted for activating both hops with an equal probability of 50%. However, the above identical TS allocation scheme and the hop-activation scheme are not optimal. An interesting problem that requires further investigation is constituted by a joint TS-activation strategy, where the RN may choose different coding rates. The joint TS-activation strategy is also a challenging problem, which may exploit the OP analysis proposed in this paper and would be investigated in our future works.

8 Acknowledgments

The financial support of the RCUK under the auspices of the India-UK Advanced Technology Centre (IU-ATC), of the EU’s Concerto project as well as that of the European Research Council’s Advanced Fellow Grant and that of the Royal Society’s Wolfson Research Merit Award was gratefully acknowledged.

9 References

Author Queries
Bo Zhang, Chen Dong, Jing Lei, Mohammed El-Hajjar, Lie-Liang Yang, Lajos Hanzo

Q1 All footnotes have been moved to text as required by journal style. Please check and confirm that they have been located correctly within the text.

Q2 Please update refs. [3 and 11]

Q3 Please provide author names for refs. [8 and 9].

Q4 Please check the e-mail id of the corresponding author.